Advanced Data Processing
in Super-resolution Microscopy

vorgelegt von
Simon Christoph Stein aus Gehrden, Deutschland

Advanced Data Processing in Super-resolution Microscopy

SUBMITTED BY
SIMON CHRISTOPH STEIN FROM GEHRDEN, GERMANY

THESIS COMMITTEE:

Prof. Dr. Jörg Enderlein (1st Referee)

Third Intsitute of Physics - Biophysics
Georg-August-University Göttingen

Prof. Dr. Holger Stark (2nd Referee)

Department of Structural Dynamics
Max Planck Institute for Biophysical Chemistry

Prof. Dr. Reiner Kree

Institute of Theoretical Physics
Georg-August-University Göttingen

OTHER MEMBERS OF THE EXAMINATION BOARD:

Dr. Alexander Egner

Laser Laboratory Göttingen

Dr. Michael Habeck

Institute for Mathematical Stochastics
Georg-August-University Göttingen

Dr. Florian Rehfeldt

Third Intitute of Physics - Biophysics
Georg-August-University Göttingen

DATE OF ORAL EXAMINATION: 14th August 2017

Advanced Data Processing
in Super-resolution Microscopy

ABSTRACT

Super-resolution fluorescence techniques enable the study of structures smaller than the diffraction limit with visible light microscopy. Their introduction and development within the last two decades opened up the possibility to ask entirely new questions in cell biology. The field of super-resolution microscopy is still growing rapidly, and many improvements and novel methods have been proposed in recent years.

The major part of this thesis revolves around the advancement of Super-resolution Optical Fluctuation Imaging (SOFI), a fairly new technique which enhances the spatial resolution of an image by evaluating the temporal fluctuations of blinking fluorescent emitters. SOFI enables optical sectioning with wide-field microscopes and is compatible with a large range of experimental conditions. The following SOFI-related contributions are presented: First, a comprehensive analysis of the convergence properties of SOFI and its dependence on experimental parameters is shown, including an estimation of the necessary recording time and particle density for a targeted resolution enhancement. Second, we present *Fourier SOFI*, a new approach to generate super-resolved images on a finer pixel grid than the original camera recording. In contrast to established algorithms relying on spatial cross-cumulants, this method is practically free of artifacts and does not require any post-processing corrections. Next, an algorithm is laid out that corrects for the contributions of noise in zero-time-lag SOFI images. This extends the applicability of auto-cumulant SOFI – which correlates values only in time, not in space – to recordings where the time scale of photoblinking is on the same order as the exposure time. We also show how the lateral microscope Point Spread Function (PSF) can be estimated from SOFI data of thin samples. If multiple focal planes are imaged at the same time, the full three-dimensional PSF can be recovered. The fifth contribution revolves around *Fourier Preweighting*, a method that increases the resolution of SOFI images by pre-processing the original data. We show that this outperforms current techniques and surprisingly also improves the density dependence of SOFI. A related algorithm is presented which automatically matches the degree of resolution enhancement to the data quality,

avoiding artifacts. The last SOFI contribution demonstrates its applicability to cells labeled with carbon nanodots, a fairly new cost-effective and bio-compatible class of fluorophores, and shows that their blinking behavior is qualitatively similar to that of fluorescing semiconductor crystals (quantum dots).

Unrelated to SOFI, we show how single Atto647N molecules can be localized with sub-nanometer precision by exploiting the increased photon yield of samples at liquid nitrogen temperature. We propose a method to resolve spectrally identical, non-blinking fluorophores spaced only few nanometers apart with a special cryo-fluorescence setup and validate our concept with simulations.

We also developed *TrackNTrace*, an open-source MATLAB framework to support the development of fluorescence imaging applications. Its design is focused on easy extensibility through plugins, simplicity of coding, and rich visual feedback. We demonstrate competitive performance and execution speed in Single-Molecule Localization Microscopy applications compared to established software and include many state-of-the-art algorithms out-of-the-box.

Finally, we present two projects reliant on single-molecule imaging: First, we developed a model for the intensity distribution of fluorescent molecules imaged while flowing through a nanochannel and verified it experimentally. This can be used to extract the ratios of differently labeled species from a measurement of their mixture. In the second project, we demonstrate the first simultaneous measurement of the excitation and emission dipole axes of single molecules.

Contents

List of Acronyms

AC-SOFI Auto-Cumulant SOFI

CND Carbon Nanodot

DFT Discrete Fourier Transform

EMCCD Electron Multiplying Charge-Coupled Device

fSNR Frequency Signal-to-Noise Ratio

fSOFI Fourier Super-resolution Optical Fluctuation Imaging

GUI Graphical User Interface

ISM Image Scanning Microscopy

MLE Maximum Likelihood Estimation

NSIM Nonlinear Structured Illumination Microscopy

OTF Optical Transfer Function

PALM Photoactivated Localization Microscopy

PSF Point Spread Function

RMS Root Mean Square (Error)

SMLM Single-Molecule Localization Microscopy

SNR Signal-to-Noise Ratio

SOFI Super-resolution Optical Fluctuation Imaging

SPT Single Particle Tracking

STED Stimulated Emission Depletion

STORM Stochastic Optical Reconstruction Microscopy

TIRF Total Internal Reflection Fluorescence

TNT TrackNTrace

XC-SOFI Cross-Cumulant SOFI

Affidavit

Hereby, I declare that the presented thesis has been written independently and with no other sources and aids than quoted.

List of Related Publications

Stein, S.C.; Huss, A.; Hähnel, D.; Gregor, I.; Enderlein, J.; "Fourier interpolation stochastic optical fluctuation imaging", Optics Express, 23, 16154-16163, 2015.

Stein, S.C.; Thiart, J.; "TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms.", Scientific Reports, 6, 37937, 2016.

Weixing, L.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Enderlein, J.; "Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy", Optics Express, 23, 3770-3783, 2015.

Chizhik, A.M.; **Stein, S.C.**; Dekaliuk, M.O.; Battle, C.; Li, W.; Huss, A.; Platen, M., Schaap, I.A.T.; Gregor, I.; Demchenko, A.P.; Schmidt, C.F., Enderlein, J.; Chizhik, A. I.; "Super-Resolution Optical Fluctuation Bio-Imaging with Dual-Color Carbon Nanodots", Nano Letters, 16, 237-242, 2016.

Karedla, N.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Chizhik, A.; Enderlein, J.; "Simultaneous Measurement of the Three-Dimensional Orientation of Excitation and Emission Dipoles", Physical Review Letters, 115, 173002, 2015.

Isbaner, S.; Karedla, N., Ruhlandt, D.; **Stein, S.C.**; Chizhik, A.; Gregor, I.; Enderlein, J.; "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging", Optics Express, 24, 9429-9445, 2016.

1

Introduction

The development of microscopy up to this day is driven by a fascination with the workings of nature on length scales too small to be observed by the human eye. Microscopes exposed this previously invisible microcosm, leading to countless insights in a diverse range of fields, especially biology, medicine, and material sciences. For a long time, the desire to study ever smaller objects was opposed by the limited quality of every microscope's core optical instrument: the lens.

The earliest depictions of lenses are found in Egyptian hieroglyphs dating back to 800 B.C.[1]. Back then, they were mostly known for their ability to ignite fires by concentrating sunlight, giving them the name "burning glasses". Only a fortunate few realized their capability to magnify objects and used this fact to facilitate the creation of fine engravings and seals [2]. However, missing knowledge about what determines a lens' optical quality prevented its widespread application. In spite of this, some craftsmen discovered early on how to fabricate lenses with impressive imaging properties, such as the *Visby lenses* made by Vikings in the 11th or 12th century, the best of which come close in quality to aspherical lenses produced today [3]. In the 13th century, the art of lens making became common practice through the invention of spectacles, but was still entirely based on empirical knowledge [4]. Unknown in the western world, the Persian scientist Ibn Sahl was responsible for several profound advancements in optical theory. In 984 he derived both the law of refraction and the shape of an ideal lens for the first time [5], a knowledge not rediscovered until the 17th century. Around this time, the first compound microscopes carrying multiple lenses were developed, which allowed for more detailed observations through higher magnification. The next leap in image fidelity was made possible by the invention of achromatic systems, which correct the color-dependency of a microscope's magnification. Towards the end of the 19th century, new illumination methods could provide even sample lighting, improving image clarity and contrast by removing the previously visible afterimage of the light source [6].

At this point, the best optical microscopes could resolve such small structures that two fundamental problems were brought into focus: First, the absorption of light vanishes with decreasing size of the specimen, making direct observation impossible. This issue was resolved by new ways to create image contrast, such as phase contrast microscopy, which converts small variations in the traveling time of light through the sample into perceptible brightness differences. Its inventor Frits Zernike was awarded the Nobel Prize in Physics in 1953. One of the most important techniques for biological microscopy today, and also the basis of this thesis, is fluorescence microscopy. Here, the sample is labeled with molecules, which are brought to an excited state by absorbing light of a suitable wavelength. After losing a certain amount of energy, they return to the ground state, emitting a photon of higher wavelength. This so-called *Stokes shift* makes it possible to separate emission from excitation, yielding images with high contrast. Another advantage of fluorescence microscopy is its specificity, as different binding sites on molecules can be precisely targeted by modern labeling techniques.

The second problem was brought to attention by Ernst Abbe in 1883, who could prove that the resolution of an optical microscope is universally limited due to the wave properties of light [7]. Owing to the finite wavelength and the phenomenon of diffraction, even a point-like light source appears as a spot of finite width in the image, causing a loss of details on length scales below the spot size. In practice, this prevents conventional far-field light microscopes from resolving features smaller than 200 nanometers in size. The resolution limit was unchallenged for over 100 years, but finally circumvented around the turn of the millennium by the invention of two fluorescence microscopy techniques: Stimulated Emission Depletion (STED) microscopy and Single-Molecule Localization Microscopy (SMLM). The former exploits the phenomenon of stimulated emission [8], while the latter relies on single-molecule sensitive imaging and the ability to control the on/off switching behavior of fluorophores [9, 10]. In 2014, Stefan W. Hell, Eric Betzig and William E. Moerner received the Nobel Prize in Chemistry for enabling this breakthrough. Since the introduction of super-resolution microscopy, a plethora of different techniques has been invented, many of which can, in principle, achieve unlimited resolution. However, they can usually be applied only within a narrow spectrum of experimental conditions, limiting their performance in practice. As the field of super-resolution is quite new, there is still considerable room for improvement, which is demonstrated by its rapid development in recent years.

In consequence, the continuous improvement of fluorescence microscopy today is not tied to the refinement of optical components anymore, but the development of clever new experimental techniques, better ways to control the fluorophore behavior and advanced data processing methods, the latter being the focus of this work.

This thesis presents improvements for different super-resolution and single-molecule applications, with a focus on Super-resolution Optical Fluctuation Imaging (SOFI, Chapter 2). First published in 2009 [11], SOFI is one of the newest super-resolution techniques. Its main advantages are its simplicity – requiring only a standard wide-field microscope with a fast recording camera – and compatibility with a wide range of experimental conditions. Chapter 3 demonstrates that single molecules can be localized with sub-nanometer precision when imaging samples that are cooled to liquid nitrogen temperatures. Additionally, a concept for the accurate co-localization of multiple fluorophores is proposed

and discussed. Chapter 4 introduces TrackNTrace (TNT), a software I developed with my colleague Jan Thiart to facilitate the development of new data evaluation methods for fluorescence imaging applications, with a focus on Single-Molecule Localization Microscopy and Single Particle Tracking (SPT). Chapter 5 deals with two novel single-molecule applications: The first section presents a detailed analysis of the brightness distribution of single-molecules flowing through nanochannels. This precise characterization is useful to determine the ratio of differently labeled samples when imaging a mixture. The second section presents an imaging-based method for the simultaneous measurement of the excitation and emission dipole of single fluorophores. Finally, Chapter 6 summarizes all findings and concludes the thesis.

In addition to this general introduction, each chapter features a brief overview of the topic and current developments. Most sections end with a list of bullet points, summarizing the most important aspects to improve comprehensibility and facilitate quick look-up of information. The following sections explain the basics of fluorescence and super-resolution microscopy, which are relevant to all chapters.
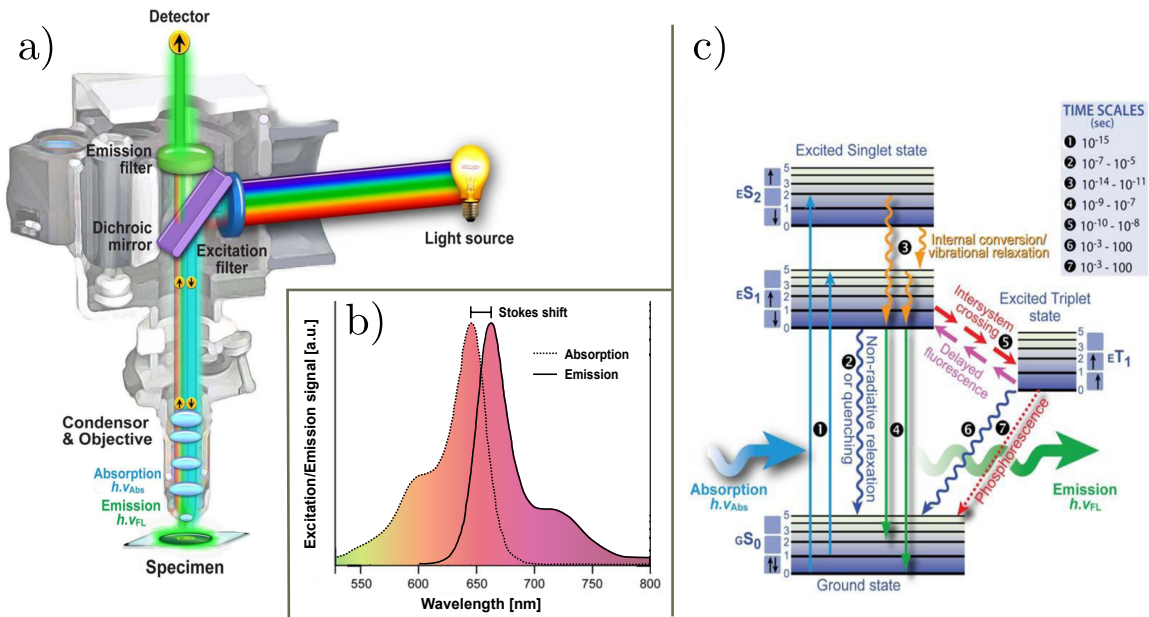
Figure 1.1: Fundamentals of fluorescence microscopy. a) Scheme of a typical fluorescence microscope. b) Jablonski diagram illustrating the electronic states of a typical fluorophore and the transitions between them. Additionally, the time-scales of these transitions are depicted. Small black arrows indicate the spin state of electrons. c) Absorption and emission spectra of the fluorophore BODIPY[650/665], showing a stokes shift of ≈18 nm. [Original images from Ishikawa-Ankerhold *et al* [12], re-edited for this figure; Creative Commons Attribution license 3.0 (CC BY 3.0): https://creativecommons.org/licenses/by/3.0/]

In fluorescence microscopy, special molecules, called fluorophores, are attached to points of interest in the sample and subsequently imaged with a customized microscope, as illustrated in Figure 1.1 a). First, the excitation light travels through an optical filter that selects a specific wavelength. Next, it is reflected by a dichroic mirror, before the objective focuses it into the specimen. A dichroic mirrors is transparent for a range of wavelengths while reflecting others. The fluorophores absorb a certain spectrum of the incoming light, elevating the molecules into an excited electronic state. They can return to the original ground state by emission of a photon of higher wavelength, leading to a color shift between excitation and emission. As an example, Figure 1.1 b) shows the absorption and emission spectrum of the fluorophore BODIPY[650/665]. Due to this wavelength shift, the *Stokes shift*, the emitted light can be separated from the excitation, passing the dichroic mirror on the return path. An additional emission filter further reduces unwanted background light, before the signal is collected by the detector (e.g. a camera or avalanche photodiode). As only structures of interest are labeled, the recorded signal exhibits high contrast and specificity.

With the exception of Section 5.2, the experimental data in this thesis was recorded with a *wide-field microscope*, the most basic type of fluorescence microscope. Here, the sample is homogeneously illuminated and the emission is detected by focusing the collected light on a camera. In contrast, Section 5.2 also makes use of a *confocal laser scanning microscope*, which scans the sample with a focused laser beam and collects the emission on a single-pixel detector. The final image is assembled only later in

post-processing by combing the measurements at all scan positions.

Figure 1.1 c) shows the Jablonski diagram – a scheme illustrating a molecule's electronic states and the transitions between them – of a typical fluorophore. Absorption of the excitation light pumps the molecule from the ground state GS_0 into a high energy state, followed by a quick relaxation into the lowest excited singlet state ES_1 through intramolecular energy conversion or intermolecular energy transfer. This energy loss is responsible for the Stokes shift between excitation and emission. Subsequently, the detected fluorescence photon is generated by a radiative transition from the first excited singlet state ES_1 to the ground state GS_0. The maximum rate of photons the fluorophore can emit is determined by the lifetime of this singlet state, which is typically on the order of nanoseconds, and is reduced when additional non-radiative relaxation channels are available. Additionally, the overall photon output of a molecule is limited by a chance of the excited fluorophore to bleach irreversibly, where it transitions into a new molecular structure that prohibits further fluorescence.

The transition to the excited triplet state is also of high relevance to fluorescence microscopy. This "intersystem crossing" requires a flip of the electron spin, changing the net spin of the system from zero to one. The spin flip has a low probability to occur, leading to a long lifetime of the triplet state. In this state, the molecule does not fluoresce and is thus not detectable by the microscope. The spontaneous transition into and out of the triplet state leads to a visible blinking of the fluorophore, a phenomenon that is exploited in certain super-resolution microscopy methods (see Section 1.3). The slow radiative relaxation out of the triplet state is called phosphorescence, a process that is deliberately used, for example, for the fabrication of special coatings for clock dials that glow in the dark after being "charged" by sunlight.

The two most commonly used fluorophores are organic dyes and fluorescent proteins. Organic dyes have a size of about 1 nm and are in general attached via immunolabeling. In this method, the fluorophore is attached to an antibody that binds with high specificity to a target antigen or, to another antibody inserted in a previous preparation step. Fluorescent proteins are incorporated via transfection, where the DNA of a living cell is manipulated, leading to the expression of the protein in a structure of interest. These dyes are slightly larger, with a size of about 3 nm to 5 nm. In the last 20 years, semiconductor quantum dots have been established as a new dye for cell biology. Quantum dots are comparatively huge, with sizes between 10 nm and 30 nm, but multiple times brighter and more photostable than the conventional dyes. Like organic dyes, they are attached using immunolabeling. Many types of quantum dots also display pronounced blinking, which is however not related to triplet state dynamics, but different mechanisms of electron charge transfer.

1.2 THE RESOLUTION LIMIT OF AN OPTICAL MICROSCOPE

In general, the emission of a fluorescent molecule is well described by an ideal electrical dipole[1]. Although the molecule has a finite size, its emission originates from a single point on its dipole axis; a

---

[1]Usually the molecules rotate fast and freely, and the emitted electromagnetic wave shows no preferred direction.
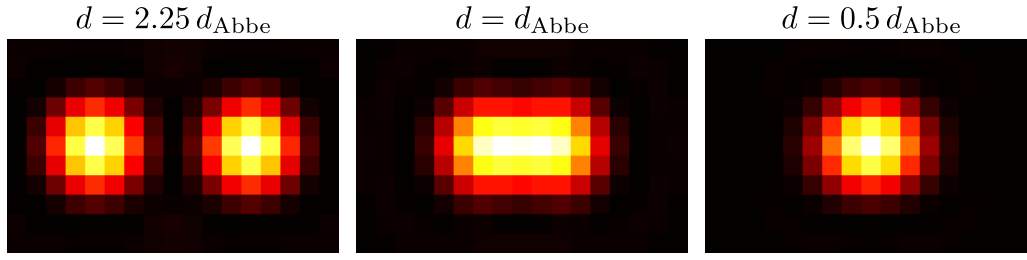
Figure 1.2: Two emitters spaced at different distances, illustrating the Abbe resolution limit.

fact that is also true for macroscopic antennas. Despite this, imaging this point source with an optical microscope spreads its intensity to a blurred spot of finite size. In consequence, two emitters at a small distance from each other are hard to distinguish (Figure 1.2). The best-known formula quantifying the minimum resolvable distance is called the "Abbe resolution limit" or "diffraction limit"

$$d_{\text{Abbe}} = \frac{\lambda_{\text{em}}}{2 \cdot n \sin \theta} = \frac{\lambda_{\text{em}}}{2 \cdot \text{NA}}, \tag{1.1}$$

where λ_{em} is the wavelength of the emitted light, n the refractive index of the medium between sample and objective[2] and θ the half-angle of the largest cone of light that can enter or exit the objective lens. Although the equation was first derived by others [13], its name-giver Ernst Abbe was the first to publish this formula together with matching experimental findings [7]. The numerical aperture (NA) is the most important measure to indicate the resolving power of an objective. Common NA values range from 0.8 for air immersion objectives to 1.4 for oil immersion. This corresponds to a resolution of about 310 nm and 180 nm respectively, when fluorophores with an emission wavelength of $\lambda_{\text{em}} = 500$ nm are imaged. Wavelengths shorter than visible light, such as UV or X-ray radiation, offer a better resolution, but require more expensive equipment, often provide less contrast in biological specimen and damage the sample more quickly.

The image of a point source is the microscope's Point Spread Function (PSF) $U(\mathbf{r})$. In fluorescence microscopy, the emission from different fluorophores is incoherent and their intensities simply add up in image space. In other words, the image formation process is linear in intensity. Furthermore, it is common to assume that the process is shift invariant, i.e. that the PSF does not change over the field of view[3]. In this case, the overall image is simply the sum of the images of all point sources, and the image formation can be expressed as

$$\text{image}(\mathbf{r}) = \text{sample}(\mathbf{r}) \circledast U(\mathbf{r}) = \int \text{sample}(\mathbf{r}') \cdot U(\mathbf{r} - \mathbf{r}') \, d\mathbf{r}', \tag{1.2}$$

where "$\circledast$" denotes convolution and the function "sample$(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$" is a set of delta distributions encoding the positions $\mathbf{r}_i$ of the fluorophores. Thus, the PSF provides a complete characterization of a shift-invariant linear imaging system.

---

[2]An objective must be used that is suitable for this refractive index.

[3]Due to lens imperfections, the PSF shows slight variations for emitters at different distances to the optical axis.

The PSF of a wide-field microscope is conceptually simple to compute, we follow the derivation from [14]: Any electric field for which the Fourier transform exists can be expressed as the sum of infinitely many plane waves

$$E(\mathbf{r}) = \mathcal{F}^{-1}\left[\tilde{E}(\mathbf{k})\right] = \int \tilde{E}(\mathbf{k})\exp(i\mathbf{k}\mathbf{r})\ \mathrm{d}\mathbf{k}, \tag{1.3}$$

where $\tilde{E}(\mathbf{k})$, the Fourier transform of $E(\mathbf{r})$, is the wave-vector dependent amplitude distribution. Here we use the complex number notation for the electromagnetic waves, keeping in mind that the actual physical waves are the real part of the expression, and we omit the factor $\exp(-i2\pi ft)$, which encodes their temporal evolution with the oscillation frequency f.

If we illuminate one side of the (objective) lens with a plane wave that propagates along the optical axis, the wavefront is bent into a spherical wave section that converges at the focal point, where we place the origin of our coordinate system. In terms of the plane wave decomposition, this spherical section can only consist of waves that travel within the cone of light that is defined by the opening angle θ of the objective. Thus, the amplitude distribution $\tilde{E}(\mathbf{k})$ is zero for wave vectors $\mathbf{k}$ outside this cone. The PSF is proportional to the intensity distribution of the electric field which would be observed when moving a screen along the light path. In turn, the intensity distribution is proportional to the magnitude of the complex electric field

$$U(\mathbf{r}) \propto E(\mathbf{r}) \cdot \overline{E(\mathbf{r})} = |E(\mathbf{r})|^2, \tag{1.4}$$

where $\overline{a + ib} = a - ib$ denotes the complex conjugate. Consequently, the PSF is computed by summing up all available plane waves and calculating the squared magnitude. The Fourier transform of the PSF, the Optical Transfer Function (OTF) $\tilde{U}(\mathbf{k})$, provides a different look at the image formation process. Using Eq. (1.2) and the convolution theorem, the Fourier transform of the image is given by

$$\mathcal{F}\left[\text{image}(\mathbf{r})\right] = \mathcal{F}\left[\text{sample}(\mathbf{r})\right] \cdot \mathcal{F}\left[U(\mathbf{r})\right] = \mathcal{F}\left[\text{sample}(\mathbf{r})\right] \cdot \tilde{U}(\mathbf{k}). \tag{1.5}$$

The OTF shows how the spatial frequencies of the sample structure are attenuated by the microscope and how the image formation manipulates their phase. Another use of the convolution theorem combined with Equations (1.3) and (1.4) shows that the OTF can be expressed as

$$\tilde{U}(\mathbf{k}) = \mathcal{F}\left[U(\mathbf{r})\right] = \mathcal{F}\left[E(\mathbf{r}) \cdot \overline{E(\mathbf{r})}\right] = \mathcal{F}\left[E(\mathbf{r})\right] \circledast \mathcal{F}\left[\overline{E(\mathbf{r})}\right] \tag{1.6}$$

$$= \tilde{E}(\mathbf{k}) \circledast \overline{\tilde{E}(-\mathbf{k})} = \tilde{E}(\mathbf{k}) \star \tilde{E}(\mathbf{k}) := \int \overline{\tilde{E}(\mathbf{k}')}\tilde{E}(\mathbf{k}' + \mathbf{k})\ \mathrm{d}\mathbf{k}', \tag{1.7}$$

where "$\star$" denotes correlation. That is, the OTF is the auto-correlation function of the amplitude distribution. The auto-correlation has a simple geometric meaning: it reports the integrated overlap of the original function with a shifted copy of itself as a function of the shift vector.

When imaging a monochromatic fluorophore, the wave vectors of all plane waves have the same magnitude $|\mathbf{k}| = 2\pi n/\lambda_{\text{em}}$. As mentioned, the possible wave vectors are limited by the opening angle of the

objective, and the function $\tilde{E}(\mathbf{k})$ is a spherical cap (section of a spherical surface) that is formed by all vectors of length $|\mathbf{k}|$ inside the cone that is collected by the objective. The wide-field OTF, which is the spatial auto-correlation of this spherical cap, and the resulting PSF are illustrated in Figure 1.3.
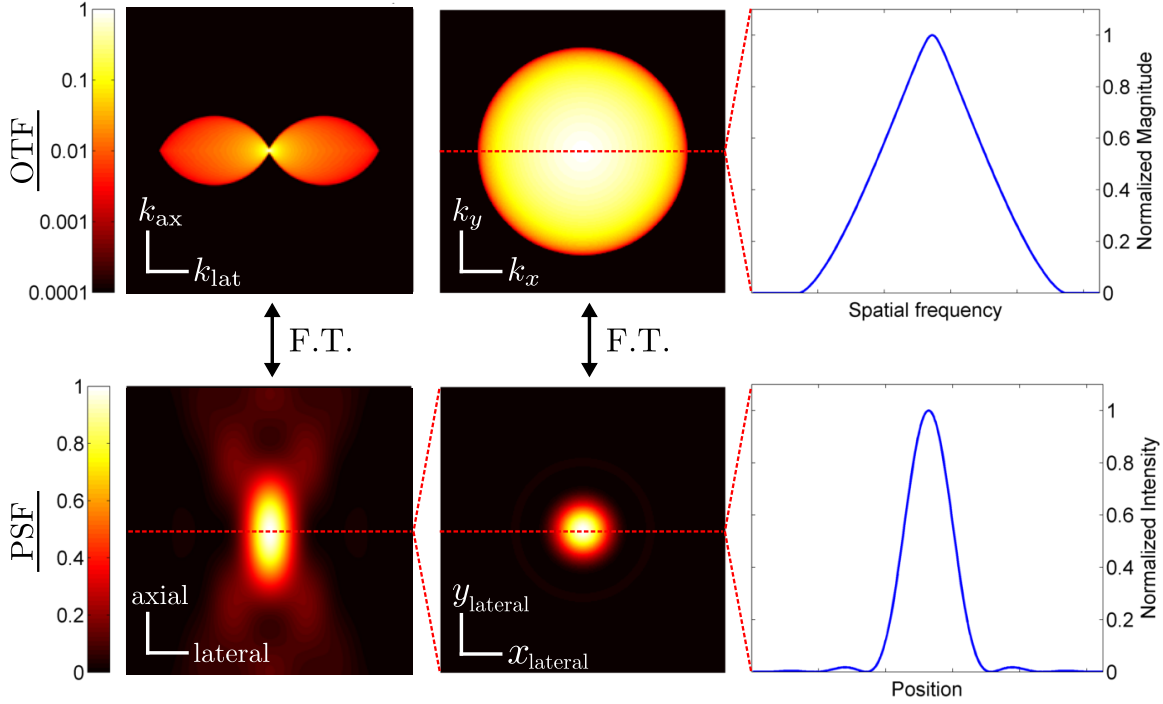
Figure 1.3: Optical Transfer Function (OTF) (frequency space, color-scale in decadic logarithm) and Point Spread Function (PSF) (real space) of an ideal wide-field microscope. The first column shows a cross-section of the 3D functions, which are radially symmetric around the central axial direction (optical axis). The second column shows the two-dimensional lateral PSF in the focal plane and its Fourier transform. The third column shows a one dimensional cross-section of both. F.T. denotes a relationship by the Fourier transform. Note that the Fourier transform of a cross-section is not equal to the cross-section of the Fourier transform, but equal to the projection of the Fourier transform orthogonal to the slice orientation. The OTF is zero outside of a finite support, limiting the optical resolution of the microscope. Due to the "missing cone" of frequencies in the axial direction, a wide-field microscope cannot resolve densely labeled three-dimensional structures, even when scanning the focal plane through the sample.

Due to the symmetry of the problem, the OTF and PSF are rotationally symmetric around the optical axis. The finite resolution of the microscope is coupled to the finite support of the Optical Transfer Function, as structures smaller than the wavelength of the highest available spatial frequency can not be encoded. This can also be understood from the image formation in Fourier space (Equation (1.5)), as the multiplication with the OTF removes all spatial frequencies beyond its support. A noticeable feature of the wide-field OTF is the "missing cone" of frequencies in the axial direction, which removes all spatial frequencies of the sample structure that are close to the optical axis. As a consequence, any densely labeled three-dimensional sample will be heavily blurred in the axial direction, prohibiting a three-dimensional reconstruction by shifting the focal plane. In contrast, this can be achieved with a confocal microscope – a capability called "optical sectioning" or "z-sectioning".

Note that even an infinite aperture microscope does not posses infinite resolution, as the smallest wavelength contributing to $|\tilde{E}(\mathbf{r})|^2$ is $\lambda_{em}/(2n)$, corresponding to the squared magnitude of waves that travel orthogonal to the optical axis. In general, it is not possible to create structures smaller than half

the wavelength of the superposed waves[4]. In this sense, the name "diffraction limit" is actually a bit confusing, as the common meaning of diffraction – the bending of light around corners into the regions of geometrical shadow of an obstacle – is not the major problem responsible for limiting the optical resolution of the microscope, but rather the wave-nature of light itself.

The main reason why a perfect image of the sample can not be reconstructed is the loss of near-field information. In addition to the light that propagates into the far-field, the near-field of a dipole also contains non-propagating components that decay exponentially with increasing distance from the source. In contrast to the fixed wavelength of the propagating electromagnetic wave, these components can carry arbitrarily high frequencies, a fact exploited in *near-field scanning optical microscopy* [16].

In most cases, 2D images from a sample in the focal plane are recorded. Here the two-dimensional cross-section through the PSF has the smallest width, corresponding to the highest lateral resolution. This two-dimensional in-focus PSF is given by the so called Airy pattern [17, 18]

$$
U_{\text{airy}}(\mathbf{r}_\perp) = \left[\frac{2 \cdot J_1 \left(\frac{2\pi}{\lambda_{\text{em}}} \text{NA} \cdot |\mathbf{r}_\perp| \right)}{\frac{2\pi}{\lambda_{\text{em}}} \text{NA} \cdot |\mathbf{r}_\perp|} \right]^2 , \tag{1.8}
$$

where $\mathbf{r}_\perp$ is the two-dimensional vector of the lateral coordinates and J_1 is the Bessel function of the first kind of order one. Its OTF is

$$
\tilde{U}_{\text{airy}}(\mathbf{k}_\perp) = \mathcal{F} \left(U_{\text{airy}} \right) = \text{circ} \left(\mathbf{k}_\perp; \frac{\text{NA}}{\lambda_{\text{em}}} \right) \circledast \text{circ} \left(\mathbf{k}_\perp; \frac{\text{NA}}{\lambda_{\text{em}}} \right) \tag{1.9}
$$

$$
\text{with circ} \left(\mathbf{x}; a \right) = \begin{cases} 1/ \left(\pi a^2 \right) & |\mathbf{x}| \leq a \\ 0 & |\mathbf{x}| > a \end{cases} , \tag{1.10}
$$

the auto-convolution of a circular area. For displacements larger than twice the radius during convolution, the circle and its shifted copy do not overlap anymore. Thus, the highest frequency of the OTF is $k_{\text{max}} = 2\text{NA}/\lambda_{\text{em}} = 1/d_{\text{Abbe}}$, the inverse of the resolution limit. Looking back at Equation (1.3), the electric field in the focal plane at $\mathbf{r} = (x, y, 0)$ can be expressed in the form

$$
E(\mathbf{r}_\perp) = \int \int \int \tilde{E}(\mathbf{k}) \exp(i(k_x x + k_y y)) \, dk_x \, dk_y \, dk_z \tag{1.11}
$$

$$
= \int \int \left[\int \tilde{E}(\mathbf{k}) \, dk_z \right] \exp(i(k_x x + k_y y)) dk_x \, dk_y \tag{1.12}
$$

$$
= \int \int \tilde{E}_\perp(k_x, k_y) \cdot \exp(i(k_x x + k_y y)) \, dk_x \, dk_y. \tag{1.13}
$$

If $\tilde{E}(\mathbf{k})$ is generated by a monochromatic spherical wave, its axial projection $\tilde{E}_\perp(k_x, k_y)$ represents the shape of the confining lens aperture. Thus, the two-dimensional in-focus OTF is the auto-convolution

---

[4]This is not strictly impossible however, as demonstrated by the curious phenomenon of *superoscillations*, which breaks the resolution limit in the far field [15]. Although this can be used to create sub-diffraction structures, they carry only a fraction of the energy of the diffraction limited structures that always surround them, which currently prohibits their application in practice.

of the aperture, and the form of Equation (1.9) reflects the circular shape of the lens. Equation (1.13) is identical to the Fraunhofer approximation of diffraction, in which the far-field of an illuminated obstacle is simply its Fourier transform. For this reason, the PSF is also called "diffraction pattern" (of the aperture).

Lastly, the Airy pattern $(Eq.(1.8))$ is commonly approximated by a 2D Gaussian function

$$U_{\text{gauss}}(\mathbf{r}_\perp) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{\mathbf{r}_\perp}{2\sigma^2}\right), \tag{1.14}$$

which is simpler to treat both theoretically and numerically, and shows only minor deviations from the true shape. In a least-squares sense, the best-fitting width of the Gaussian is given by [17]

$$\sigma = \frac{1}{2\cdot(2\sqrt{2\log 2})}\frac{\lambda_{\text{em}}}{\text{NA}} \approx 0.21\frac{\lambda_{\text{em}}}{\text{NA}}, \tag{1.15}$$

which yields a *full width at half maximum* (FWHM) that is identical to the resolution limit

$$\text{FWHM} = 2\sqrt{2\log 2}\cdot\sigma = \frac{\lambda_{\text{em}}}{2\cdot\text{NA}} = d_{\text{Abbe}}. \tag{1.16}$$

1.3   Super-resolution microscopy

The classical resolution limit was derived for a static situation, where the microscope essentially takes a snap-shot of the whole scene at one instant in time. Super-resolution techniques circumvent the diffraction limit by making use of time-dependent processes, trading time-resolution against spatial resolution. Currently three major categories can be identified:

1. **Deterministic linear-excitation-response:** A temporally varying intensity pattern is used to excite the sample. The emission intensity is still proportional to the excitation strength. Although these methods achieve better resolution than the classical diffraction limit, there is an upper bound of the lateral resolution, which is twice that of an ideal wide-field microscope. Thus, they do not provide "true" super-resolution.

2. **Deterministic nonlinear-excitation-response:** Like before, a temporally varying excitation pattern is used, but the sample has a nonlinear excitation response. The achievable resolution is, in principle, unbounded.

3. **Stochastic sample behavior:** These methods utilize labels which exhibit some kind of stochastic temporal behavior with regard to their emission. Again, there is no theoretical upper bound for the resolution that can be achieved.

Although only methods from the third category are applied in this thesis, we give a brief overview over all categories. Technically, the widely-used *confocal laser scanning microscope* is an example for the first category. Here the sample is scanned with a focused laser beam, and the emission is refocused through

a small pinhole before being concentrated on a single-pixel detector. For an infinitely small pinhole, its OTF is approximately the auto-convolution of the wide-field OTF, yielding a doubled resolution [14]. However, very small pinhole sizes are associated with losing the major part of the signal, resulting in a low Signal-to-Noise Ratio (SNR). The moderate pinhole sizes used in practice provide only a minor resolution improvement, but efficiently reject background light, enabling optical sectioning. *Image Scanning Microscopy* (ISM) works in a similar way, but uses a camera instead of a single-pixel detector, which records an image at every scan position [19]. In this approach, each camera pixel is treated as a small pinhole, which allows a doubled resolution image to be computed from the image series in postprocessing. Another technique of this category is *Structured Illumination Microscopy* (SIM)[20], where a high-frequency standing wave excitation pattern is generated and scanned across the sample along different directions. At each excitation state an image is taken, and the output image is reconstructed from the recorded set once the measurement is finished.

The most prominent technique of the second category is *Stimulated Emission Depletion* (STED) [8], which was the first true super-resolution method. In STED the sample is scanned with two spots simultaneously or pulsed in short succession: a normal focused excitation spot is aligned with a donut-shaped "depletion beam" (or STED beam), which is usually generated by passing the laser beam through a spiral phase plate. The depletion beam is driven with high power and its wavelength is commonly chosen to be in the upper tail of the emission spectrum of the utilized dye. After the sample is excited by the diffraction limited excitation spot, the STED beam induces stimulated emission inside its donut profile, leaving only molecules inside the "hole" in the donut center in the excited state, from which they can fluoresce. The effective hole size is not diffraction limited and depends on the intensity of the depletion beam. Photons from stimulated and spontaneous emission can be separated based on their wavelength, using an optical filter. Alternatively, in time-gated STED, they can also be distinguished by their arrival times at the detector, as stimulated emission is the faster process, leaving only the spontaneously emitted photons after a certain time. Higher depletion beam intensities increase the resolution, but increasingly damage the sample. While resolution in the nanometer range has been demonstrated using highly-photostable fluorescent nitrogen vacancies in diamonds [21], typical values in biological system are in the 50 nm range [22]. The second technique of this category is *Nonlinear Structured Illumination Microscopy* (NSIM), which is in principle the same as normal SIM, but at such high excitation intensities that a part of the fluorophores becomes saturated [23]. The emission of the saturated molecules can not be further boosted by increasing the excitation power, which effectively deforms the crests of the originally sine-like excitation pattern to plateaus. This allows reconstructions with resolution around 50 nm in practice [24].

The third category revolves around two methods that are important for this thesis. The first one is *Single-Molecule Localization Microscopy* (SMLM), which is illustrated in Figure 1.4. Here the photoblinking of the fluorophores is tuned in a precise way, such that only a small subset of them is fluorescing in every frame of a recorded movie. The tuning is mostly achieved by manipulating the chemical environment of the fluorophores, or by using dyes with photosensitive behavior, which can be controlled via excitation with additional wavelengths. A plethora of techniques for this has been invented, the first
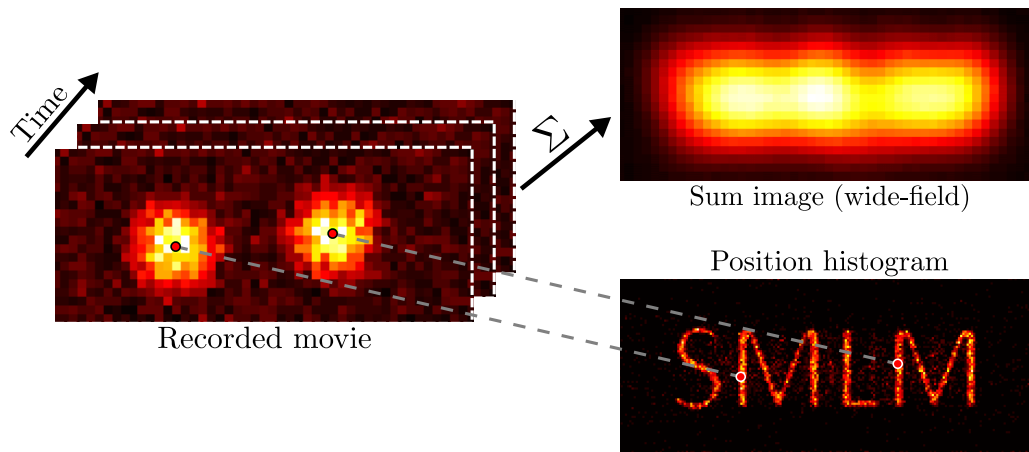
Figure 1.4: In Single-Molecule Localization Microscopy (SMLM), a movie of blinking fluorophores is recorded and the fluorescence intermittency is tuned in such a way that in each frame (statistically) only one molecule is active per diffraction limited area. In this case, the center positions of the single-molecule patterns can be localized with much higher precision than their width, allowing sub-diffraction structures to be captured. The example shows a simulation that was evaluated with TrackNTrace (see Chap. 4).

and most prominent being Photoactivated Localization Microscopy (PALM) [10] and Stochastic Optical Reconstruction Microscopy (STORM) [25], but they all follow the same concept: The center of an isolated diffraction limited spot can be extracted with much higher precision than its width. If the dark state occupation time of the fluorophores is high, the chance to observe two molecules in a diffraction limited area is low. After localizing the subsequently activated spots in each frame, a super-resolved image can be assembled from the gathered positions (e.g. using a position histogram in the simplest case). The precision of the localization – and thus the optical resolution – increases with the number of photons that can be collected from an emitter during the exposure time [26]. In practice, typical resolutions are again in the 50 nm range [22]. SMLM techniques are applied in Chapters 3, 4 and 5. A conceptually different method is *Super-resolution Optical Fluctuation Imaging* (SOFI), which exploits the fact that the photoblinking of different emitters is uncorrelated and computes high resolution images by correlating the intensity fluctuations in recorded movies. SOFI requires no special hardware, only a normal wide-field microscope and a fast recording camera, and is compatible with a wide range of experimental conditions [27]. SOFI also enhances wide-field microscopes with optical sectioning capabilities and typically achieves a resolution improvement between two- and threefold. The advancement of this technique is one of the main aspects of this thesis, and its basics are explained in detail in the following chapter.

2

Advancements in Super-resolution Optical Fluctuation Imaging

Super-resolution Optical Fluctuation Imaging (SOFI) was invented in 2009 by Dertinger *et al* and is thus one of the newest methods in the field of super-resolution microscopy [11]. Similar to single-molecule-localization-based approaches [10, 25], the method exploits the fluorescence intermittency of emitters to produce images with enhanced spatial resolution, but in a conceptually very different way. While the former methods reconstruct an image by combining the observations of spatio-temporally isolated single dyes, SOFI relies on the computation of temporal correlation functions to disentangle the overlapping intensity patterns of molecules that are spatially close to each other. Thus, SOFI allows higher labeling densities, works better in the presence of high background signal and low signal-to-noise ratio, and is applicable to a wider range of photo-switching kinetics [27]. However, this comes at the cost of much lower resolution improvement compared to SMLM methods under optimal conditions. SOFI can be used with conventional wide-field microscopes and any labels that exhibit statistically independent fluctuations, as long as the employed camera is fast and sensitive enough to resolve them. Furthermore it equips wide-field microscopes with the ability for optical sectioning, which allows the recording of true three-dimensional images [28, 29]. The technique has been used with quantum dots [30], photo-switching dyes [27, 28, 31] and fluorescent proteins [32–34]. A very recent publication by Zhang *et al* from February 2017 used SOFI to quantify and visualize the activity of protein kinase A (PKA) in living cells [35]. This is the first reported super-resolution visualization of dynamic biochemical activities.

Currently, there are two popular approaches to SOFI: Auto-Cumulant SOFI (AC-SOFI) calculates

correlations between different points in time for each pixel position individually. In contrast, Cross-Cumulant SOFI (XC-SOFI) exploits the spatial correlation introduced by the PSF and correlates different pixels in the same frame. Both methods have their advantages and drawbacks, which is briefly discussed in the first section of this chapter. However, this chapter of the thesis concentrates on improvements for AC-SOFI.

First, the basic ideas behind SOFI are explained (Sec. 2.1) and how to calculate SOFI images from the raw movies (Sec. 2.2). Then, we will explore how SOFI performs under different imaging conditions and how many frames are needed for a desired quality (Sec. 2.3). Section 2.4 shows how to calculate artifact free virtual subpixels that increase the resolution of SOFI images beyond the pixel resolution of the camera. Next, we demonstrate how to correct for the influences of noise in zero-time-lag SOFI images provided that the distribution of the noise is known (Sec. 2.5). In Section 2.6 we show how to estimate the OTF/PSF directly from the measured data. Section 2.7 deals with boosting the resolution of SOFI images and suggest a method that adapts the resolution improvement to the data quality, avoiding artifacts. In the last section, we apply SOFI to a relatively new kind of dye, carbon nanodots, and discuss their blinking properties.

2.1 THE THEORETICAL BASIS OF SOFI

In order to use SMLM methods such as STORM, emitters have to be spatio-temporally isolated, so that (at best) the recorded signals of two emitters never overlap. Mostly, this is achieved by adjusting the chemical environment of the sample to alter the blinking behavior of the utilized dye. The positional information encoded in all images is than extracted in post-processing and collected into a single super-resolved image. If the spatio-temporal unmixing works well and the PSF model used to extract each emitter's position closely matches the experimental data, this method yields the best possible position estimates for a given photon count and thus the best possible resolution. However, although the handling of overlapping signals is possible [36, 37], it is currently very limited and generally leads to a large decrease in performance or the occurrence of artifacts. As spatio-temporal separation is a very special case which limits recording to dyes with well adjustable blinking behavior in specific environments, a method is desirable that makes no assumption about spatio-temporal isolation but handles the general case of mixed signals.

In Super-resolution Optical Fluctuation Imaging, the recorded intensity of mixed signals of stochastically fluctuating emitters is processed directly. In its basic form, the method is purely statistical and does not need any assumptions about the PSF of the microscope. We want to compute a quantity with two important characteristics:

1. The result for mixed signals should be the superposition of the results of the individual signals.

2. The resolution of the computed image should be increased compared to the original recording.

These requirements are fulfilled by a mathematical set of quantities called *cumulants*. Similar to moments, the set of all orders of cumulants fully determines a probability distribution. Most commonly

known is the cumulant of second order, which is also called the *variance*. All cumulants share certain properties which are important for our scenario, the first one being additivity: If X and Y are independent random variables, the n-th order cumulant κ_n yields

$$\kappa_n(X + Y) = \kappa_n(X) + \kappa_n(Y). \tag{2.1}$$

Thus, the cumulant of a time trace recording two independently fluctuating emitters equals the sum of the cumulant of their individual time traces (see Fig. 2.1), effectively treating the signals as if they were disentangled, which meets our first requirement. Note that the assumption of independence is crucial here and while reasonable for most imaging scenarios, one should be aware that this breaks down if two signal sources are physically coupled in a way that their blinking behavior is influenced by one another.
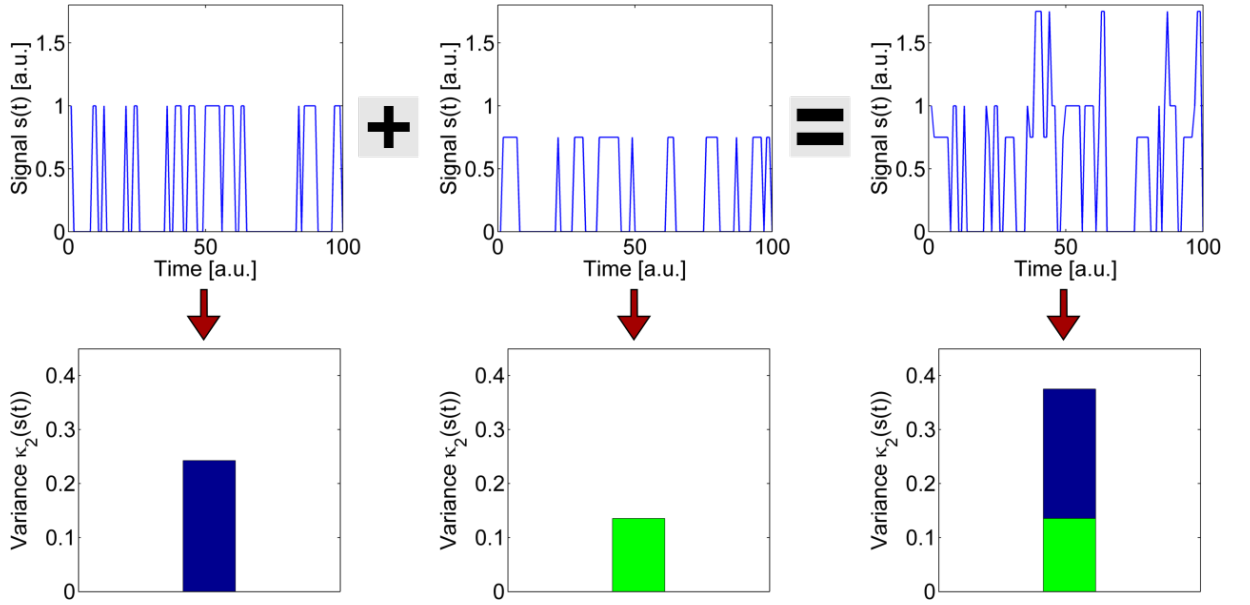
Figure 2.1: Example of additivity of the second order cumulant (the variance) for independent stochastic processes. The top row shows two independently fluctuating signals and their sum. The bottom row shows the individual cumulant values which were calculated using a simulation of 10000 frames length.

The second important property is that the n-th order cumulant is homogeneous of degree n. If c is a constant, then:

$$\kappa_n(cX) = c^n \kappa_n(X) \tag{2.2}$$

Let us take $U(r)$ to be the Point Spread Function of a detector which records the image of an emitter at $r = 0$ whose intensity is modulated by the signal $x(t)$ that randomly fluctuates over time according to $P(X)$. Furthermore the variable ε is a measure of the intensity of the emitter. When computing the n-th order cumulant over time, $U(r) \cdot \varepsilon$ is constant and we get the result

$$\kappa_n(U(r)\varepsilon \cdot x(t)) = U(r)^n \varepsilon^n \cdot \kappa_n(x(t)). \tag{2.3}$$

Temporarily assuming that the PSF can be approximated by a Gaussian, it is easy to see that this im-

proves the resolution:

$$U(\mathbf{r})^n = \left(e^{-\frac{\mathbf{r}^2}{2\sigma^2}} \right)^n = e^{-\frac{\mathbf{r}^2}{2(\sigma/\sqrt{n})^2}} = e^{-\frac{\mathbf{r}^2}{2\sigma_*^2}}. \tag{2.4}$$

The new PSF is also a Gaussian, whose standard deviation is a factor of $\sqrt{n}$ smaller than the original one, resulting in a sharper image for each recorded emitter. This resolution increase arises naturally when computing the n-th order cumulant. Looking at this in the Fourier domain, we see that a resolution increase can be expected not only for a Gaussian but for every PSF with a band-limited OTF[1]:

$$\mathcal{F}\left(U(\mathbf{r})^n \right) = \underbrace{\tilde{U}(\mathbf{k}) \circledast \cdots \circledast \tilde{U}(\mathbf{k})}_{(n-1) \text{ convolutions}} \tag{2.5}$$

The new OTF is the original one convolved $n - 1$ times with itself. The support of this function is now n times larger than before, which is consequently also true for the highest non-vanishing spatial frequency. For a perfect in-focus PSF, the highest frequency is in a linear relationship with its resolution (see Section 1.2). Thus, a linear resolution increase with the cumulant order n could be expected. However, as can be seen from the example above (Eq. (2.4)), the resolution increase for the Gaussian is only $\sqrt{n}$. This difference comes from the changed shape of the new OTF's magnitude, where the higher frequencies are dampened compared to the original shape (see Fig. 2.2). In Section 2.7 we will discuss how a proper reweighting of the frequencies in pre- or post-processing can be used to increase the resolution compared to the raw computation of cumulants.
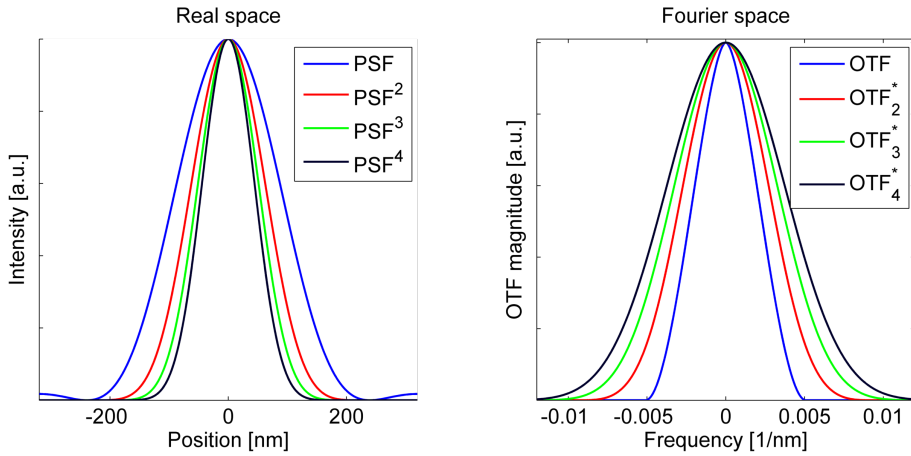
Figure 2.2: Powers of a one dimensional Airy disc point spread function and their corresponding Fourier transforms (optical transfer functions). Higher powers lead to a smaller PSF and thus exhibit a higher resolution, which corresponds to higher frequencies being present in the OTF. Note that the highest available frequencies in the OTF are more and more dampened with higher powers of the PSF.

Figure 2.3 shows a one-dimensional example of two independently blinking emitters whose signals overlap and which become better distinguishable by the cumulant computation. This illustrates how additivity and homogeneity together lead to higher resolution images. For higher order cumulants, the resolution improvement is also higher (see Fig. 2.5), but more frames are needed for the cumulant to

---

[1]An OTF which is zero outside of a finite support.

converge to a meaningful result (see Sec. 2.3).
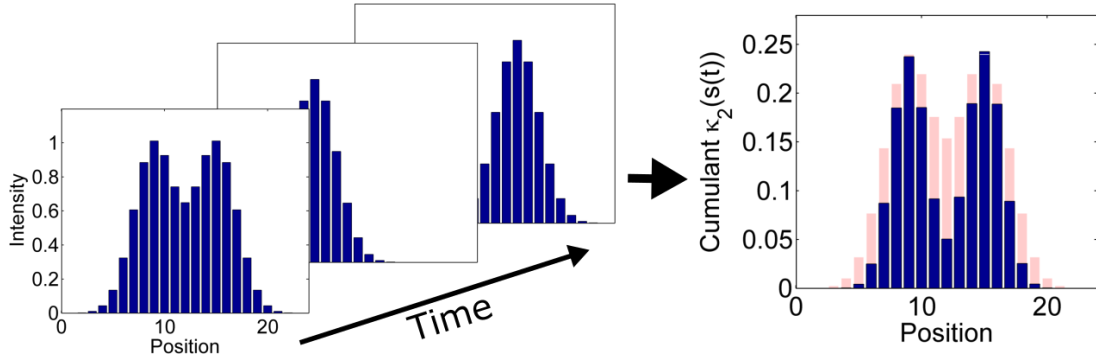
Figure 2.3: One dimensional simulation of two independently blinking emitters with identical blinking statistics and brightness. The per-pixel 2nd order cumulant values are shown on the right side. The emitters show up with a squared profile (homogeneity property) without artifacts in the mixing region (additivity property). This results in a better distinguishability compared to the original mixture (light red, scaled to same maximum).

One very interesting property of SOFI imaging that arises from the exponentiation of the wide-field PSF is demonstrated in Figure 2.4. As previously discussed (see Section 1.2) the OTF of a conventional wide-field microscope has a missing cone of frequencies along directions close to the optical axis, which leads to a complete lack of z-sectioning capability. This makes it impossible to reconstruct a three dimensional image of a sample by recording a z-stack of images with shifted focal positions. Because the OTF in SOFI imaging is an auto-convolution of the original OTF, this missing cone is closed and z-sectioning becomes possible with an ordinary wide-field microscope. This was experimentally demonstrated using a z-stack [29] as well as using special optics that allowed recording multiple focal planes simultaneously [28].
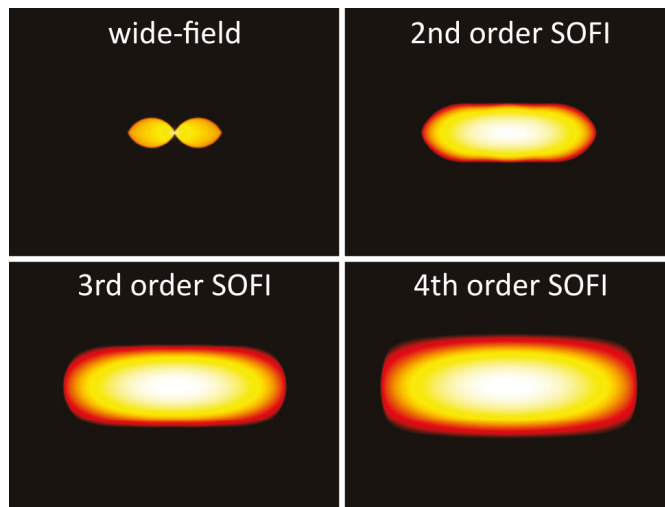
Figure 2.4: The OTF of a wide-field microscope compared to SOFI with horizontal axis along the lateral Fourier coordinate and vertical axis showing the axial Fourier coordinate. The color-scale is logarithmic from 0 (black) to 1 (white). SOFI removes the "missing cone" and widens the frequency support. Also visible is the strong relative attenuation of high frequencies with increasing order, which was illustrated in Fig. 2.2 in the one dimensional case. [Re-edit of the original image published in the article [38].]

Another important effect is the exponentiation of the brightness ε of each individual emitter, which leads to a nonlinear intensity response in the SOFI image. A dye that is twice as bright as another in the average image will thus appear four times brighter in second order SOFI. For higher orders the dynamic range of the image gets very large, leaving either only the brightest dyes visible or showing a very over-exposed image, depending on how the contrast is adjusted when viewing the image. Note that relative brightness differences within a sample caused by inhomogeneous fluorophore density will simply carry over to the SOFI image, and do not cause this nonlinear response due to the additivity of cumulants. Both situations are compared in Fig. 2.5. Another parameter that influences the brightness visible in the final image is the blinking behavior, which essentially determines the value of $\kappa(x(t))$ in Eq. (2.3). Usually the blinking behavior is assumed to be homogeneous. However, if three or more orders of cumulants can be reliably extracted, the molecular brightness, on-time ratio of blinking and molecular density can be calculated [39].
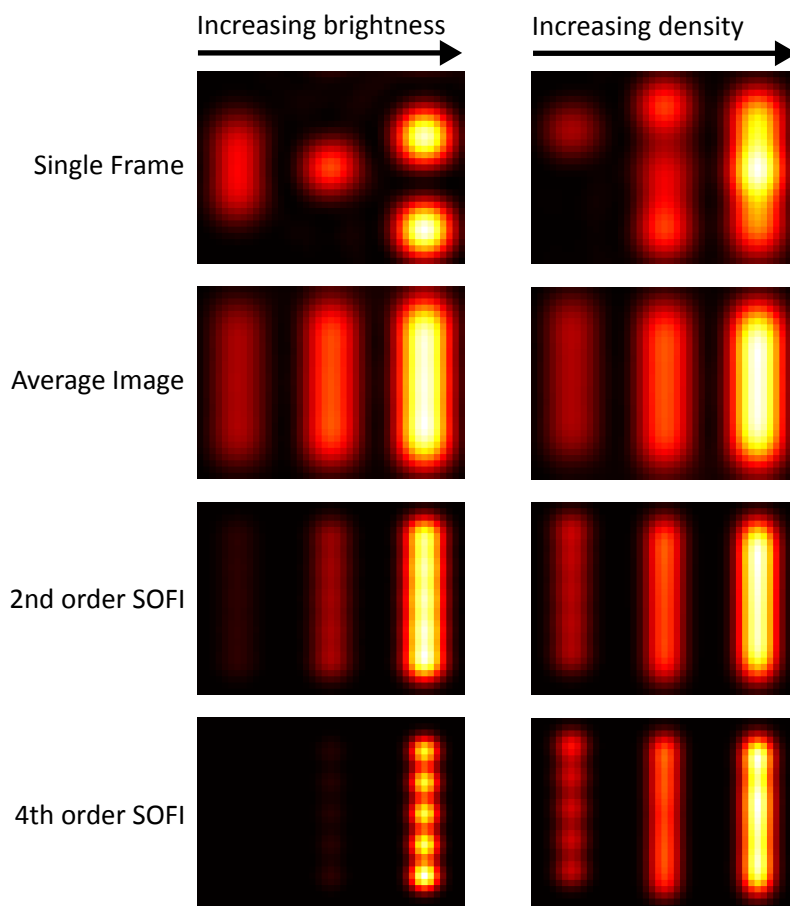
Figure 2.5: The influence of emitter density and brightness on SOFI images is visualized using two simulations showing emitters in three lines. **Left:** constant emitter density, emitter brightness increases by factor of two for each step to the right. **Right:** emitter density increases by factor of two for each step to the right, constant emitter brightness. Although both movies look identical in the time average, their SOFI images are quite different. Intensity differences caused by emitter brightness are increased by SOFI (homogeneity), while the ratio of those due to emitter density stays untouched. Note also the increasing resolution with higher SOFI orders, revealing individual emitters in the 4th order.

All cumulant orders except the first (which is equivariant) are shift invariant:

$$\kappa_1(X + c) = \kappa_1(X) + c \tag{2.6}$$

$$\kappa_n(X + c) = \kappa_n(X), \; n \geq 2 \tag{2.7}$$

Thus, adding a constant to the signal does not change the higher order cumulants. In SOFI imaging this leads to the removal of non-blinking background signal, such as contributions from autofluorescence (see Fig. 2.6).
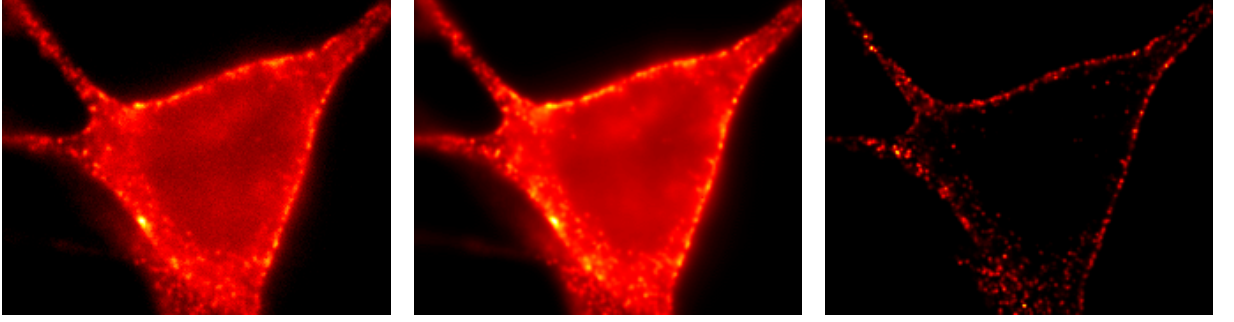
Figure 2.6: Demonstration of SOFIs background removal on experimental data. From left to right: First frame of the movie, time average and second order SOFI image. The sample is a mouse hippocampal neuron stained with quantum dots (QD_{605}) recorded by *Anja Huss*. It is clearly visible that background fluorescence in the center is greatly reduced.

This type of background is a problem for localization based super-resolution like STORM, as it leads to false or missing localizations and introduces a potential bias of estimated positions towards regions of higher background.

One last problem to address is the influence of shot noise on the estimation. Similar to the auto-correlation of the signal at zero lag time, the cumulant value is influenced by the presence of noise in the signal. For example, the second order cumulant (also known as the variance) for a noisy signal I at position $\mathbf{r}$ in time t

$$\kappa_2\left(I(\mathbf{r}, t)\right) = \left\langle \delta I(\mathbf{r}, t)^2 \right\rangle = \left\langle \left(I(\mathbf{r}, t) - \langle I(\mathbf{r}, t) \rangle\right)^2 \right\rangle = \left\langle I(\mathbf{r}, t)^2 \right\rangle - \langle I(\mathbf{r}, t) \rangle^2 \tag{2.8}$$

simply estimates the power of the total signal fluctuation around the mean, including the contributions from noise. Here $\langle \bullet \rangle$ denotes the time average. The noise contaminates the super-resolution image with diffraction limited information (see Section 2.5). A possible solution to this problem is correlating different points in time and/or space, for which *joint cumulants K_n* can be used. The joint cumulant of the n-th order is computed for a set of n random variables instead of one, but shares all properties of ordinary cumulants that were previously discussed. For example, the second order joint cumulant can be computed as

$$K_2\left(I(\mathbf{r}_1, t + \tau_1), I(\mathbf{r}_2, t + \tau_2)\right) = \left\langle \delta I(\mathbf{r}_1, t + \tau_1) \delta I(\mathbf{r}_2, t + \tau_2) \right\rangle \tag{2.9}$$

$$= \left\langle I(\mathbf{r}_1, t + \tau_1) I(\mathbf{r}_2, t + \tau_2) \right\rangle - \left\langle I(\mathbf{r}_1, t + \tau_1) \right\rangle \left\langle I(\mathbf{r}_2, t + \tau_2) \right\rangle \tag{2.10}$$

which suppresses the shot noise contribution if $\mathbf{r}_1 \neq \mathbf{r}_2$ **or** $\tau_1 \neq \tau_2$, as the noise is not correlated in time or space. The n-th order joint cumulant in SOFI can be computed for n different positions with n time-lags. Based on this, we define the two most commonly used approaches to SOFI:

$$\text{Auto-cumulant SOFI (AC-SOFI): } \mathbf{r}_1 = \mathbf{r}_2 = \cdots = \mathbf{r}_n, \text{ and } \tau_i \neq \tau_j, \forall i \forall j | i \neq j$$

$$\text{Cross-cumulant SOFI (XC-SOFI): } \mathbf{r}_i \neq \mathbf{r}_j, \forall i \forall j | i \neq j, \text{ and } \tau_l \text{ arbitrary}$$

Although arbitrary time-lags could be used with cross-cumulants, in most cases they are simply set to zero. The n-th order cross-cumulant XC_n carries information about the location of the geometric average of the positions of all correlated pixels if the PSF $U(\mathbf{r})$ is a symmetric Gaussian [40]:

$$K_n\left(I(\mathbf{r}_1, t+\tau_1), \ldots, I(\mathbf{r}_n, t+\tau_n)\right) = XC_n(\mathbf{r}_1, \ldots, \mathbf{r}_n, \tau_1, \ldots, \tau_n) =$$

$$\left[\prod_{j<l}^{n} U\left(\frac{\mathbf{r}_j - \mathbf{r}_l}{\sqrt{n}}\right)\right] \cdot \sum_{v=1}^{N_{\text{em}}} U^n\left(\mathbf{r}_v - \frac{\sum_{k=1}^{n} \mathbf{r}_k}{n}\right) \cdot \varepsilon_v^n \cdot K_{n,v}(s_v(t+\tau_1), \ldots, s_v(t+\tau_n)) \qquad (2.11)$$

Here ε_v is the brightness of the v-th emitter, $s_v(t) \in [0, 1]$ its time dependent fluctuation, and N_{em} the overall number of emitters in the image. This relationship can be used to generate an image on a finer pixel grid than the original data. Note that there are multiple pixel combinations which contribute to the same position and the number of possibilities grows exponentially with the joint cumulant order. For higher orders, it is computationally prohibitive to calculate every possibility and thus a finite subset must be chosen. This subset can be combined using weighting factors to compose the final result. It is not immediately clear how to these weights have to be chosen, leading to an optimization problem [41]. Additionally, equation (2.11) shows that each pixel combination has an intrinsic weighting factor based on the distance of all possible pairs of pixels. This leads to grid-like artifacts in the final image, which have to be corrected. If the sample is in focus as well as thin, and if the PSF is homogeneous in the field of view, a global correction of these weighting factors can be performed. Out-of-focus emitters as well as PSF-inhomogeneities violate Eq. (2.11) and are hard to correct, leaving artifacts in the image. However, the possibility of using different combinations to extract information on the same position has the potential to strongly improve the SNR of the final image.

It is also possible to generate images using different values of the time-lags τ_n. Note that for XC-SOFI negative time-lags can be chosen, because the averages for a finite movie of T frames

$$\frac{1}{T}\sum_{t=1}^{T} \delta I(\mathbf{r}_1, t)\delta I(\mathbf{r}_2, t+\tau) \neq \frac{1}{T}\sum_{t=1}^{T} \delta I(\mathbf{r}_1, t+\tau)\delta I(\mathbf{r}_2, t)$$

are unequal, while in AC-SOFI each set of time-lags including negative values is identical to another set using only positive values. Potentially, images computed for different time-lags could be combined to improve the SNR of the final image, but it is again not clear how to do this in an optimal way. Many approaches use one fixed set of time-lags, which are usually the smallest possible values, as the signal correlation drops exponentially over time. This reflects the time-scale of the blinking kinetics of the

fluorophore.

XC-SOFI has the advantage of being able to use time-lags of zero without the influence of shot noise. This is important when using dyes with very fast blinking dynamics on the order of the exposure time, where the cumulant decays quickly with the time-lag.

To sum up this section about the theoretical basics of SOFI:

- SOFI is based on the calculation of statistical quantities called cumulants from recordings of samples with stochastically varying intensity.

- Cumulants are additive for independent signal sources. Thus, measuring a mixture of those signals yields the same result as measuring each individual signal and adding them up later. Fluorescent molecules couple only in special situations, making this property useful for the analysis of recordings in fluorescence microscopy – even if the images of molecules overlap on the sensor.

- The n-th order cumulant is homogeneous of degree n: $\kappa_n(cX) = c^n \kappa_n(X)$. Calculating the cumulant for a movie with the PSF $U(\mathbf{r})$ produces an image with the new PSF $U^n(\mathbf{r})$. This increases the optical resolution, surpassing the diffraction limit.

- The SOFI PSF $U^n(\mathbf{r})$ corresponds to a resolution increase of factor $\sqrt{n}$ for in-focus molecules, despite having an n-fold higher cutoff frequency in Fourier space. This is caused by a disadvantageous shape compared to that of the stretched PSF $U(n \cdot \mathbf{r})$, which has the same frequency support. Thus SOFI can theoretically achieve a linear resolution increase of order n when additional processing is performed compared to the pure calculation of cumulants.

- SOFI closes the *missing cone* of frequencies present in the OTF of a wide-field microscope. This enables z-sectioned 3D recordings of samples even when a wide-field microscope is used for recordings.

- SOFI images show non-linear brightness with respect to emitter intensity ε^n, but scale linear with emitter density (because cumulants are additive). Thus, molecular brightness differences of emitters are exaggerated, while brightness differences caused by varying density are mapped correctly.

- Non-varying signals are removed by the cumulant calculation. This removes background signals, for example from autofluorescence, and enhances image contrast.

- Shot noise leads to the superposition of the desired high resolution image with lower resolution information when ordinary cumulants are calculated (see Sec.2.5). To circumvent this, correlations between different frames (AC-SOFI) and/or spatial positions (XC-SOFI) can be calculated using joint cumulants. This removes contributions from signals uncorrelated in time and/or space.

2.2 From cumulants to SOFI computation

In order to calculate a SOFI image, we need to derive the formula for the computation of the cumulants: The *moment generating function* of a random variable X is given by

$$c(t) = \langle \exp(tX) \rangle, \tag{2.12}$$

where $\langle \bullet \rangle$ denotes the ensemble average. The logarithm of this function

$$g(t) = \log c(t) = \log \langle \exp(tX) \rangle. \tag{2.13}$$

is called the *cumulant generating function*. This is, because the cumulants are defined using a power series expansion of $g(t)$:

$$g(t) = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}. \tag{2.14}$$

The cumulant κ_n contains all terms proportional to a certain power t^n of t, times the factorial $n!$. Using the series representation of the exponential function

$$e^{tX} = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^n = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} X^n \tag{2.15}$$

we rewrite eq. (2.13):

$$g(t) = \log \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right) \tag{2.16}$$

We can further expand this using the Taylor series of the logarithm

$$\log(1 + x) \approx x - \frac{1}{2} x^2 + \frac{1}{3} x^3 - \ldots \tag{2.17}$$

to get

$$\log \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right) \approx \tag{2.18}$$

$$\sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle - \frac{1}{2} \sum_{n,k=1}^{\infty} \frac{t^{n+k}}{n!k!} \langle X^n \rangle \langle X^k \rangle + \frac{1}{3} \sum_{n,k,l=1}^{\infty} \frac{t^{n+k+l}}{n!k!l!} \langle X^n \rangle \langle X^k \rangle \langle X^l \rangle - \ldots \tag{2.19}$$

Now we can read out the cumulants using the terms proportional to powers of t multiplied by $n!$. For example, the first three cumulants (called *mean*, *variance* and *skewness*) are given by:

$$t^1 \quad \rightarrow \quad \kappa_1 = 1! \cdot \langle X \rangle = \langle X \rangle$$

$$t^2 \rightarrow \kappa_2 = 2! \cdot \left(\frac{\langle X^2 \rangle}{2} - \frac{1}{2} \langle X \rangle \langle X \rangle \right) = \langle X^2 \rangle - \langle X \rangle^2 = \left\langle (X - \langle X \rangle)^2 \right\rangle$$

$$t^3 \rightarrow \kappa_3 = 3! \cdot \left(\frac{\langle X^3 \rangle}{6} - \frac{1}{2} \left(2 \cdot \frac{\langle X^2 \rangle \langle X \rangle}{2} \right) + \frac{1}{3} \langle X \rangle \langle X \rangle \langle X \rangle \right)$$

$$= \langle X^3 \rangle - 3 \langle X^2 \rangle \langle X \rangle + 2 \langle X \rangle^3 = \left\langle (X - \langle X \rangle)^3 \right\rangle$$

Although the second and third order cumulant are equivalent to the same order *central moment* $\mu_n = \langle (X - \langle X \rangle)^n \rangle$, this is not true for higher orders. An alternative way to compute the cumulant is using derivatives of the cumulant generating function evaluated at $t = 0$.

$$\kappa_n = \frac{\partial^n}{\partial t^n} g(t) \bigg|_{t=0} = \frac{\partial^n}{\partial t^n} \log \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right) \bigg|_{t=0} \tag{2.20}$$

For the first cumulant we simply get

$$\kappa_1 = \frac{\partial}{\partial t} \log \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right) \bigg|_{t=0} = \left(\sum_{n=1}^{\infty} \frac{n t^{n-1}}{n!} \langle X^n \rangle \right) \cdot \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right)^{-1} \bigg|_{t=0}$$

$$= \left(\langle X \rangle + \sum_{n=2}^{\infty} \frac{n t^{n-1}}{n!} \langle X^n \rangle \right) \cdot \left(1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \langle X^n \rangle \right)^{-1} \bigg|_{t=0} = \langle X \rangle$$

Using the time series of intensity recorded at one pixel as input, these (ordinary) cumulants yield the SOFI image values without time-lag. As stated in the previous section, this calculation also includes contributions from noise. To eliminate these contributions, we have to correlate different data points in space and/or time, for which *joint cumulants* are necessary. The joint cumulant $K_m(X_1, X_2, \ldots, X_m)$ of m signals X_i is a function that it is nonzero only if there is correlation between **all** signals. If one of the variables is stochastically independent from the others, the cumulant yields always zero.

The approach for computing the joint cumulant is similar to before, but the new moment generating function depends on the sum of all random variables with individual prefactors. To give an example, we first compute the joint cumulant for two random variables. The joint cumulant generating function is given by:

$$g(t_1, t_2) = \log c(t_1, t_2)$$
$$= \log \langle \exp(t_1 X_1 + t_2 X_2) \rangle$$
$$= \log \langle \exp(t_1 X_1) \cdot \exp(t_2 X_2) \rangle$$
$$= \log \left(\sum_{n,k=0}^{\infty} \frac{t_1^n t_2^k}{n! k!} \langle X_1^n X_2^k \rangle \right)$$
$$= \log \left(1 + \underbrace{\sum_{n=1}^{\infty} \frac{t_1^n}{n!} \langle X_1^n \rangle + \sum_{k=1}^{\infty} \frac{t_2^k}{k!} \langle X_2^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n! k!} \langle X_1^n X_2^k \rangle}_{\hat{=} x} \right)$$

To get the joint cumulant, we now search for all terms proportional to $t_1 t_2$ in the power series expansion of this expression. Note that in the two-variable-case, these terms can only exist up to the second order in the expansion, which is given by:

$$
\begin{aligned}
\log c(t_1, t_2) \approx & \sum_{n=1}^{\infty} \frac{t_1^n}{n!} \langle X_1^n \rangle + \sum_{k=1}^{\infty} \frac{t_2^k}{k!} \langle X_2^k \rangle + \underbrace{\sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n X_2^k \rangle} \\
& - \frac{1}{2} \left[\sum_{n,k=1}^{\infty} \frac{t_1^{n+k}}{n!k!} \langle X_1^n \rangle \langle X_1^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_2^{n+k}}{n!k!} \langle X_2^n \rangle \langle X_2^k \rangle + \sum_{n,k,l,i=1}^{\infty} \frac{t_1^{n+l} t_2^{k+i}}{n!k!l!i!} \langle X_1^n X_2^k \rangle \langle X_1^l X_2^i \rangle \right. \\
& + 2 \left(\sum_{n,k,l=1}^{\infty} \frac{t_1^{n+l} t_2^k}{n!k!l!} \langle X_1^n X_2^k \rangle \langle X_1^l \rangle + \sum_{n,k,l=1}^{\infty} \frac{t_1^n t_2^{k+l}}{n!k!l!} \langle X_1^n X_2^k \rangle \langle X_2^l \rangle + \underbrace{\sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n \rangle \langle X_2^k \rangle} \right) \right] \\
& + \frac{1}{3} \dots
\end{aligned}
\tag{2.21}
$$

Only the under-braced parts contain terms proportional to $t_1 t_2$. Collecting these, we get for the joint cumulant of both variables:

$$
K_2(X_1, X_2) = \langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle
\tag{2.22}
$$

Setting $X_1 = X_2 = X$ yields

$$
K_2(X, X) = \langle X^2 \rangle - \langle X \rangle^2 = \kappa_2(X),
\tag{2.23}
$$

which is the second order ordinary cumulant for one variable, as calculated before. An easy algorithm to compute the formulas for higher order joint cumulants is explained in Appendix E.

To ease the implementation of SOFI algorithms, we now give the first four cumulant formulas. In the simplest case with zero time-lag at one pixel position, ordinary cumulants are used. With the fluorescent signal $F = F(\mathbf{r}, t)$ at pixel position $\mathbf{r}$ at time t and $\delta F = F - \langle F \rangle$ the ordinary cumulants up to order $n = 4$ are given by:

n	Cumulant $\kappa_n \left[F(\mathbf{r}, t) \right]$
1	$\langle F \rangle$
2	$\langle \delta F^2 \rangle = \langle F^2 \rangle - \langle F \rangle^2$
3	$\langle \delta F^3 \rangle = \langle F^3 \rangle - 3 \langle F \rangle \langle F^2 \rangle + 2 \langle F \rangle^3$
4	$\langle \delta F^4 \rangle - 3 \langle \delta F^2 \rangle^2 = \langle F^4 \rangle - 4 \langle F \rangle \langle F^3 \rangle - 3 \langle F^2 \rangle^2 + 12 \langle F \rangle^2 \langle F^2 \rangle - 6 \langle F \rangle^4$

Usually multiple points in time and/or space should be correlated to get rid of the influences of shot noise, for which joint cumulants must be used. With $F_i = F(\mathbf{r_i}, t + \tau_i)$ and $\delta F_i = F_i - \langle F_i \rangle$, the first four joint cumulants are given by:

n	Joint cumulant $K_n\left[F(\mathbf{r}_1, t+\tau_1), \ldots, F(\mathbf{r_n}, t+\tau_n)\right]$
1	$\langle F_1 \rangle$
2	$\langle \delta F_1 \delta F_2 \rangle = \langle F_1 F_2 \rangle - \langle F_1 \rangle \langle F_2 \rangle$
3	$\langle \delta F_1 \delta F_2 \delta F_3 \rangle =$ $\langle F_1 F_2 F_3 \rangle - \langle F_1 \rangle \langle F_2 F_3 \rangle - \langle F_2 \rangle \langle F_1 F_3 \rangle - \langle F_3 \rangle \langle F_1 F_2 \rangle + 2 \langle F_1 \rangle \langle F_2 \rangle \langle F_3 \rangle$
4	$\langle \delta F_1 \delta F_2 \delta F_3 \delta F_4 \rangle - \langle \delta F_1 \delta F_4 \rangle \langle \delta F_2 \delta F_3 \rangle - \langle \delta F_1 \delta F_3 \rangle \langle \delta F_2 \delta F_4 \rangle - \langle \delta F_1 \delta F_2 \rangle \langle \delta F_3 \delta F_4 \rangle$ $= \langle F_1 F_2 F_3 F_4 \rangle$ $- \Big(\langle F_1 \rangle \langle F_2 F_3 F_4 \rangle + \langle F_2 \rangle \langle F_1 F_3 F_4 \rangle + \langle F_3 \rangle \langle F_1 F_2 F_4 \rangle + \langle F_4 \rangle \langle F_1 F_2 F_3 \rangle$ $\qquad + \langle F_1 F_2 \rangle \langle F_3 F_4 \rangle + \langle F_1 F_3 \rangle \langle F_2 F_4 \rangle + \langle F_1 F_4 \rangle \langle F_2 F_3 \rangle \Big)$ $+ 2 \cdot \Big(\langle F_1 \rangle \langle F_2 \rangle \langle F_3 F_4 \rangle + \langle F_1 \rangle \langle F_3 \rangle \langle F_2 F_4 \rangle + \langle F_1 \rangle \langle F_4 \rangle \langle F_3 F_2 \rangle$ $\qquad + \langle F_2 \rangle \langle F_3 \rangle \langle F_1 F_4 \rangle + \langle F_2 \rangle \langle F_4 \rangle \langle F_1 F_3 \rangle + \langle F_3 \rangle \langle F_4 \rangle \langle F_1 F_2 \rangle \Big)$ $- 6 \cdot \langle F_1 \rangle \langle F_2 \rangle \langle F_3 \rangle \langle F_4 \rangle$

The tables show two equal formulas for every order, which differ just in the way they are computed. From a numerical standpoint, the two-pass algorithm which first calculates the mean and then averages the fluctuations δF_i has increased stability over the one-pass summation that uses the values F_i directly. If the number of frames is very large, the roundoff error can lead to poor results for both formulas and techniques such as *compensated summation* should be used. However, the number of frames is usually sufficiently small due to the limited photostability of the sample.

The aforementioned properties of cumulants that SOFI builds upon need the underlying process of the signal to be stationary, i.e. its statistical properties should not change over time. This is not strictly true for fluorescent molecules whose behavior can rapidly change due to photobleaching. The change in intensity caused by photobleaching contaminates the SOFI image with diffraction limited information, thus lowering the image quality. For this reason, each image stack should be subdivided into sub-stacks with N_w frames each, which are processed individually. The number of frames N_w must be smaller than the bleaching time-scale in order to avoid artifacts, but should be greater than the typical correlation time-scale of the emitters' intensity fluctuations. Finally, the sub-stacks can simply be added up to the final SOFI image. A visual comparison of this interval-based computation to a computation from all frames is shown in Fig. 2.10 in the next section. For very short intervals the cumulant estimator gets biased (see Appendix D), and further research is required to check if this bias can be exploited in any useful way.

To sum up this section about the computation of SOFI images using cumulants:

- Ordinary cumulants are derived from a power series expansion of the cumulant generating function $\log \langle \exp(tX) \rangle$. The n-th order cumulant is built by collecting all terms proportional to t^n times the factorial $n!$.

- Joint cumulants are derived from a power series expansion of the joint cumulant generating func-

tion $\log\langle\exp(t_1 X_1 + \cdots + t_n X_n)\rangle$. The n-th order joint cumulant is constructed by collecting all terms proportional to $t_1 \cdot \cdots \cdot t_n$.

- The important properties of cumulants only hold if the signal is stationary. To correct for bleaching of the sample (non-stationary behavior), an image stack should be divided into sub-stacks which are processed individually. The final result can be obtained by averaging. For this to work as intended, the substack size should be smaller than the bleaching time-scale, but larger than the typical time-scale of the emitters' intensity fluctuations.

2.3 CONVERGENCE OF CUMULANTS

The theoretical ideas which SOFI is based on are not only built on the assumption of stationarity that was just mentioned, but also assume an infinite recording time, i.e. that the cumulants have converged to their theoretical values. For finite recordings, computing a statistical measure gives only an estimation of the true value with some error depending on the length of the sampling and its properties. For example, computing the average image of a recording of blinking emitters is not expected to give a good approximation of the true average if only a handful of frames are taken, but for a few thousand frames the result is expected to be much closer. The same is true for the computation of cumulants – the more frames the better. The experimentally important question is how fast the cumulants converge to their true values, as this determines how many frames have to be recorded in a given experimental setting to achieve a certain quality of the result, which ultimately limits the imaging speed. This is also an important question for choosing the right fluorophore, as it has to live long enough and produce enough photons for the desired result to be achievable. The performance of SOFI in different imaging conditions was previously investigated by Geissbuehler *et al* [27] using a self-defined "visibility" measure. We extend these results by calculations relying on fundamental statistics, and estimate the necessary recording time for different imaging conditions and SOFI orders. More recently, Vandenberg *et al* proposed a method to estimate the signal-to-noise ratio of SOFI images from the raw data using methods from numerical statistics [41]. This is used to increase the SNR of XC-SOFI by optimizing the weights of different types of spatial cross-correlations. Here, we will use a similar SNR measure to lookt at the performance from a more general point of view and also compare different SOFI orders. Very recently, a similar analysis was published by Wang *et al*, who did a quantitative analysis of the variance of cross-cumulants of different order [42]. Their results are mostly complementary to the findings presented here. However, the analysis has a few flaws, as their statistical analysis does not take into account that the time series is correlated (i.e. it is only valid for very fast dynamics) and they wrongly assume that the variance of cumulants is simply additive. This is not true, as we will illustrate by investigating the density dependence of the statistical error. Additionally, we present some qualitative results, assuming a two-state-model for the fluorescence intermittency.

Hereinafter we look at the quality of SOFI images with respect to blinking dynamics, time-lag, emitter density, emitter brightness and noise. As a tool, we use a simple simulation in which each emitter is either completely on or completely off during a frame and follows a two-state model with fixed transition

rates between the states. The visual assessment is complemented by a new error measure to quantify the convergence. To simplify, we start without taking shot noise into account, i.e. we analyze images that would be expected for extremely high photon yield per emitter per frame. The color contrast for every SOFI image in this chapter is set individually. In cases where the theoretical SOFI result is negative, the sign of all pixel values was switched before display to ease comparison between images. If the background surrounding the emitters turns red, this is an indication for positive and negative values being present in the image at the same time. Such images might seem worse, compared to strictly positive images, than they really are (i.e. when showing the absolute values), but as different signs of the values indicate bad convergence (all emitters follow the same statistics and should have the same sign) it is left as is. The following two-state blinking dynamics set by their on-/off-state lifetimes ($\tau_{\mathrm{on}}, \tau_{\mathrm{off}}$) will be considered in this section:

- Slow dynamics: $\tau_{\mathrm{on}} = 6.67$ frames, $\tau_{\mathrm{off}} = 10$ frames, $\frac{\tau_{\mathrm{on}}}{\tau_{\mathrm{off}}} \approx \frac{2}{3}$
- Fast dynamics: $\tau_{\mathrm{on}} = 1.67$ frames, $\tau_{\mathrm{off}} = 2.5$ frames, $\frac{\tau_{\mathrm{on}}}{\tau_{\mathrm{off}}} \approx \frac{2}{3}$
- Similar on-/off-times: $\tau_{\mathrm{on}} = 1.96$ frames, $\tau_{\mathrm{off}} = 2.04$ frames, $\frac{\tau_{\mathrm{on}}}{\tau_{\mathrm{off}}} \approx 0.96$
- Long off-time: $\tau_{\mathrm{on}} = 1.67$ frames, $\tau_{\mathrm{off}} = 25$ frames, $\frac{\tau_{\mathrm{on}}}{\tau_{\mathrm{off}}} \approx 0.067$

Time is always given in frames here and can be converted into seconds by multiplication with an arbitrary exposure time. The sample composed of emitters arranged in three lines with increasing density ($0.4 \frac{1}{\mathrm{px}}$, $0.8 \frac{1}{\mathrm{px}}$, $1.6 \frac{1}{\mathrm{px}}$; similar to Fig. 2.5), with all emitters having the same intrinsic brightness. The number of emitters contributing to a single pixel is higher than the density, as the PSF spreads over multiple pixels. The standard deviation (i.e. width) of the PSF is given by $\sigma_{\mathrm{PSF}} = 3$ pixels.
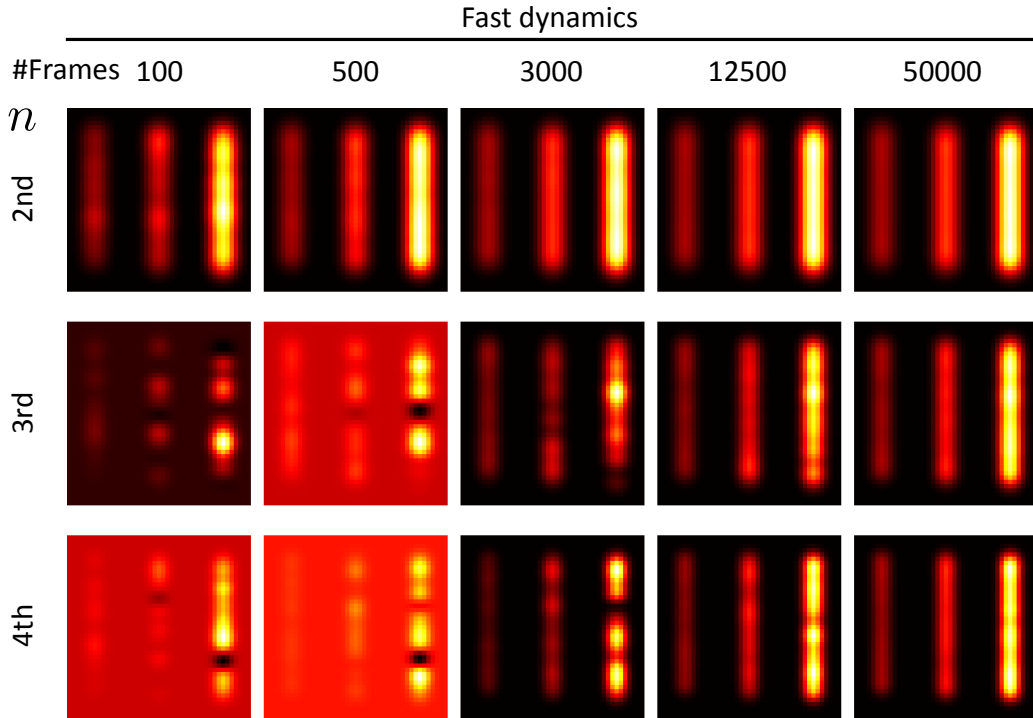
Figure 2.7: Comparison of the convergence of different SOFI orders with respect to the recording time of a noise-less movie. While the second order converges very quickly, achieving good quality at 500 frames, the higher orders need 50000 frames for a similar visual impression. The performance of 3rd and 4th order is very similar under these blinking dynamics.

Figure 2.7 shows the convergence of different SOFI orders calculated from noise-less data with fast dynamics. It is clearly visible that the second order converges very quickly, with almost no change in the quality being perceptible for movies with more than 500 frames. The 3rd and 4th order yield very similar quality in this setting and are much slower in convergence, needing approximately 100 times more frames than the second order for a comparable result. As higher orders extract higher resolution, it is intuitive that better statistics are needed to extract this information. The similarity of the 3rd and 4th order is surprising in this context, but can be connected to the way the blinking dynamics are set up, as we will shortly see. Additionally, the cumulant values of pixels in regions of higher emitter density seem to converge slower. This is analyzed in more detail below.
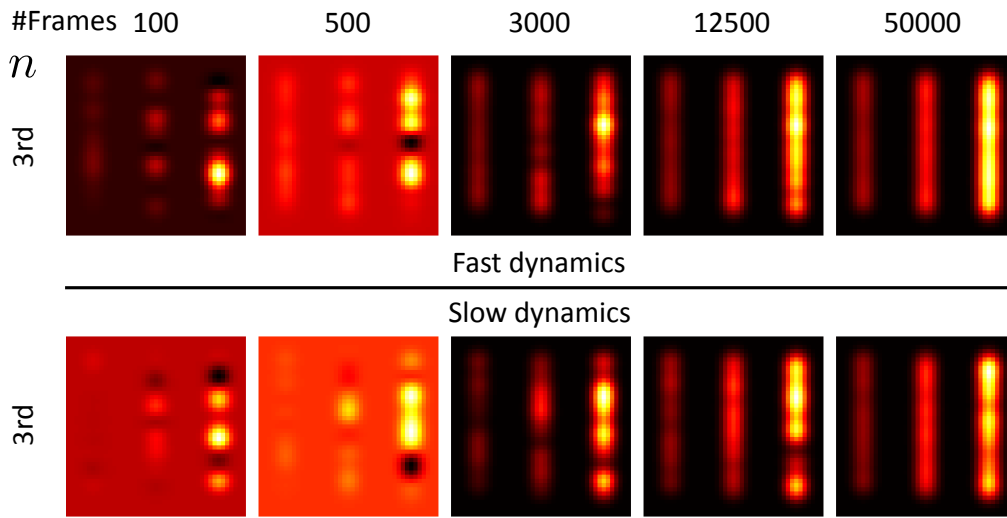
Figure 2.8: Comparison of 3rd order convergence for fast and slow dynamics, where the expected period $\tau_{on} + \tau_{off}$ of blinking events is about 4 times longer. The ratio of on-/off-state lifetimes $\frac{\tau_{on}}{\tau_{off}} \approx 2/3$ is constant . The SOFI images are calculated from a noise-less simulation. As expected, for the slower dynamics roughly four times more frames are needed to yield a similar results.

For slower dynamics, more frames are needed, which is shown in Fig. 2.8. The expected period of blinking events $\tau_{on} + \tau_{off}$ is about 4 times longer in the slow dynamics setting than in the fast one. Unsurprisingly, a comparison of the 3rd order SOFI results for 3000, 12500 and 50000 frames shows that the slow dynamics are always one image (about factor 4 in frames) behind in terms of quality. As a consequence, fluorophores that blink faster should be preferred over slower ones, and the exposure time should roughly match the characteristic time-scale of the blinking dynamics. However, if time-lags greater than zero are used for computation, it can be beneficial to spread the on-state of the emitters over multiple frames (see end of this section).

Changing the ratio of the on-/off-state times impacts the convergence of the SOFI orders as well as the relative quality difference between them at a fixed recording time. A comparison between three different scenarios is shown in Figure 2.9, where the 2nd order is left out, as it converges very quickly in all cases. For very similar on-/off -times, the 3rd order does not converge within the first 50000 frames, while the 4th order yields a good result for 12500 frames or more. As the off- and on-time become more dissimilar, the 3rd order converges much quicker and yields quality similar to the 4th order. This is the
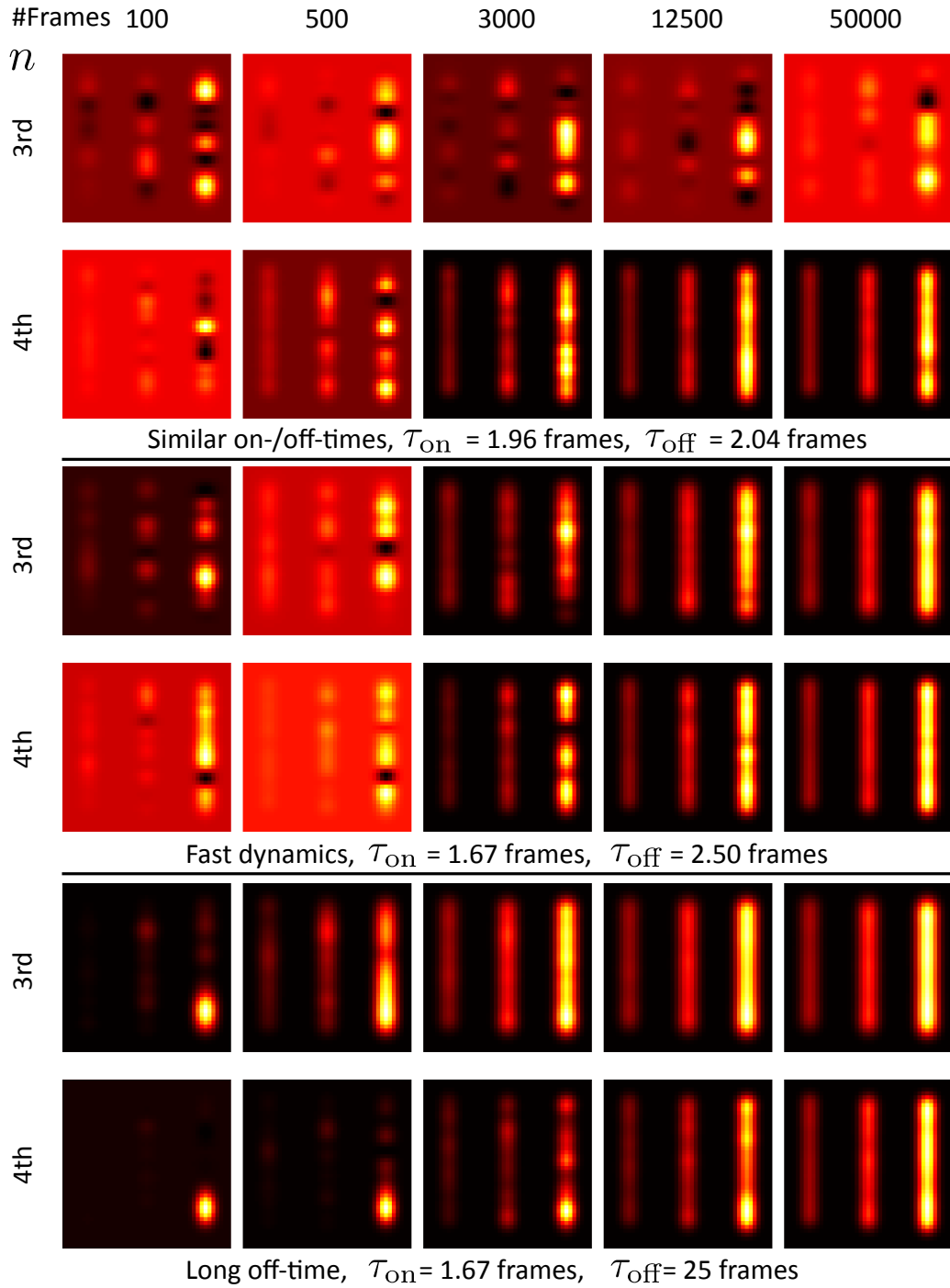
Figure 2.9: Comparison of SOFI convergence computed from noise-less simulations with different on-/off-state lifetimes. The 2nd order is not shown as it converges very quickly in all cases. Top: If the on-/off-times are very similar, no convergence is observed in the 3rd order for up to 50000 frames, while the 4th order yields good results. Mid: As the on- and off-time becomes more dissimilar the third order converges and yields similar quality than the 4th order. Bottom: For off-times much larger than the on-time, the 3rd order now converges much more quickly than the 4th order, giving a good image at 3000 frames. In all cases the 4th order needs between 12500 and 50000 frames to achieve a good quality with high homogeneity of the image.

situation we already observed in Fig. 2.7. For off-times much larger than the on-time, the third order achieves a good image with 3000 frames and thus much quicker than the 4th order. This agrees with the results of Geissbühler *et al* [27], where the quality of SOFI images was analyzed in dependence of the on/off ratio for a fixed number of frames. According to their analysis, even cumulant orders perform

best when the on- and off-time are equal, while odd orders perform the worst at this point and generally profit from ratios that differ from one. For the 3rd order this is actually quite easy to see by looking at the corresponding cumulant formula (without time-lag):

$$\kappa_3(F(t)) = \left\langle \delta F^3(t) \right\rangle = \left\langle (F(t) - \langle F \rangle)^3 \right\rangle \tag{2.24}$$

If $F(t)$ follows a two state model with $F_{on} = I$, $F_{off} = 0$ and the on-state is observed as often as the off-state, then the average is exactly in the center $\langle F \rangle = I/2$, leading to $\delta F_{off} = -I/2$ and $\delta F_{on} = +I/2$. As $\delta F^3(t)$ takes the sign from δF, and both states are observed the same number of times, the average adds up exactly to zero. Thus, the 3rd order will never produce a useful SOFI image in this singular case.
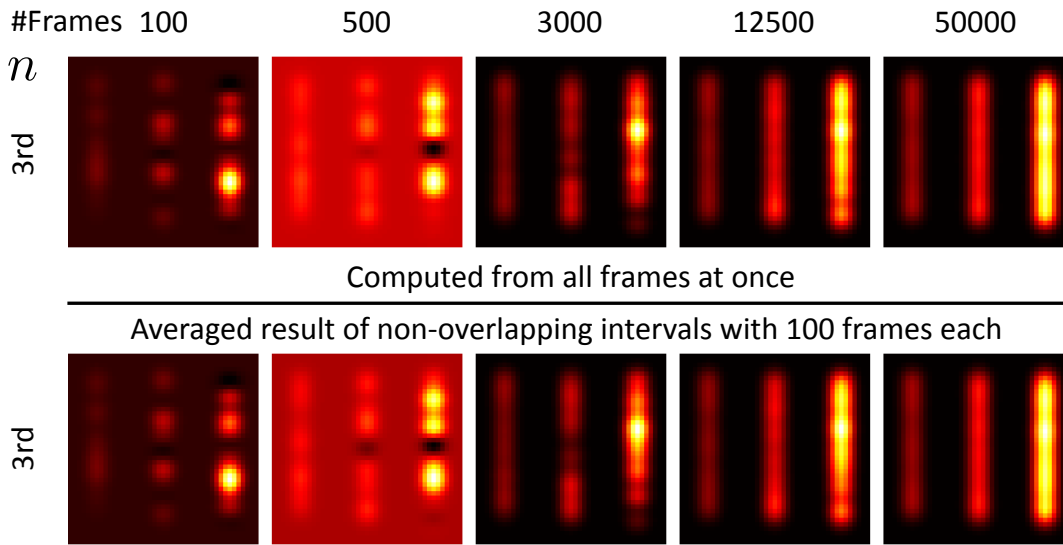
Figure 2.10: Comparison of 3rd order SOFI computation from all available frames at once against an interval-based approach, where an individual image is computed for successive intervals of 100 frames and the result is calculated by averaging all interval computations. The underlying data is noise-less and shows fast dynamics, giving $100/(\tau_{on} + \tau_{off}) \approx 24$ observed blinking events per interval. The differences between both approaches are minor and both converge to the same result. When larger intervals are used, the differences become imperceptible. For very small intervals the cumulant calculation becomes biased (see Appendix D), leading to different images especially at high emitter densities. This bias could potentially be beneficial to the reconstruction, but further research is required to check this.

Although the visual inspection of results gives a good idea on the general convergence properties, it is useful to introduce a more broadly applicable quantitative measure for the quality of SOFI, to compare different orders and give estimations for the expected number of frames needed for a certain quality. It was stated in Section 2.2 that it is often useful to subdivide a movie into intervals from which individual SOFI images are computed, and take the final result as the average of all intervals. This is usually done to counter the effects of bleaching, which violates stationarity and leads to the mean signal decreasing over time. For this to be useful, the interval-based approach must yield the same result as a computation from all frames when no bleaching present. Figure 2.10 shows a comparison of these two approaches. In this example, roughly 24 blinking events are captured per interval and the differences between both ways of computation are minor. For even larger intervals they become imperceptible. Minor differences arise from events close to the interval border, which slightly change the observed distribution

as the state's occupancy time is cut to fit within the interval. As long as the intervals are larger than the typical time-scale of blinking, they contain sufficient statistics and yield results very similar to the single-block computation, making interval-based computation an acceptable way to compute SOFI images. It is very important to note, that the cumulant calculation gets biased (see Appendix D) for very short intervals. Simulations revealed that this bias can potentially be exploited in a useful way, improving the convergence of higher order cumulants at high densities, but further research is required to check this.

As shown later for the 2nd order, it is possible to derive the theoretical formulas for the statistical error of each SOFI order for a finite measurement time. If the theoretical values of the cumulants are known for a specific blinking model, they can be used to calculate the number of frames necessary to estimate the true value with a defined precision. However, these formulas quickly get very extensive for higher SOFI orders and must be adapted when the stochastic model of the fluorophore is changed. Therefore, we first look at a simpler approach to the problem which is more general. Fortunately, the possibility of the interval-based computation helps us in simplifying our analysis of the convergence of SOFI values: The computation for each interval yields an estimate κ_j of the true value κ that is obtained by averaging over a large number of intervals $\kappa = \lim_{T \to \infty} \frac{1}{T} \sum_{j=1}^{T} \kappa_j$. As all intervals are qualitatively the same, the κ_j are independently identically distributed. To quantify the convergence of SOFI we simply need to look at the error of the average over a finite number T of interval computations compared to the true value. This approach to analyze the convergence is valid for all SOFI orders, as long as interval-based and single-block computation yield similar results, i.e. it is independent of the blinking dynamics as long as the underlying distributions can be normalized (i.e. have a finite time-scale). Let e be the index of a series of N_e experiments / simulations that we repeat under identical conditions. The squared error $\tilde{E}_S$ of the interval averaged cumulant value for one experiment is given by:

$$\tilde{E}_{S,e}(T) = \left(\frac{1}{T} \sum_{j=1}^{T} \kappa_{j,e} - \kappa \right)^2 \tag{2.25}$$

To make the error of different SOFI orders comparable, it is useful to normalize this error with the true value, i.e. to look at the relative deviation $E_{S,e}(T) = \kappa^{-2} \tilde{E}_{S,e}(T)$.

$$E_{S,e}(T) = \left(\frac{1}{\kappa T} \sum_{j=1}^{T} \kappa_{j,e} - 1 \right)^2 \tag{2.26}$$

We can interpret each interval's estimate as the true value plus noise $\kappa_j = \kappa + n_j$. Like the original values, this noise is independently identically distributed.

$$E_{S,e}(T) = \left(\frac{1}{\kappa T} \sum_{j=1}^{T} (\kappa + n_{j,e}) - 1 \right)^2 = \left(\frac{1}{\kappa T} \sum_{j=1}^{T} n_{j,e} \right)^2 \tag{2.27}$$

$$= \frac{1}{\kappa^2 T^2} \left(\sum_{j=1}^{T} n_j^2 + 2 \sum_{j=1}^{T} \sum_{k<j}^{T} n_{j,e} n_{k,e} \right) \tag{2.28}$$

The first term can be identified with the variance $\tilde{\sigma}_e^2$ of the intervals values $\kappa_{j,e}$

$$\tilde{\sigma}_e^2 = \frac{1}{T} \sum_{j=1}^{T} \left(\kappa_{j,e} - \kappa \right)^2 = \frac{1}{T} \sum_{j=1}^{T} n_{j,e}^2, \tag{2.29}$$

which yields

$$E_{S,e}(T) = \frac{\tilde{\sigma}_e^2}{\kappa^2 T^2} + \frac{1}{\kappa^2 T^2} \left(2 \sum_{j=1}^{T} \sum_{k<j}^{T} n_{j,e} n_{k,e} \right). \tag{2.30}$$

Averaging the squared error over a large number of experiments, the second term becomes zero, as the n_j are independent $\langle n_j n_k \rangle_e = \delta_{jk}$, and the average variance approaches its true value $\tilde{\sigma}$. We define the mean squared error E_{MS}

$$E_{MS}(T) := \frac{1}{N_e} \sum_{e=1}^{N_e} E_{S,e}(T) \tag{2.31}$$

which converges to

$$\lim_{N_e \to \infty} E_{MS}(T) = \frac{\tilde{\sigma}^2}{\kappa^2 T} \tag{2.32}$$

for a large number of experiments. We will omit the limit to keep the notation simple and assume N_e will be large enough to ensure convergence. Furthermore with the relative variance $\sigma^2 = \tilde{\sigma}^2 / \kappa^2$ we get

$$E_{MS}(T) = \frac{\sigma^2}{T}. \tag{2.33}$$

The derivation we just followed is helpful to understand, how we calculate this error from simulations: Take the true cumulant value κ (either by deriving it analytically or by approximating it using a very long simulation) and compute the squared error for one simulation following Eq. (2.26). Repeat this for a number of independent simulations N_e and average the individual errors to get the final result. For interpretation purposes it can be helpful to look at a simpler derivation of this result: The variance is additive for independent stochastic processes and homogeneous of degree two (Sec. 2.1). Following

this, the variance of the average $\overline{\kappa}_T = \frac{1}{T} \sum\limits_{j=1}^{T} \kappa_j$ is simply given by

$$\tilde{\sigma}_{\overline{\kappa}_T}^2 = \text{Var}\left(\overline{\kappa}_T\right) = \text{Var}\left(\frac{1}{T}\sum_{j=1}^{T}\kappa_j\right) = \frac{1}{T^2}\sum_{j=1}^{T}\text{Var}\left(\kappa_j\right) = \frac{T\tilde{\sigma}^2}{T^2} = \frac{\tilde{\sigma}^2}{T} \tag{2.34}$$

$$= \lim_{N_e \to \infty} \frac{1}{N_e} \sum_{e=1}^{N_e} \left(\overline{\kappa}_{N,e} - \kappa\right)^2 = E_{\text{MS}}(T) \cdot \kappa^2 \tag{2.35}$$

The mean squared error is simply the (relative) variance of the average SOFI value $\overline{\kappa}_N$ calculated from T available intervals. As SOFI values are computed from averaging a lot of statistically independent events in time, it is reasonable to assume that the individual values κ_j follow a distribution very close to a Gaussian, making the variance a well-suited error measure. Finally, we define the relative root mean square error of SOFI:

$$E_{\text{RMS}}(T) = \sqrt{E_{\text{MS}}(T)} = \sqrt{\frac{1}{e}\sum_{e=1}^{N_e}\left(\frac{1}{\kappa}\left[\frac{1}{T}\sum_{j=1}^{T}\kappa_{j,e} - \kappa\right]\right)^2} = \frac{\sigma}{\sqrt{T}}. \tag{2.36}$$

Note that the length of the regarded intervals is arbitrary for this derivation, as long as they are larger than the typical time-scale of blinking. In fact, considering our core assumption, that computing SOFI values from all available frames is practically identical to the interval-based approach, we can forget about intervals entirely and simply interpret T as the number of available frames. It is important to keep in mind that this assumption breaks down as the interval size comes close to the time-scale of blinking. If the standard deviation is computed using intervals of h frames, it can be converted by a simple rescaling $\sigma_{\text{Intervals}}/\sqrt{T} = \sigma_{\text{Frames}}/\sqrt{T \cdot h} \Rightarrow \sigma_{\text{Frames}} = \sigma_{\text{Intervals}} \cdot \sqrt{h}$, making the result independent of the interval length. To ease interpretation, all (relative) standard deviations in this section are given for T in units of frames.

Due to the underlying Gaussian distribution, the error value can be easily interpreted. For example, if $E_{\text{RMS}}(N = 400) = 0.1$, the SOFI value calculated from a simulation with 400 frames is with $\approx 68\%$ chance within a 10% error margin off the true value. Appendix C shows the analytical derivation of the true cumulant values κ, which are necessary to compute the error function, for the two-state blinking model used to generate the simulations. Different orders and experimental situations are compared using their E_{RMS} directly or the associated values of the standard deviation σ.

Figure 2.11 shows the RMS error calculated from $N_e = 1000$ noise-less simulations of a time trace capturing two equally bright independent emitters. The SOFI standard deviation σ for each case is determined by fitting curves to the data points using Eq. (2.36). The quantitative analysis confirms what was qualitatively seen in Fig. 2.9: For very similar on-/off-times, the 3rd order converges much slower than 2nd and 4th order. As the time-scales become more dissimilar, 3rd and 4th order yield almost similar performance, the 3rd order still being a bit worse. For off-times much longer than the on-time,
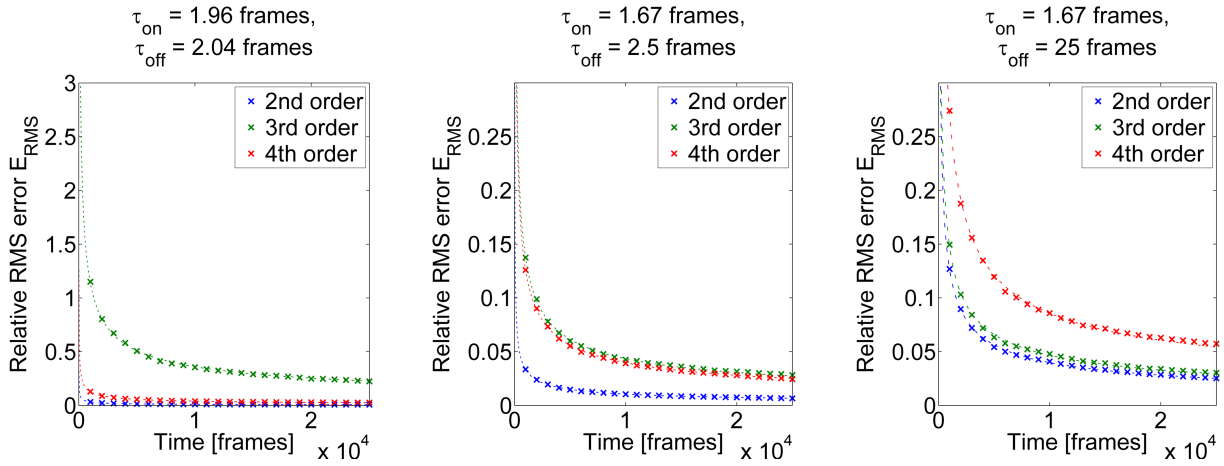
Figure 2.11: Comparison of the relative root-mean-square error for a noise-less time trace capturing two equally bright independent emitters using three different blinking dynamics and SOFI orders. Data points are calculated as averages from $N_e = 1000$ simulations. The interval size for the SOFI calculation is 1000 frames. Dotted lines show fitted curves using Eq. (2.36). Left: When on- and off-times are very similar, the 3rd order shows much slower convergence compared to 2nd and 4th order (note the scale on the y-axis). Mid: As both times become more dissimilar, the 3rd and 4th order are almost alike in terms of the error. The 2nd order is still much faster. Right: For off-times much longer than the on-time, 2nd and 4th order converge slower, making the 2nd and 3rd order perform similarly. Note that in the last situation, the time-scale of blinking ($\tau_{on} + \tau_{off}$) differs by roughly a factor of six from the others.

	$\tau_{on} = 1.96, \tau_{off} = 2.04$	$\tau_{on} = 1.67, \tau_{off} = 2.5$	$\tau_{on} = 1.67, \tau_{off} = 25$
σ, 2nd order	1.01	1.05	3.98
σ, 3rd order	35.96	4.43	4.67
σ, 4th order	3.99	3.97	8.57

	$\tau_{on} = 1.96, \tau_{off} = 2.04$	$\tau_{on} = 1.67, \tau_{off} = 2.5$	$\tau_{on} = 1.67, \tau_{off} = 25$
$T_{10\%}$, 2nd order	101	111	1587
$T_{10\%}$, 3rd order	129280	1885	2181
$T_{10\%}$, 4th order	1590	1572	7350

Table 2.1: SOFI standard deviation σ and number of necessary frames $T_{10\%}$ to converge to the 10% error margin $E_{RMS}(T_{10\%}) = 0.1$ when simulating a time trace of two overlapping independent emitters. The convergence in this case is shown in Fig. 2.11.

the 3rd order now converges almost as fast as the 2nd order, while the 4th order takes much longer than before. Note that 2nd and 4th order both converge slower in the last case than in the other two, which is expected, as the time-scale of blinking ($\tau_{on} + \tau_{off}$) is roughly a factor of six longer. Table 2.1 shows the calculated standard deviations and the resulting number of frames necessary to achieve 10% error. The 2nd order is usually quite fast, needing only about 100 frames in two of three cases. The exception occurs at long off-times, where convergence is roughly 16 times slower. As expected, the 3rd order is impractical for cases with $\tau_{on} \approx \tau_{off}$, needing almost 130000 frames. In the other cases it needs roughly 2000 frames to converge, about a factor 20 to the best case 2nd order. Despite the factor of six between the blinking time-scales in the two cases, the increase in necessary frames is minor. This further points

to the 3rd order being better suited for situations where the difference between τ_{on} and τ_{off} is large. The 2nd and 4th order show a slowdown in convergence which exceeds this factor of four (16 and 8 respectively) and thus seem less suited for this type of dynamics. Note that when comparing their best performance, the 4th order shows faster convergence than the 3rd.
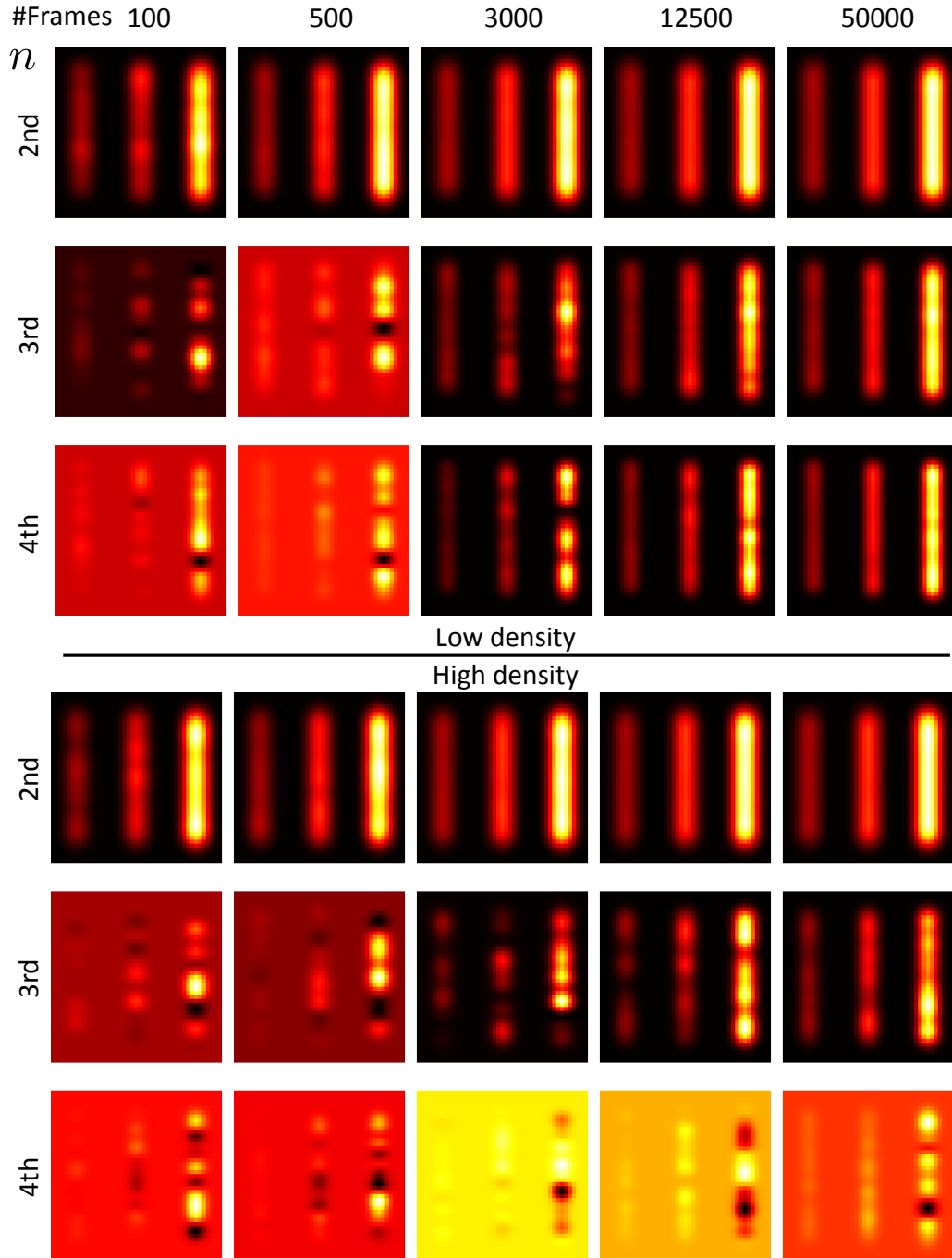
Figure 2.12: Comparison of SOFI convergence computed from noise-less simulations with different emitter densities (fast dynamics). The densities are $0.4 \frac{1}{\text{px}}$, $0.8 \frac{1}{\text{px}}$, $1.6 \frac{1}{\text{px}}$ in the low density case and $3.2 \frac{1}{\text{px}}$, $6.4 \frac{1}{\text{px}}$, $12.8 \frac{1}{\text{px}}$ in the high density case (factor of 8). While the second order is mostly unaffected by this difference, the third converges noticeably slower. The result of the fourth order has degraded so much that the 50000 frame image is comparable to 500 frames at the lower density.

Next we look at the density dependence of the convergence. Figure 2.12 illustrates the different orders in a low and a high density scenario (factor of 8 in density). The 2nd order is mostly unaffected

by the high density. The 3rd order convergence is much slower, roughly a factor of four compared to the low density case. Most affected is the 4th order result, which at 50000 frames with high density is comparable to the 500 frames low density result. Increasing the density by 8-fold thus slowed down its convergence by roughly 100-fold. The density dependence of the convergence seems to increase with the SOFI order, and high-density-samples require longer imaging times.
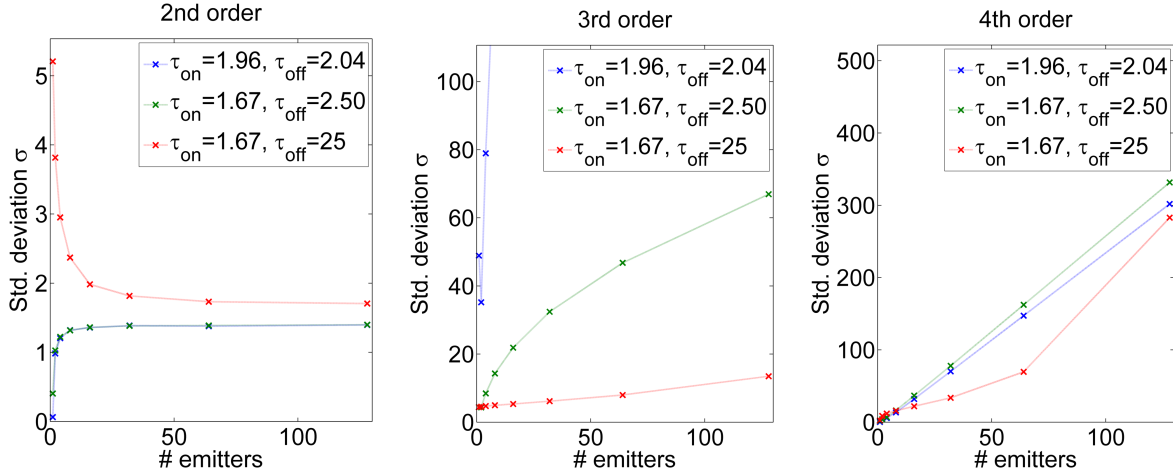
Figure 2.13: Density dependent convergence of SOFI standard deviation for different orders and blinking dynamics. For each density $N_e = 1000$ noise-less simulations were performed and the standard deviation extracted by a fit to the root mean squared error following Eq. (2.36). The interval size for the SOFI calculation is 1000 frames. For the second order and high off- compared to on-time, the error decreases when the number of simultaneously measured emitters increases. In every other case higher density leads to higher error. The second order error shows asymptotic behavior, while the higher orders increase fast within the observed range. Higher orders thus scale noticeably worse with density (note the values on the axis of ordinates). The third order scales worse as off- and on-time become more similar.

To analyze this further, time traces for different blinking dynamics and number of simultaneously measured emitters N_{em} were simulated. The relative SOFI standard deviation was calculated from $N_e = 1000$ simulations for every emitter count using Eq. (2.36). Figure 2.13 shows the results, which are summarized in Table 2.2 together with the number of frames $T_{10\%}$ necessary to lower the RMS error to 10%. The first peculiarity to notice is a qualitative difference in the behavior of the second order compared to the others. As the emitter density rises, the error asymptotically approaches a constant. Equally surprising, when the off-time is long, the error drops with density, making convergence faster, while the opposite is seen in the other two cases. In the first case, the number of necessary frames drops from 2711 for one emitter to 300 for 64 emitters. In the other cases roughly 200 frames are needed for the same error. Thus, a few hundred frames seem to be enough to get a good 2nd order result if the density of emitters is high enough. As before, the 3rd order is basically unusable for $\tau_{on} \approx \tau_{off}$, requiring at minimum few hundred thousand frames. It also scales worst with the density in this case, which improves as τ_{on} and τ_{off} get more dissimilar. The number of needed frames rises from 1949 for one emitter to 218940 for 64 emitters under fast dynamics, while only 6355 frames are necessary when the off-time is long. The 4th order scales worst, quickly requiring more than 10000 frames and going up to millions of frames for higher densities. It is important to note here, that even a few thousand frames are often hard to achieve experimentally, as the sample bleaches. This makes SOFI orders higher than

$\tau_{\mathrm{on}} = 1.96, \tau_{\mathrm{off}} = 2.04$							
$N_{\mathrm{em}} =$	1	2	4	8	16	32	64
$T_{10\%}$, 2nd order	1	101	145	174	184	192	189
$T_{10\%}$, 3rd order	239260	129280	623330	$2\cdot10^6$	$4.9\cdot10^6$	$10.6\cdot10^6$	$22.4\cdot10^6$
$T_{10\%}$, 4th order	6	1590	3218	18189	102340	492570	$2.2\cdot10^6$

$\tau_{\mathrm{on}} = 1.67, \tau_{\mathrm{off}} = 2.5$							
$N_{\mathrm{em}} =$	1	2	4	8	16	32	64
$T_{10\%}$, 2nd order	16	111	149	173	185	191	192
$T_{10\%}$, 3rd order	1949	1885	7139	20435	47845	105100	218940
$T_{10\%}$, 4th order	294	1572	4761	26084	134600	607290	$2.6\cdot10^6$

$\tau_{\mathrm{on}} = 1.67, \tau_{\mathrm{off}} = 25$							
$N_{\mathrm{em}} =$	1	2	4	8	16	32	64
$T_{10\%}$, 2nd order	2711	1587	870	561	393	329	300
$T_{10\%}$, 3rd order	1930	2181	2243	2475	2847	3802	6355
$T_{10\%}$, 4th order	595	7350	14055	24777	48185	12920	484000

Table 2.2: SOFI standard deviation σ and number of necessary frames $T_{10\%}$ to converge to the 10% error margin $E_{\mathrm{RMS}}(T_{10\%}) = 0.1$ when simulating a noise-less time trace of N_{em} overlapping independent emitters. The convergence in this case is shown in Fig. 2.13. Numbers comparable or smaller than the time-scale of blinking violate our assumption of large enough statistics, but indicate the convergence is not limited by the cumulant calculation.

the 2nd only applicable for scenarios of comparably low density, while the 2nd order seems to be feasible in almost all cases. Even at low densities, it is advisable to tune the blinking dynamics for the 3rd or 4th order to make convergence as fast as possible. It will be shown in Section 2.7.3 (e.g. Fig. 2.36), that the quality of images at high densities improves, even for higher orders, when the data is pre-processed to extract higher resolution. However, a more detailed analysis is required to quantify the effect.

To get more insights into the convergence behavior of SOFI, we derived the standard error for the 2nd order analytically. The only assumption here is, that the time trace $I(t)$ is long enough to measure the average brightness $\langle I \rangle$ with low error, i.e. that the recording time is longer than the typical blinking time-scale. Otherwise, the estimator for the second order cumulant is biased, which makes an additional correction necessary. For the derivation of this bias and the standard error, refer to Appendix D. A similar formula was published very recently by Wang *et al*, but they did not take into account that the recorded time series is correlated and assume that the variance is simply additive when multiple emitters are present. This is not the case, as we will see below. As a reminder, κ_i denote ordinary cumulants of i-th order and K_i the joint cumulant of i signals. The variance of the second order cumulant (i.e. the

variance of the variance) for a time-correlated stationary time series $I(t)$ of length T is given by

$$\text{Var}\left[\kappa_2\left(I(t)\right)\right] = \frac{1}{T}\left(\kappa_4\left(I(t)\right) + 2\kappa_2^2\left(I(t)\right)\right)$$
$$+ \frac{2}{T^2}\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i}\left[K_4\left(I(t), I(t), I(t+\tau), I(t+\tau)\right) + 2K_2^2\left(I(t), I(\tau)\right)\right],$$

(2.37)

where the first line is the variance for an uncorrelated time series and the second line is a correction for correlated signals. To compare to our previous results, we divide by κ_2^2 to get the relative variance. Inserting a two-state model, where k_+ is the transition rate from the off- to the on-state, k_- its counterpart, and N_{em} is the number of simultaneously measured (identical) emitters, this can be written as (see Appendix D for details)

$$\frac{\text{Var}\left[\kappa_2\left(I(t)\right)\right]}{\kappa_2^2(I(t))} = \frac{1}{T}\left(\frac{1}{N_{\text{em}}k_+k_-}\left[k_+^2 - 4k_+k_- + k_-^2\right] + 2\right)$$
$$+ \frac{2}{T^2}\left[\frac{1}{N_{\text{em}}}\frac{(k_+ - k_-)^2}{k_+k_-}\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i}G(\tau) + 2\cdot\left(1 - \frac{1}{N_{\text{em}}}\right)\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i}G^2(\tau)\right].$$

(2.38)

$G(\tau) = \langle\delta I(t)\delta I(t+\tau)\rangle / \langle\delta I^2(t)\rangle$ is the autocorrelation function of our process. If a continuous process is modeled, the double sums must be converted into integrals, and $G_{\text{cont}}(\tau) = \exp(-(k_+ + k_-)\tau)$ is an exponential function. For our discrete simulation, in which emitters can only be off or on during a complete frame, but never switch in between, we have $G(\tau) = [1 - (k_+ + k_-)]^\tau = a^\tau$ with $a := [1 - (k_+ + k_-)]$ (see Appendix C). The double sum can be solved using the known solution for the partial sum of the geometric series, yielding our final solution:

$$\frac{\text{Var}\left[\kappa_2\left(I(t)\right)\right]}{\kappa_2^2(I(t))} = \frac{1}{T}\left(2 + \frac{1}{N_{\text{em}}k_+k_-}\left[k_+^2 - 4k_+k_- + k_-^2\right]\right)$$
$$+ \frac{2}{T^2}\left\{\frac{1}{N_{\text{em}}}\frac{(k_+ - k_-)^2}{k_+k_-}\left[\frac{a(a^T - Ta + T - 1)}{(a-1)^2}\right]\right.$$
$$\left. + 2\cdot\left(1 - \frac{1}{N_{\text{em}}}\right)\left[\frac{a(a^{2T} - Ta^2 + T - 1)}{(a^2 - 1)^2}\right]\right\}.$$

(2.39)

Taking the square root of this, we get the relative standard deviation of second order SOFI, which is equivalent to the relative root-mean-square error we defined above (in the limit of infinite number N_e of averaged simulations).

$$E_{\text{RMS}}^{\text{theo.}}(T, N_{\text{em}}) = \frac{\text{Std}\left[\kappa_2\left(I(t)\right)\right]}{\kappa_2(I(t))} = \sqrt{\frac{\text{Var}\left[\kappa_2\left(I(t)\right)\right]}{\kappa_2^2(I(t))}}$$

(2.40)

Figure 2.14 shows the near perfect agreement between the theoretical convergence and our simulations with respect to the recording time and density dependence. Small residual deviations are caused by the finite sample size. Note that the result is independent of emitter brightness, because we assume noise-less conditions and normalize by κ^2. The formula explains the peculiar properties of the
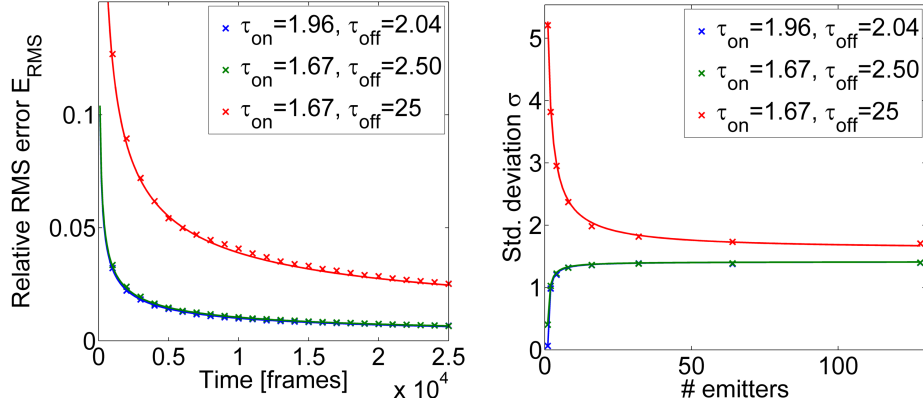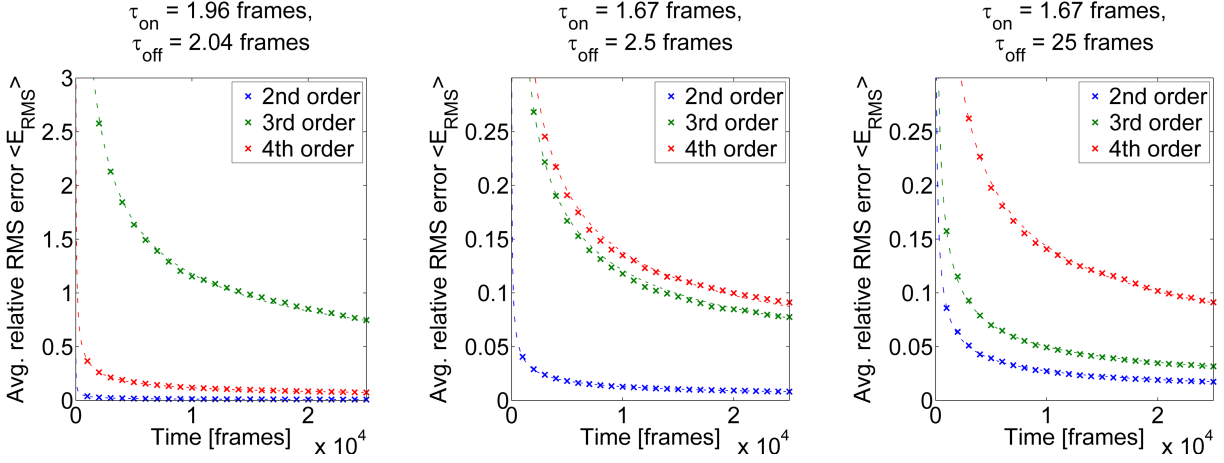
Figure 2.14: Comparison of the relative root mean squared error between simulations ($N_e = 1000$, interval size for SOFI 1000 frames) and the analytically derived formula (Eq. (2.39)) for 2nd order. Left: Recording time dependence. Right: Density dependence. As the standard deviation of simulations is computed using an interval size of 1000 frames, the theoretical values are computed for 1000 frames and then rescaled using $E_{\text{RMS}}^{\text{theo.}}(T = 1000, N_{\text{em}}) \cdot \sqrt{1000}$ (see Eq. (2.40)).

2nd order with respect to density: The first line of Eq. (2.39) (the variance for uncorrelated time-series) dominates the rest for usual imaging parameters. Its rate dependent part is negative if the ratio $r := k_-/k_+ = \tau_{\text{on}}/\tau_{\text{off}}$ is between the values $2 - \sqrt{3} < r < 2 + \sqrt{3}$, while all other parts of the equation are always positive. In this regime, a higher density leads to a larger variance and thus slower convergence. Outside this range the term becomes positive, lowering the variance with higher density and thus leading to faster convergence. Another important information that we can extract from the theoretical formula is that second order SOFI is suitable for arbitrary high densities, as the variance quickly approaches the asymptotic value $(2/T)$ when the density is increased. Thus, the density can be chosen high when imaging with second order only and must only be low when very fast imaging speeds should be achieved. In the last case the off- and on-time should be tuned to be as similar as possible.

Similar formulas can be derived for higher order cumulants, but quickly rise in complexity. As the expected variance of the n-th order cumulant is calculated using the squared deviation of the original estimator from its expected value, it depends on (joint) cumulants up to order $2n$. Given the complexity of the joint cumulant formulas (see Sec. 2.2) this should be done with symbolic algebra software.

Instead of analyzing individual time traces, we can also look at movies, which is closer to the experimental situation. The density can now vary locally and neighboring pixels are correlated by the spatial extent of the PSF. As the simplest quality measure we calculate the pixel-averaged error $\langle E_{\text{RMS}} \rangle$, where we exclude pixels with theoretical values smaller than 5% of the maximum of the theoretical image to avoid averaging over empty regions. Figure 2.15 shows the convergence of the average error for movies as shown in Fig. 2.9. The number of necessary frames for good quality is given in Tab. 2.3. The results are similar to the single-pixel, two-emitters simulations (Fig. 2.11). The main difference is a larger gap between the different orders and that the 3rd order now converges faster than the 4th for $\tau_{\text{on}} = 1.67$, $\tau_{\text{off}} = 2.5$. As we just saw, higher orders scale worse with the emitter density (Fig. 2.13), which is now spatially varying in the movie and higher in some parts than the two-emitter case. The calculated number of necessary frames matches the visual impression of the convergence illustrated in Fig. 2.9.

Figure 2.15: Comparison of the pixel-averaged relative root mean squared error for noise-less movies as shown in Fig. 2.9. Dotted lines show fitted curves (using Eq. (2.36)) through the data points that are calculated as averages from $N_e = 50$ simulations. The interval size for the SOFI calculation is 1000 frames. The results are similar to the single-pixel two-emitters simulations from Figure 2.11. The main difference is a larger gap between the different orders and the 3rd order now converging faster than the 4th for $\tau_{on} = 1.67$, $\tau_{off} = 2.5$ (center). This is caused by higher orders scaling worse with the emitter density (Fig. 2.13), which is spatially varying in the movie and higher in some parts than the two-emitter case.

|	$\tau_{on} = 1.96, \tau_{off} = 2.04$	$\tau_{on} = 1.67, \tau_{off} = 2.5$	$\tau_{on} = 1.67, \tau_{off} = 25$
$\langle\sigma\rangle$, 2nd order	1.23	1.28	2.74
$\langle\sigma\rangle$, 3rd order	117.70	12.13	5.00
$\langle\sigma\rangle$, 4th order	11.81	13.82	14.34

	$\tau_{on} = 1.96, \tau_{off} = 2.04$	$\tau_{on} = 1.67, \tau_{off} = 2.5$	$\tau_{on} = 1.67, \tau_{off} = 25$
$T_{10\%}$, 2nd order	150	165	751
$T_{10\%}$, 3rd order	1385300	14706	2500
$T_{10\%}$, 4th order	13957	19088	20575

Table 2.3: Pixel-averaged relative SOFI standard deviation $\langle\sigma\rangle$ and number of necessary frames $T_{10\%}$ to converge to the 10% error margin $\langle E_{RMS}(T_{10\%})\rangle = 0.1$ for simulated movies as shown in Fig. 2.9. The convergence of the error in these cases is shown in Fig. 2.15.

Thus, the calculated error seems to be a good quality criterion. The second order needs 751 frames for convergence in the long off-time case, indicating that convergence could be sped up by higher emitter density. In the other cases, around 160 frames are needed for the second order. The third order needs 2500 frames in the best situation (long off-time) compared to 14000 frames for the 4th order (for similar off-/on-time). This further illustrates why higher orders will not be suitable for most imaging scenarios which suffer from photobleaching. Note also that up to now, every simulation was performed without noise, making the number of frames needed here a lower bound for the experimental situation.

AC-SOFI images are usually computed using nonzero time-lags $\tau_i > 0 \, \forall i$. If all time-lags are chosen differently, only the correlations between different points in time are calculated, which eliminates the influence of temporally uncorrelated shot-noise. The simplest choice is to set $\tau_i = i \, \forall i$ with the n-th
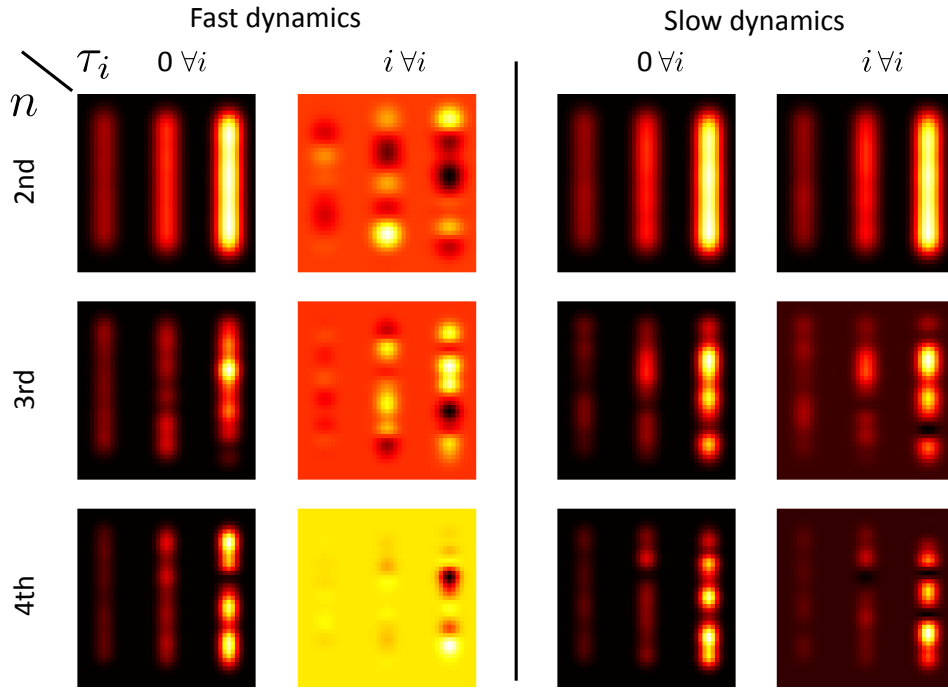
Figure 2.16: Comparison of SOFI result without and with time-lag under fast ($\tau_\text{on} = 1.67$ frames, $\tau_\text{off} = 2.5$ frames, $\frac{\tau_\text{on}}{\tau_\text{on}} \approx \frac{2}{3}$) and slow dynamics ($\tau_\text{on} = 6.67$ frames, $\tau_\text{off} = 10$ frames, $\frac{\tau_\text{on}}{\tau_\text{on}} \approx \frac{2}{3}$) for a movie of 3000 frames. It is easy to see that the quality of the result breaks down completely if a nonzero time-lag is used for signals with very fast dynamics. For slow dynamics, we see slightly different images with and without time-lag, which is caused by a slower convergence for nonzero time-lag.

order joint cumulant needing $(n-1)$ time-lags (equivalent to the first time-lag being zero). Although we will see in Sec. 2.5 how to correct zero time-lag images, we first have a look how the time-lag influences the resulting images. Figure 2.16 shows a comparison between zero and nonzero time-lag for fast and slow dynamics. It is easy to see that the quality of the result breaks down completely if a nonzero time-lag is used for signals with very fast dynamics. For slow dynamics, we see slightly different images with and without time-lag, which is caused by a slower convergence for nonzero time-lag. This intuitively makes sense, as the correlation between two points in time drops with their distance. For a time-continuous two-state model, the auto-correlation function is the exponential function $\exp(-(k_+ + k_-)\tau)$. As τ increases, the number of correlated events that we observe with this time-lag decreases, leading to worse statistics. Therefore, more measurement time is needed. If the dynamics are very fast, we almost never observe an emitter which stays in the on state for two frames. As the underlying process does not have memory, the time series is essentially uncorrelated in time, and no correlation can be measured even for the minimum time-lag $\tau = 1$ frame. The exposure time must be reduced well below the typical time-scale of blinking if nonzero time-lag should be used.

As time-lag-based computation is used to remove noise, we look at the effects of nonzero time-lag in the presence of shot noise. Figure 2.17 shows a comparison between a noise-less measurement and the same simulation with added shot noise. The zero time-lag images look less noisy, thus giving a better visual impression. However, a closer look reveals that the nonzero time-lag images have better resolution, yielding finer structures (compare the white dotted lines). The seemingly better quality is caused by a superposition with images from lower cumulant orders (including the time average), which is caused
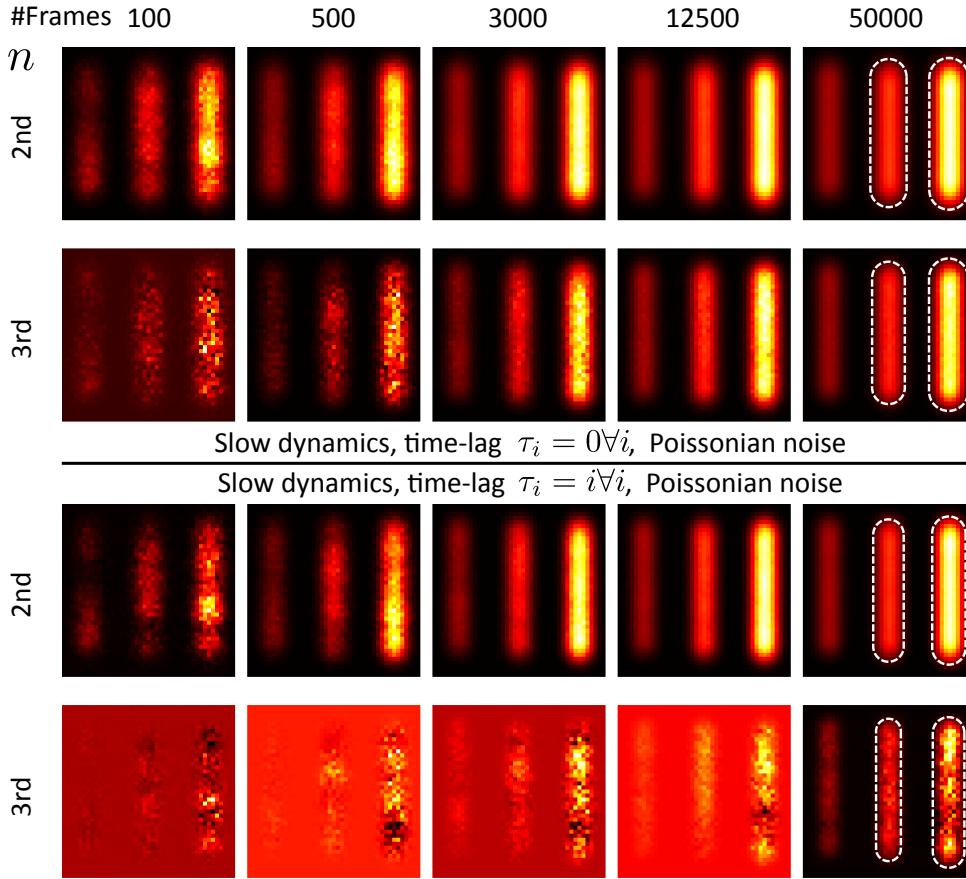
Figure 2.17: Comparison of convergence with zero and nonzero time-lag under slow dynamics (same simulation in both cases). The brightness of each emitter is set to 200 photons per frame. The zero time-lag images look less noisy, thus giving a better visual impression. However, a closer look reveals that the nonzero time-lag images are better resolved, yielding finer structures (compare white dotted lines). The seemingly better quality is the result of a superposition with lower resolution images caused by the noise. While both orders show slower convergence, the 3rd order seems to suffer more from noise than the 2nd. This is partly because the third order result depends on the observation of events correlated over three consecutive frames ($\tau_1 = 1, \tau_2 = 2$). Such events are rarer than events lasting two frames used for the second-order calculation.

by the noise. This lowers the resolution of the image for movies capturing dim emitters; for brighter samples the effect will be less noticeable (see Sec. 2.5). While both orders show slower convergence, the 3rd order seems to suffer more from noise than the 2nd. This is partly because the third order result depends on the observation of events correlated over three consecutive frames ($\tau_1 = 1, \tau_2 = 2$). Such events are rarer than events lasting two frames used for the second order calculation.

To isolate the impact of noise on the convergence, Fig. 2.18 shows a comparison between noise-free and shot-noise-limited measurements when using the same time-lags for both cases. The large scale image structure is preserved for the noisy measurements, showing differences only on the small scale. Thus, shot noise does not seem to have a large impact on the convergence to the final structure but rather worsens the SNR of the final image. This matches the experiences that we got from many simulations that were performed during this thesis. It is important to keep this in mind when analyzing the error in the same way we did before. Figure A.1 in the appendix shows the pixel-averaged RMS error for movies with different emitter intensities. The corresponding estimated number of needed frames is
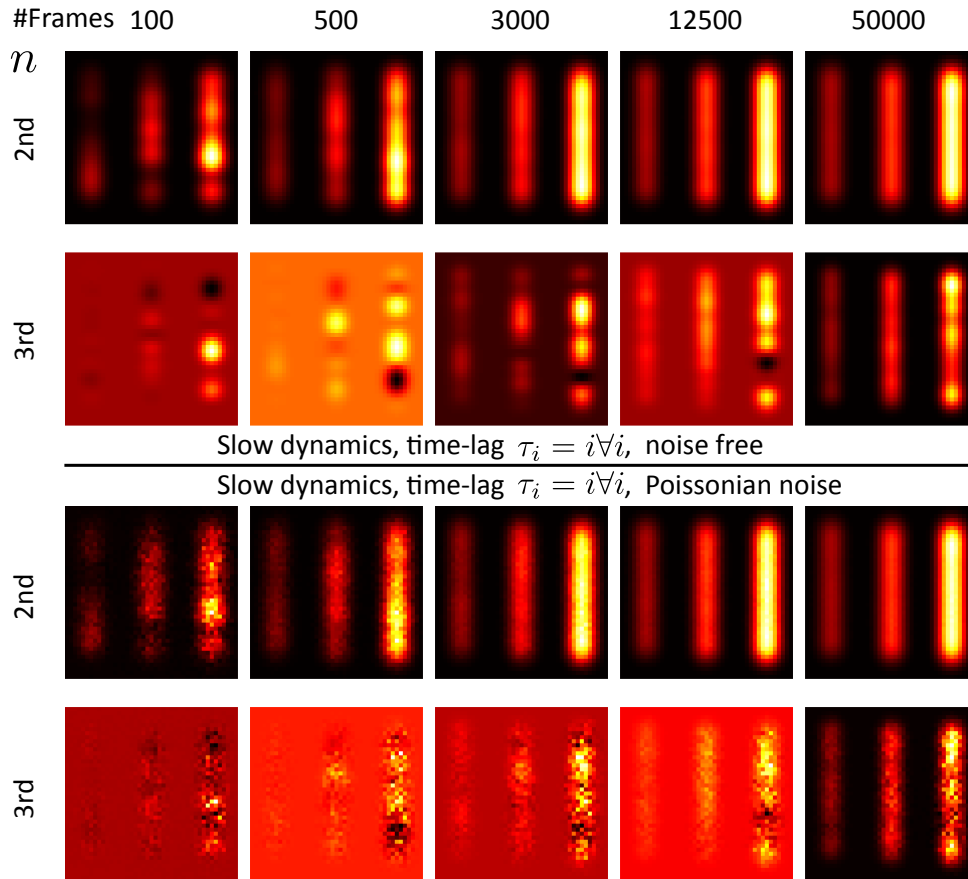
Figure 2.18: Comparison of convergence with nonzero time-lag in a noise-less measurement to the same measurement subject to shot noise. The brightness of each emitter is set to 200 photons per frame. The images do not change on a larger scale, but become more noisy on the small scale. Thus, shot noise does not seem to have a large impact on the convergence to the final structure but rather worsens the SNR of the final image.

given in Tab. A.1. Using the same methods as before, we would estimate a large increase in the number of frames needed to achieve good quality. Although the variance of individual pixel values is indeed increased considerably[2], the perceived image quality is much better. The spatial correlation introduced by the size of the PSF preserves the large-scale structure of the image in the presence of higher variations on the single pixel level. We suggest using the noise-free error as a basis when estimating the number of necessary frames in practice, and to consider the finite emitter brightness simply as a separate source of noise.

To sum up, from simulations and theoretical modeling of two-state blinking dynamics we learned the following about the properties of SOFI up to the 4th order. Note that the term "density" here must be understood as the number of emitters whose signals overlap on camera pixels (i.e. it will be bigger than the density of fluorophores because of the extent of the PSF).

- The performance of SOFI depends on the blinking dynamics.

- Fast dynamics are generally to be preferred, as they minimize imaging time / improve the time

---

[2]For the theoretical variance of cumulants for non-correlated time series (i.e. in the approximation of very fast dynamics in our case) subject to Poisson noise (*factorial cumulants*, see Sec. 2.5) see reference [43].

resolution of the measurement.

- Computation of SOFI images by subdivision of a movie into non-overlapping intervals which are individually processed and averaging the results is similar to processing the whole movie at once if the interval length is multiple times larger than the typical time-scale of blinking. This is important, as this procedure can be used to correct for bleaching in the sample as explained in Sec. 2.2. As was shown, this property can be used to simplify the convergence analysis of SOFI.

- Higher orders need more frames to converge and scale worse with rising emitter density. The blinking dynamics do not only influence the speed of convergence for a fixed number of overlapping emitters, but also determines how the orders scale with the density. This makes tuning of the blinking dynamics important for higher order SOFI. While the 2nd and 4th order work best if the off- and on-time are the same, the 3rd order gives the worst results in this case and needs them to be dissimilar. This is in accordance with the findings of [27], where it was reported that even orders require similar off- / on-times, while odd orders require them to be dissimilar. Section 2.7.3 (e.g. Fig. 2.36) shows that the quality of images at high densities improves, even for higher orders, when the data is pre-processed to extract higher resolution. However, a more detailed analysis is required to quantify this effect.

- The 2nd order shows fast convergence, usually between 150 and 300 frames. It is suitable for arbitrarily high densities as the speed of convergence quickly approaches an asymptotic value when the density is increased. Only when very fast imaging speed is desired, the density should be kept low while the fluorophores must be tuned for comparable off-/on-times. Surprisingly, when the ratio $r = \tau_{\mathrm{on}}/\tau_{\mathrm{off}}$ is outside the range $2 - \sqrt{3} < r < 2 + \sqrt{3}$, higher densities lead to faster convergence.

- If 3rd order SOFI is targeted, tuning the blinking dynamics is very important. Otherwise the convergence speed is too slow and scales very badly with rising density. In a well tuned environment, between 1000 and 6000 frames seem plausible, depending on the density. The 4th order quickly needs more than 10000 frames for acceptable quality and is extremely sensitive to the emitter density. In the best case, where density is kept very low (e.g. less than four overlapping emitter per pixel) and the blinking is optimally tuned, around 2000-3000 frames might be enough to achieve acceptable results. This is already very demanding for most available fluorophores and can not be achieved in most experimental situations. Even higher orders are infeasible in almost every imaginable case except for special non-bleaching samples like fluorescent diamond centers at not too high density. In most cases, only the 2nd order seems to be applicable.

- Zero time-lag calculations causes a superposition of lower cumulant order to be added to the final image if shot noise is present, which lowers the resolution of the image especially when imaging dim samples (see Sec. 2.5 for details).

- Non-zero time-lags can only be used if the exposure time is shorter than the typical blinking time-scale, and need more frames. For typical fluorophores, the chance to spend a certain time in the

on-state decreases (drastically) with the duration. Thus, for larger time-lags less of these events are observed, which decreases the available statistics and increases convergence time.

- Shot noise does not influence the convergence of the SOFI image on larger spatial scales but rather decreases its SNR. To estimate the number of necessary frames in practice, analyzing the noise-less situation will usually suffice and gives a good lower bound.

2.4   FOURIER SOFI

The resolution of SOFI increases with the order of the calculated cumulant. As shown in the previous section, higher order cumulants need more measurement time to yield trustworthy results and are more sensitive to increasing emitter density. In principle, however, SOFI allows arbitrarily high spatial resolution limited only by the measurement time and thus the photostability of the sample. The raw input data for the SOFI calculation are frames with pixels of finite size and this discretization is kept for the final SOFI image. In order to fully capture a signal with discrete samples, the sampling rate must be twice as high as the highest frequency component of the signal, a condition known as the *Nyquist sampling criterion*. Consequently, the pixel size of a SOFI image must be at least two times smaller than the desired resolution, or information is lost in the process. One solution to this problem is the XC-SOFI algorithm as first proposed by Dertinger at al. in ref. [40], and later perfected by Geissbuehler et al. in ref. [27, 28, 39]. Here, virtual pixels are generated between the physical pixels of the recording camera by cross-correlating signals from the original pixel grid (see Eq. (2.11)). Although this allows to generate images with arbitrarily small pixels, the distance of the correlated pixels as well as the axial positions of the fluorophores lead to different <u>types</u> of virtual pixels with different weighting factors for their brightness. This results in grid-like artifacts which have to be corrected and usually rely on the assumption of a Gaussian PSF and that all fluorophores are located in the focal plane. Furthermore, the number of possible high-order cross-correlations for obtaining one desired virtual pixel increases exponentially and there is no unambiguous way to choose which cross-correlation geometry will be optimal.

This section presents a much simpler solution which is straightforward, exact and based on the fact that the Optical Transfer Function of a microscope has a finite support. Thus, when performing a Fourier transform on an image, the Fourier amplitudes will drop to zero at large values of the Fourier vector. Padding the Fourier-transformed image with zeros does not change or alter its information content, but after back-transforming such a padded image into real space, one obtains an image with increased pixel resolution. This procedure is a method of exact interpolation of band-limited signals, without the introduction of artifacts. Using this scheme, the pixel size of the original frames of a recorded movie can be adapted to the desired spatial resolution delivered by SOFI. We call this combination of SOFI with Fourier-transform-based interpolation Fourier-SOFI or fSOFI.

Parts of this section have been published in the article [44]:

Stein, S.C.; Huss, A.; Hähnel, D.; Gregor, I.; Enderlein, J.; "Fourier interpolation stochastic optical fluctuation imaging", <u>Optics Express</u>, 23, 16154-16163, 2015.

Real space   Fourier space   Real space
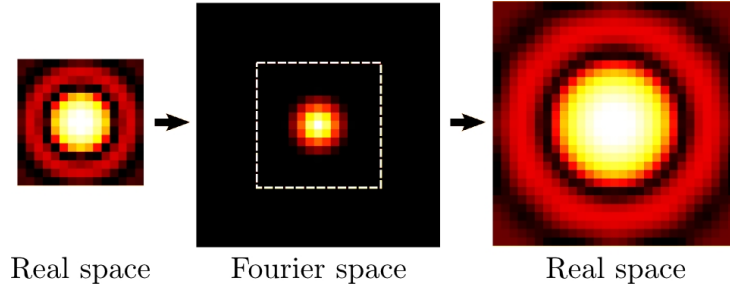
Figure 2.19: Algorithm of Fourier interpolation: Each frame (example: single emitter represented by an airy disc) is first Fourier-transformed. For sufficiently small detector pixel sizes, the Fourier transform is zero on the borders due to finite support of the OTF. The Fourier transform can thus be padded with zeros without changing the frequency information, here separated by a dashed white line from the original Fourier transform. Transforming back into real space gives an artifact-free image with more pixels, where each "virtual" pixel corresponds to a smaller area than that of the original detector pixels. [This image has been published in the article [44].]

When transforming an image with $N \times N$ pixels using the discrete Fourier transform, the relation between the frequency spacing Δf in Fourier space and the spatial sampling frequency (i.e. pixel size) Δx along one direction is simply given by

$$\Delta x = \frac{1}{N \Delta f} \Leftrightarrow \Delta f = \frac{1}{N \Delta x}. \tag{2.41}$$

If the pixel size of the recording camera was properly adjusted to be smaller than half the available spatial resolution, the OTF drops to zero within the observed interval, and symmetrically padding the Fourier transform with an overall number of ΔN_0 zeros does not alter its information content. Note that images subject to noise (e.g. shot noise) usually show variations on the per-pixel level and thus contain frequencies above the cutoff of the microscopes OTF, up to the highest available frequency. Although this violates the assumption of a band-limited signal needed for exact interpolation, zero-padding does not influence the noise-less portion of the signal, and the contributions from noise are removed by the SOFI calculation, leaving an exact result if the measurement is long enough. Back-transformation into real space after zero padding yields a new pixel size

$$\Delta x_{\text{new}} = \frac{N}{N + \Delta N_0} \tag{2.42}$$

An example of this method of interpolation is illustrated in Figure 2.19. The interpolation is applied to every frame of the original movie, before SOFI processing.

The new pixel size can be arbitrarily small, but must be chosen small enough to prevent information from being lost. For a recording with the original PSF $U(\mathbf{r})$, the SOFI image of n-th order produces an image with the the new PSF $U(\mathbf{r})^n$. As was already shown in Eq. (2.5), the Fourier transform

$$\mathcal{F}\left(U(\mathbf{r})^n\right) = \underbrace{\tilde{U}(\mathbf{k}) \circledast \cdots \circledast \tilde{U}(\mathbf{k})}_{n\text{-1 times}} \tag{2.43}$$
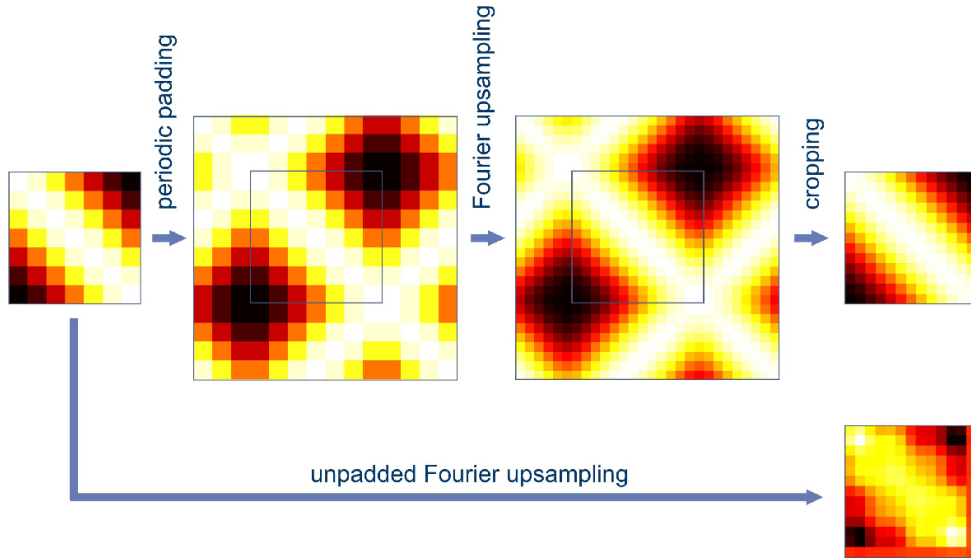
Figure 2.20: Example of periodic padding of an image for preventing boundary artifacts upon Fourier up-sampling: Instead of Fourier-upsampling the original image, one first extends the image with mirror symmetric half-copies to obtain a larger and continuously periodic image (periodic padding), which is then Fourier-upsampled, and then cropped back to the original size. The comparison between both results (unpadded Fourier upsampling, right bottom, versus padded Fourier upsampling, right top) clearly shows that the padding procedure leads to a perfectly artifact-free upsampling of the original image.[This image has been published in the article [44].]

is the $(n-1)$-times auto-convolution of the OTF. For any function with finite support, the cutoff frequency increases linearly with the number of auto-convolutions. If the original pixel size was well adjusted to the available resolution, the new pixel size must be chosen n times smaller to match the new cutoff frequency.

Note that the Fourier space is expanded naturally if the SOFI calculation is performed in Fourier space. For example the second order can be calculated using

$$\left\langle \delta F^2(\mathbf{r}, t) \right\rangle \equiv \left\langle \mathcal{F}^{-1}\left[\mathcal{F}(\delta F(\mathbf{r}, t)) \circledast \mathcal{F}(\delta F(\mathbf{r}, t)) \right] \right\rangle. \tag{2.44}$$

The discrete convolution of two signals of length N and M has $N + M - 1$ points, resulting in a larger Fourier space image that produces finer pixels when transformed back into real space. It is also important to notice that explicitly performing the convolutions to compute δF^2 is identical to a Fourier interpolation of δF via zero padding and taking the square after back-transforming into real space. For this to hold, the padding for the SOFI order n must be chosen as such:

$$\Delta N_0 = (n-1) \cdot (N-1) \tag{2.45}$$

$$\Delta N_0^{\text{post}} = \lfloor \Delta N_0/2 \rfloor \tag{2.46}$$

$$\Delta N_0^{\text{pre}} = \Delta N_0 - \Delta N_0^{\text{pre}}. \tag{2.47}$$

Here ΔN_0^{pre} is the number of zeros prepended to the part of the spectrum carrying the "negative" frequencies and ΔN_0^{post} the number appended to the positive direction. This padding must be applied to each spatial dimension. As zero padding requires only one copy operation for each of the N^d data points
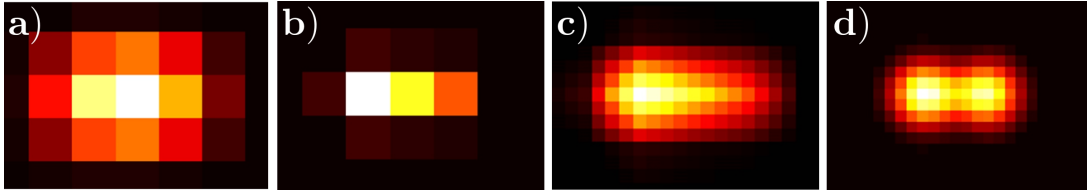
Figure 2.21: Comparison between fSOFI and interpolation of SOFI image using a simulation of two close emitters. a) Average of movie. b) 4$^{\text{th}}$ order SOFI image. c) Linear interpolation of b. d) 4$^{\text{th}}$ order fSOFI image. The coarseness of the pixel grid prevents standard SOFI to resolve the emitters. In contrast to interpolation, the fSOFI image captures true information and correctly resolves the emitters. [This image has been published in the article [44].]

of a d-dimensional image, compared to $\mathcal{O}(N^{2d})$ arithmetic operations for direct convolution in Fourier space, it is computationally much cheaper and should always be the preferred method.

Zero padding is also equivalent to sinc-interpolation in real space[45]. A continuous function $f(x)$ can be perfectly reconstructed from samples f_n spaced Δx apart if its Fourier transform $\tilde{f}(k)$ is confined to the region $|\tilde{f}(k)| < 1/(2\Delta x)$. The reconstruction formula [45]

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \cdot \operatorname{sinc}\left(\frac{x - n\Delta x}{\Delta x}\right) = \left(\sum_{n=-\infty}^{\infty} f_n \cdot \delta(x - n\Delta x)\right) \circledast \operatorname{sinc}\left(\frac{x}{\Delta x}\right) \qquad (2.48)$$

uses convolution to perform the interpolation. Usually the sinc function is cut off to include only a few of the closest pixels in practice. For small upsampling factors, this approximation can be faster than the Fourier-based approach.

The discrete Fourier transform (DFT) intrinsically assumes an image which is periodic in all directions. Images which exhibit discontinuities at the borders when stacked periodically suffer from wave-like artifacts in these border regions when Fourier interpolation is performed. The problem can be alleviated by padding the image with mirror-symmetric half-copies of itself (see Fig. 2.20) before Fourier interpolation. In most cases this effect is negligible, especially if the sample of interest is surrounded by label-free space.

In contrast to a simple interpolation of the final image, which cannot restore lost information, the interpolation in fSOFI must be performed on the original recording before SOFI processing, to prevent the loss of information. Figure 2.21 shows a comparison between simple interpolation of the result and fSOFI for a simulation of two very close emitters. It is clear that fSOFI captures the improved spatial resolution that can not be displayed on the original pixel grid and resolves the emitters, which is not the case for post-processing interpolation.

The improved image fidelity as well as correct subpixel positioning is demonstrated on two simulated examples in Fig. 2.22 and compared to conventional SOFI and raw XC-SOFI. The raw cross-correlation images exhibits the mentioned artifacts which are inherent to the method and have to be corrected in post-processing. An experimental proof of principle is shown in Fig. 2.23 using a rat hippocampal neu-
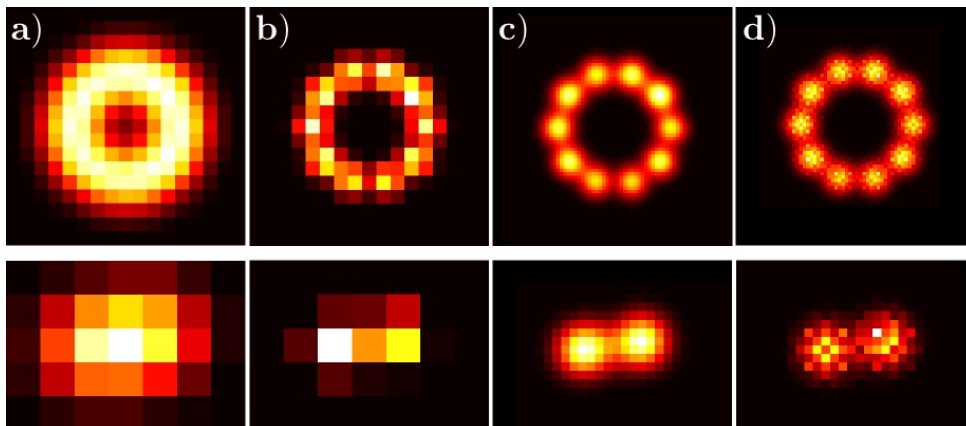
Figure 2.22: Two examples from a simulated SOFI experiment illustrating the presented Fourier interpolation in combination with SOFI. Top panels: 10 emitters in a ring. Bottom panels: Two emitters with sub-pixel shift (1.75 px, 0.25 px). a) Time average of all frames. b) 4th order SOFI. c) 4th order SOFI from 4x-super-sampled frames. d) Raw 4th order cross-correlation SOFI. It is easy to see that the Fourier interpolation improves image quality and exposes the sub-pixel positioning of the emitters without producing artifacts visible in the raw cross-correlation approach. [This image has been published in the article [44].]
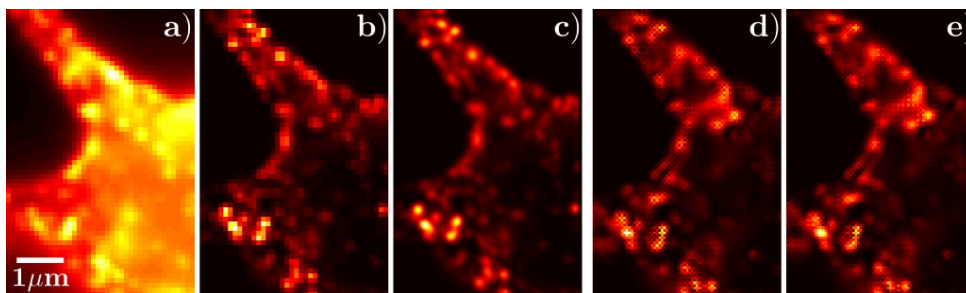
Figure 2.23: Experimental example using fSOFI. Rat hippocampal neuron with neurotransmitter receptor subunit $GABA_BR1$ immunostained with commercial quantum dots QD_{525} (Invitrogen). The raw stack of images contains 3000 frames recorded at 20 Hz frame rate. Fluorescence was excited at 401 nm wavelength and about 20 W/cm^2 using an laser (Cube401, 100 mW, Coherent). The microscope used was a commercial epi-fluorescence microscope (IX-71, Olympus) equipped with a 1.4NA oil-immersion objective (UPlanSApo, Olympus), and an EMCCD (DU-897-CS0-BV, Andor). Magnification was chosen in such a way that the effective pixel size of the recorded images was 100 nm. (a) Time average of original images. (b) 2nd order SOFI. (c) 2nd order SOFI with 3x Fourier interpolation. The last two images show the raw 2nd cross-correlation SOFI (d) and the artifact corrected image (e) calculated using publicly available software [46]. The recording was performed by *Anja Huss*. [This image has been published in the article [44].]

ron with neurotransmitter receptor subunit $GABA_BR1$ immunostained with commercial quantum dots QD_{525} as a sample. The recording was performed by *Anja Huss*. It is clear that fSOFI improves image fidelity and resolution compared to the conventional calculation. For comparison, we also show cross-cumulant based calculations and performed artifact correction using publicly available software [46]. Note that the corrected image still shows artifacts, probably because of the assumption that all emitters are located in the focal plane.

To sum up this section about artifact-free subpixel generation for SOFI,

- As conventional SOFI calculations yield results on the pixel-grid defined by the recording camera, it is possible that the improved spatial resolution cannot be fully captured due to the coarseness of the grid, and information is lost. This is the case if the final resolution is better than two times the pixel size (the Nyquist criterion).

- fSOFI is a conceptually simple way to produce images on a finer pixel grid using Fourier interpolation on the original frame stack before SOFI processing. Fourier interpolation is carried out by zero-padding in the Fourier domain and transforming back to real space.

- For band-limited (finite support in Fourier space) signals, the method produces exact results and is artifact-free. This is in contrast to other currently used methods which rely on cross-correlation of different pixel positions.

2.5   NOISE CORRECTION FOR AUTO-CUMULANTS WITH ZERO TIME-LAGS

Measurements of physical quantities are usually subject to some form of noise. In fluorescence microscopy the quantity of interest is the brightness of the fluorophores, which is inherently noisy as the emission of photons is a probabilistic process. This leads to a Poisson distribution for the number of measured photons. The detection of the photons by the imaging camera produces additional noise, determined by the working principle of the measurement as well as its internal signal processing (e.g. signal amplification). The computation of ordinary cumulants of noisy signals superimposes the noise-less value of interest with additional contributions produced by the noise. As noise present in movies produced by fluorescence microscopy is mostly *shot noise*, i.e. it is neither correlated in time nor space, these contributions can be removed by using joint cumulants to correlate different time points (AC-SOFI) or spatial positions (XC-SOFI). However, the signal-to-noise decreases with increasing spatial distance of correlated pixels or distance of correlated frames in time, lowering the image quality. Furthermore, while correlation of different points in time cannot be used when imaging very fast blinking samples (it decays too fast), spatial cross-correlation produces artifacts dependent on the local PSF which have to be corrected later. Thus, it is desirable to understand the influence of noise on ordinary cumulants and use this knowledge to correct for it.

In this section, we take a look at the expectancy value of ordinary cumulants of noisy signals and compute formulas that estimate the cumulant value of the noise-less signal. Although two specific examples are shown, one for Poisson noise and one for the measurement with an EMCCD camera (which are

currently used by many research groups), the general approach is the same for any noise distribution. The correction can also be combined with fSOFI (Section 2.4) without any alterations.

Let $x(t)$ be a randomly fluctuating signal, e.g. the intensity of a blinking fluorophore, and $y(t)$ the noise contaminated signal that is measured:

$$y(t) = y\left(x(t)\right) = \text{Noise}\left[x(t)\right]. \tag{2.49}$$

The chance to measure a certain value y can be calculated by summing up the chances of all possible situations which can yield this value:

$$P(y) = \int P(x)P(y|x)\,\mathrm{d}x. \tag{2.50}$$

Here $P(y|x)$ denotes the conditional probability to measure y given that the value of the underlying process is x. This is simply the probability distribution function of our noise, i.e. the distribution that would be measured if $x(t)$ was simply a constant. Using this relationship, the n-th moment of y reads

$$\langle y^n \rangle = \int P(y)\,y^n\,\mathrm{d}y \tag{2.51}$$

$$= \int \left[\int P(x)P(y|x)\,\mathrm{d}x\right] y^n\,\mathrm{d}y \tag{2.52}$$

$$= \int P(x)\left[\int P(y|x)y^n\mathrm{d}y\right]\,\mathrm{d}x \tag{2.53}$$

$$= \int P(x)\,\langle y^n \rangle_{\text{Noise,x}}\,\mathrm{d}x \tag{2.54}$$

$$= \int P(x)\,m_n(x)\,\mathrm{d}x \tag{2.55}$$

$$= \langle m_n(x) \rangle \tag{2.56}$$

The n-th moment of the noisy measurement y of a random signal x is the average over the n-th moment $m_n(x)$ of the noise distribution for fixed x. To connect the cumulants of the noisy and the noise-less signal, we compute the relationship between their moments and replace "noisy moments" with "noise-less moments" in the cumulant formulas.

Poisson noise

The first and second moment of a Poisson distribution with mean λ are

$$m_1^{\text{Poiss}}(\lambda) = \lambda \tag{2.57}$$

$$m_2^{\text{Poiss}}(\lambda) = \lambda^2 + \lambda \tag{2.58}$$

Thus, the first two moments of a signal subject to Poisson noise are:

$$\langle y \rangle = \int P(x) \, m_1^{\text{Poiss}}(x) \, dx \tag{2.59}$$

$$= \int P(x) \, x \, dx = \langle x \rangle \tag{2.60}$$

$$\langle y^2 \rangle = \int P(x) \, m_2^{\text{Poiss}}(x) \, dx \tag{2.61}$$

$$= \int P(x) \left(x^2 + x \right) \, dx = \langle x^2 \rangle + \langle x \rangle \tag{2.62}$$

For the Poisson distribution, the n-th moment of y is always a linear combination of equal or lower order moments of x. Using these moments, the first two cumulants are

$$\kappa_1^{\text{Poiss}} = \langle y \rangle = \langle x \rangle = \kappa_1 \tag{2.63}$$
$$\kappa_2^{\text{Poiss}} = \langle y^2 \rangle - \langle y \rangle^2 \tag{2.64}$$
$$= \langle x^2 \rangle - \langle x \rangle^2 + \langle x \rangle \tag{2.65}$$
$$= \kappa_2 + \kappa_1^{\text{Poiss}}. \tag{2.66}$$

Rearranging this to calculate the cumulants of the noise-less signal κ_i gives

$$\kappa_1 = \kappa_1^{\text{Poiss}} \tag{2.67}$$
$$\kappa_2 = \kappa_2^{\text{Poiss}} - \kappa_1^{\text{Poiss}}. \tag{2.68}$$

Using this scheme, we can derive rules how to compute the cumulants of the signal of interest from the cumulants of the noisy measurements. Note that the superposition with lower orders caused by the noise (Eq. 2.65) leads to a degradation of the resolution of SOFI images. However, as the image intensity scales with the power of the order, the highest order dominates for brighter fluorophores and the effects of noise will be less noticeable.

The linear combinations of ordinary cumulants of the noisy signal that yield the noise-less ordinary cumulants are called *factorial cumulants*. Like ordinary cumulants, factorial cumulants are additive for statistically independent variables and homogeneous. For a Gaussian distribution all cumulants above the second order vanish. Similarly, for the Poisson distribution all factorial cumulants above the first order are equal to zero, which effectively eliminates the contribution of shot-noise. The n-th order factorial cumulant reads [47]

$$\kappa_n = \sum_{i=1}^{n} s(n, i) \cdot \kappa_j^{\text{Poiss}} \tag{2.69}$$

with the *Stirling number of the first kind* $s(n, i)$, which can be generated using the recursion relation

$$s(n, i) = -(n-1) \cdot s(n-1, i) + s(n-1, i-1) \tag{2.70}$$

$$\text{with } s(0,0) = 1, \; s(0,k) = s(k,0) = 0 \; \forall k > 0, k \in \mathbb{N}.$$

The first four factorial cumulants are given in Appendix E. As a side note, the variance of factorial cumulants for uncorrelated time-series (i.e. in the limit of very fast dynamics) is calculated in [43]. The variance represents the error of estimation of the cumulant and is useful for thinking about the convergence of the cumulant estimator. As explained in Section 2.3, the error of the noise-less cumulant should probably be preferred when, for example, estimating the number of necessary frames in practice.

EMCCD CAMERA NOISE

Electron Multiplying Charge-Coupled Device (EMCCD) cameras are currently used by many research groups when imaging single molecules. The photoelectrons generated from light hitting the detector are amplified by an additional gain register before read-out. The gain register consists of hundreds of identical stages which accelerate the incoming electrons using high voltage, leading to electron multiplication via impact ionization. Thus, every photoelectron generates a high signal and the additional read-out noise becomes negligible in comparison. Although this process allows the detection of single photons, the additional random processes widens the measured intensity distribution at constant photon flux by approximately a factor of $\sqrt{2}$ compared to a pure Poisson process, which lowers the signal-to-noise in the same way as halving the sensors quantum efficiency. However, at low photon fluxes the advantages far outweigh this handicap.

The distribution of the measured input counts n_{ic} of the camera is given by [48]

$$P_{EMCCD}(n_{ic}; \lambda, g, s, \sigma_{ro}) = \tag{2.71}$$

$$\begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{ro}} \cdot \exp\left(-\lambda - \frac{(s \cdot n_{ic})^2}{2\sigma_{ro}^2}\right) + \frac{2}{g} F_\chi(2\lambda; 4, 2s \cdot n_{ic}/g) & n_{ic} > 0 \\ \frac{1}{\sqrt{2\pi}\sigma_{ro}} \cdot \exp\left(-\lambda - \frac{(s \cdot n_{ic})^2}{2\sigma_{ro}^2}\right) & n_{ic} \leq 0 \end{cases}$$

where λ is the mean number of detected photons, g the EM-gain (average number of output electrons per photoelectron), s the sensor sensitivity (electrons per input count) and σ_{ro} the standard deviation of the read out noise. $F_\chi(2\lambda; 4, 2s \cdot n_{ic}/g)$ denotes the non-central χ^2-distribution for 2λ with 4 degrees of freedom and the non-centrality parameter $2s \cdot n_{ic}/g$. Note that usually a constant bias is added during readout by the camera to prevent the occurrence of negative input counts, which simply shifts the distribution. The photon conversion factor $a := g/s$ converts from (average) measured photons to input counts and varies with the settings of the camera. We computed the first four moments of the

distribution using the software *Mathematica*:

$$m_1^{\text{EMCCD}}(\lambda) = a \cdot \lambda \tag{2.72}$$

$$m_2^{\text{EMCCD}}(\lambda) = a^2 \cdot \left(\lambda^2 + 2\lambda\right) \tag{2.73}$$

$$m_3^{\text{EMCCD}}(\lambda) = a^3 \cdot \left(\lambda^3 + 6\lambda^2 + 6\lambda\right) \tag{2.74}$$

$$m_4^{\text{EMCCD}}(\lambda) = a^4 \cdot \left(\lambda^4 + 12\lambda^3 + 36\lambda^2 + 24 \cdot \lambda\right). \tag{2.75}$$

The EMCCD distribution is often simplified using a shifted Poisson $P_{\text{Poiss}}(k + \lambda; 2\lambda)$ with mean λ and variance 2λ. This approximation usually works very well but gives slightly different higher moments. Using these moments, we can compute the linear combinations to extract the cumulants of the underlying noise-less signal. For example, the first two cumulants of the noisy signal are

$$\kappa_1^{\text{EMCCD}} = \left\langle m_1^{\text{EMCCD}}(x) \right\rangle = a \left\langle x \right\rangle \tag{2.76}$$

$$\kappa_2^{\text{EMCCD}} = \left\langle m_2^{\text{EMCCD}}(x) \right\rangle - \left\langle m_1^{\text{EMCCD}}(x) \right\rangle^2 \tag{2.77}$$

$$= a^2 \left(\left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 + 2 \left\langle x \right\rangle\right) \tag{2.78}$$

$$= a^2 \left(\left\langle x^2 \right\rangle - \left\langle x \right\rangle^2\right) + 2a^2 \cdot \left\langle x \right\rangle \tag{2.79}$$

$$= a^2 \cdot \kappa_2 + 2a \cdot \kappa_1^{\text{EMCCD}}. \tag{2.80}$$

Rearranging this for the cumulants of the noise-less signal yields

$$\kappa_1 = \frac{1}{a}\kappa_1^{\text{EMCCD}} \tag{2.81}$$

$$\kappa_2 = \frac{1}{a^2}\left(\kappa_2^{\text{EMCCD}} - 2a \cdot \kappa_1^{\text{EMCCD}}\right). \tag{2.82}$$

Again, using the moments given above we computed corrections for the first four cumulants, which are given in Appendix E.

ANALYSIS AND EXAMPLES

To better understand how the noise influences the final result, we first take second order SOFI of movies subject to Poisson noise as an example. Simulations as illustrated in Figure 2.24 are used to visualize the discussed effects and demonstrate how well the correction performs. The noise-contaminated SOFI signal is given by the true signal plus the average intensity.

$$\text{SOFI}_2^* = \text{SOFI}_2 + \text{Avg} \tag{2.83}$$

If the fluorophores follow a two-state model (see Appendix C), we know the average intensity

$$\text{Avg} = N_{\text{em}}\varepsilon\frac{k_+}{k_+ + k_-} + b \tag{2.84}$$
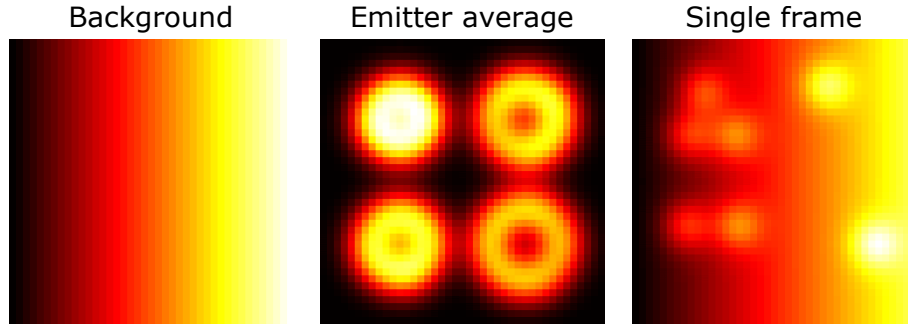
Figure 2.24: Simulation used for the demonstration of the noise correction. 10 emitters are placed in rings of different sizes on top of a slope-like background. Images are shown here without noise (infinite photon count) for clarity. The background count rises from b at the leftmost position to $4b$ at the rightmost position.

and the desired SOFI signal of the emitters to be

$$\text{SOFI}_2 = N_\text{em}\varepsilon^2 \frac{k_+ k_-}{(k_+ + k_-)^2}. \tag{2.85}$$

Here k_+ and k_- are the transition rates from the off- to the on-state and vice versa, N_em is the number of emitters, ε their molecular brightness and b the average number of background photons. The ratio of the average to the signal of the emitters is

$$\frac{\text{Avg}}{\text{SOFI}_2} = \frac{k_+ + k_-}{\varepsilon k_-} + \frac{b}{N_\text{em}\varepsilon^2} \frac{(k_+ + k_-)^2}{k_+ k_-} \tag{2.86}$$

$$= \frac{1}{\varepsilon} \cdot \left(\underbrace{(1 + r)}_{\text{emitters}} + \underbrace{\frac{b}{N_\text{em}\varepsilon} \left(\frac{1}{r} + r + 2 \right)}_{\text{background}} \right). \tag{2.87}$$

For large values of this ratio, the influence of noise in the uncorrected SOFI image is stronger. To simplify the analysis, we introduce the ratio of state lifetimes / transition rates $r = k_+/k_- = \tau_\text{on}/\tau_\text{off}$. As expected, the influence of the noise drops with increasing emitter brightness ε. The first term indicates how strong the high resolution image of each emitter is disturbed by the low resolution average image and gets large for high r ($\tau_\text{on} \gg \tau_\text{off}$). This is easy to explain: the second order signal is strongest if the lifetimes are equal ($\tau_\text{on} = \tau_\text{off}$) and is reduced in asymmetric situations. However, the average grows if the emitters stay longer in the on-state ($\tau_\text{on} > \tau_\text{off}$) and more photons are collected. For the same reason, the background related term grows for asymmetric $\tau_\text{on}, \tau_\text{off}$. Otherwise, it depends mainly on the ratio between the number of background photons to the number of photons collected from the emitters $N_\text{em}\varepsilon$.

Simulations were performed using 2000 frames with the width of the PSF set to $\sigma_\text{PSF} = 3$ px, transition rates $k_+ = 0.4, k_- = 0.6$ and varying brightness of the emitters as well as the background. Figure 2.25 shows how the resolution of the SOFI image is lowered at very low photon count rates and clearly
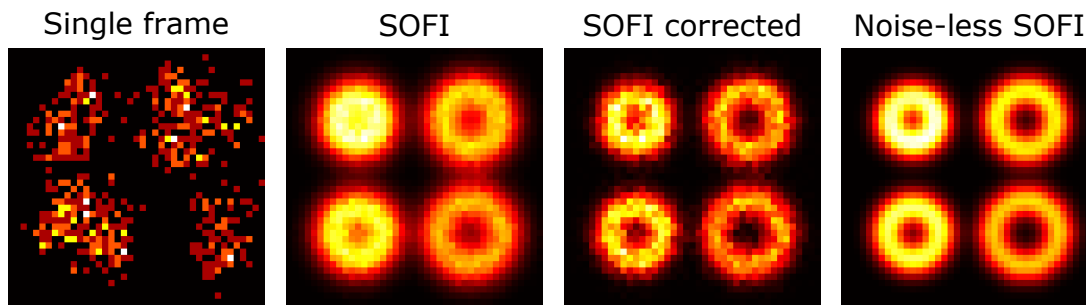
Figure 2.25: Simulation without background and low emitter brightness $\varepsilon \approx 40$ photons (amplitude: 1 photon). At low photon count rates, the resolution of the SOFI image is noticeably lowered by the overlay with the average image and clearly improved by the correction.
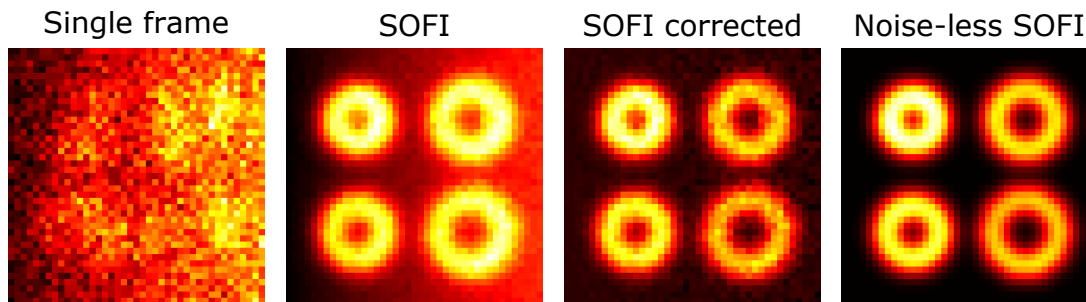
Figure 2.26: Simulation with (leftmost) background $b = 20$ photons and emitter brightness $\varepsilon \approx 400$ photons (amplitude: 10 photons). The background is clearly visible in the uncorrected image, lowering the image contrast, but is removed by the correction.
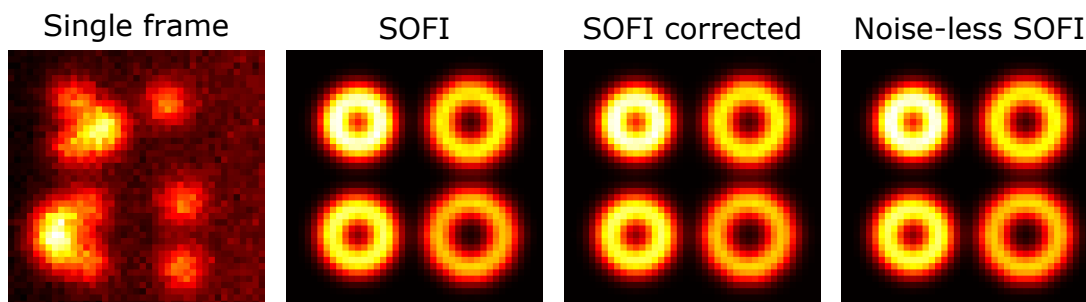
Figure 2.27: Simulation with (leftmost) background $b = 20$ photons and very high emitter brightness $\varepsilon \approx 4000$ photons (amplitude: 100 photons). At higher photon count rates, the impact of noise is negligible, making the noise correction unnecessary.
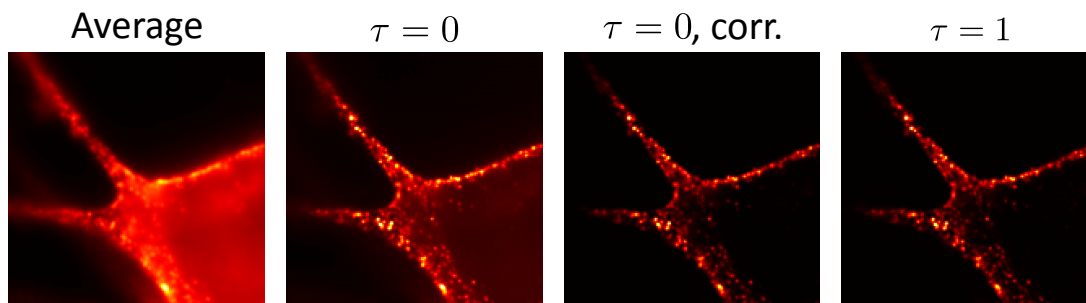
Figure 2.28: Noise correction of experimental data of a mouse hippocampal neuron stained with quantum dots (QD_{605}) recorded by *Anja Huss*. In the zero time-lag image (second order) the desired SOFI signal is overlayed with the average image. This contribution is eliminated by the noise correction, giving much better contrast and improved resolution. Due to the very long blinking time-scale of the quantum dots, the corrected image can be compared with a SOFI image calculated with a time-lag of one frame. As expected both images are virtually identical, proving the noise correction algorithm.

improves by applying the noise correction. If noticeable background is present in the recording, it will also be visible in the uncorrected SOFI images, but is easily removed (2.26). As Figure 2.27 demonstrates, for very bright samples the noise correction becomes unnecessary. In general, the influence of noise does not influence the SOFI result too much even for moderate emitter intensities because of its quadratic dependence on the brightness. However, in situations with strong background the correction is very useful. One experimental example of this situation is shown in Figure 2.28, where it is clear that the noise correction improves the result. Furthermore, the long time-scale of blinking of the utilized quantum dots allows the computation of a SOFI image with nonzero time-lag. The result is virtually identical to the corrected zero time-lag image, proving the effectiveness of the correction.

If the noise correction is used with fSOFI, the upsampled images should be multiplied with the square of the upsampling factor $u = N_{\text{new}}/N_{\text{old}}$ where $N \times N$ is the image size. Because zero padding in the Fourier domain is energy conserving and the number of pixels in real space is increased, their values drop. This change in number changes the statistics, and a conversion factor (similar to a for the EMCCD above) becomes necessary. Multiplication with u^2 ensures that pixels falling on the positions of the original sampling points retain the original values and the derived formulas apply.

The noise correction can be useful in many cases, but while it successfully removes the bias, the noise still increases the variance of the cumulants [43], resulting in a lower signal-to-noise ratio. Methods which inherently remove shot noise contributions by exploiting physical correlations (i.e. using time-lag or spatial cross-correlation) perform better at suppressing the noise. Thus, they deliver a better SNR of the final image as long as the blinking time-scale is not too fast or the PSF is not too small when using time-lag or spatial cross-correlation respectively. A proper comparison of the SNR is needed to estimate the usefulness of the correction.

To sum up this section about corrections for the effects of noise in zero-lag-time SOFI images,

- Noise causes a linear combination of lower cumulant orders to be added to a SOFI image calculated with zero time-lag (ordinary cumulants).

- If the distribution of the noise is known, formulas can be derived to correct for these contributions. For Poisson as well as EMCCD camera noise, formulas are derived here and given in Appendix E up to the 4th SOFI order.

- For moderate emitter intensities and not too high background, the influence of the noise is low and can be neglected.

- Methods exploiting the physical correlations (time-lag / spatial cross-correlation) potentially yield a higher SNR (if the time/length scale of the correlation relied upon is large enough).

2.6 Estimation of the OTF and PSF

The Point Spread Function (PSF) is one of most important characteristics of a microscope, as its shape encodes resolution and optical aberrations. Knowing the PSF as precisely as possible is also one of the prerequisites for successfully applying most schemes for "deconvolution", which try to increase the image resolution and remove the effect of aberrations in post-processing . Unfortunately, the PSF can vary for different measurements, for example due to different refractive indices of the embedding medium or different focal depth, but is hard to determine from the recorded image of the structure of interest where no isolated emitters are present. For this reason, it is usually determined using bright fluorescent beads under conditions as similar as possible to the measurement. However, this process is tedious and time consuming and the conditions are hard to match in some cases.

The stochastic fluctuations emitters exhibit in SOFI recordings can be used to directly compute a magnitude-based estimate of the OTF, even if no isolated single emitters are visible in the data. To do this, we essentially need to compute cumulants in a similar way to a normal SOFI computation, but apply this in Fourier space instead of the original data from the camera. After the OTF is known, the PSF can be computed using the Fourier transform.

2.6.1 OTF estimation for 3D-imaging

Let us assume at first that we recorded a 3D-movie similar to the work shown in [28]. The Fluorescence signal F and its Fourier transform $\tilde{F}$ with PSF $U(\mathbf{r})$ and OTF $\overline{U(\mathbf{k})}$ are given by:

$$F(\mathbf{r}, t) = \sum_j U(\mathbf{r} - \mathbf{r}_j) \cdot \varepsilon_j \cdot s_j(t) \text{ and} \tag{2.88}$$

$$\tilde{F}(\mathbf{k}, t) = \sum_j \tilde{U}(\mathbf{k}) \cdot e^{-i\mathbf{k}\mathbf{r}_j} \cdot \varepsilon_j \cdot s_j(t). \tag{2.89}$$

Here $\mathbf{r}_j$ is the three-dimensional position of the j-th emitter, ε_j its brightness and $s_j(t)$ the associated time dependent fluctuation function.

Computing the 2nd order joint cumulant of the Fourier transform with its complex conjugate yields:

$$\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)} \rangle =$$
$$= \tilde{U}(\mathbf{k}) \cdot \overline{\tilde{U}(\mathbf{k})} \cdot \sum_j \sum_l \varepsilon_j \varepsilon_l \cdot e^{-i\mathbf{k}\mathbf{r}_j} e^{i\mathbf{k}\mathbf{r}_l} \cdot \langle \delta s_j(t)\delta s_l(t + \tau) \rangle \tag{2.90}$$

$$= |\tilde{U}(\mathbf{k})|^2 \cdot \sum_j \varepsilon_j^2 \langle \delta s_j(t)\delta s_j(t + \tau) \rangle \tag{2.91}$$

$$= |\tilde{U}(\mathbf{k})|^2 \cdot \text{const} \tag{2.92}$$

With $\delta\tilde{F} = F - \langle F \rangle$ and $\delta s = s - \langle s \rangle$. Note that after the cross-terms drop out due to the cumulant computation, the terms depending on the emitter positions cancel each other because the complex conjugate is used here. Looking at this, we can get a magnitude-based estimate of the PSF simply by taking

the square root of the real part (the phase information is lost):

$$\tilde{\tilde{U}}(\mathbf{k}) \approx \sqrt{\mathrm{Re}\left\{\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)}\rangle\right\}} \tag{2.93}$$

$$= |\tilde{U}(\mathbf{k})| \cdot \mathrm{const} \tag{2.94}$$

If the PSF is well approximated by an arbitrary rotated asymmetric Gaussian, its Fourier transform should be purely real it can be perfectly reconstructed. However, due to the loss of phase information, higher order aberrations are not captured by the reconstruction (a way to recover the phase information is sketched below in 2.6.4). Back-transforming into real space gives an approximate image of the PSF.

2.6.2   OTF ESTIMATION FOR 2D-IMAGING

When imaging a 3D-structure with conventional 2D-imaging, we do not observe the three dimensional light distribution of the emitters but rather a cut through their 3D-PSF. In this case the 2D fluorescence signal and its 2D (lateral) Fourier transform can be described by:

$$F(x, y, t) = \sum_j U(x - x_j, y - y_j, z_j) \cdot \varepsilon_j \cdot s_j(t) \tag{2.95}$$

$$\tilde{F}(k_x, k_y, t) = \sum_j \int_{-\infty}^{\infty} \mathrm{d}k_z\, \tilde{U}(k_x, k_y, k_z) \cdot \mathrm{e}^{ik_z z_j} \cdot \mathrm{e}^{-i\mathbf{k}\mathbf{r}_j} \cdot \varepsilon_j \cdot s_j(t). \tag{2.96}$$

Again we compute the 2nd order cumulant of the Fourier transform with its complex conjugate:

$$\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)}\rangle =$$

$$\sum_j \int \int_{-\infty}^{\infty} \mathrm{d}k_z \mathrm{d}k_z'\, \tilde{U}(k_x, k_y, k_z) \cdot \overline{\tilde{U}(k_x, k_y, k_z')} \cdot \mathrm{e}^{i(k_z - k_z')z_j} \cdot \varepsilon_j^2 \langle \delta s_j(t)\delta s_j(t + \tau)\rangle, \tag{2.97}$$

where we again used the fact that the cross-terms between different emitters drop out, which leads to cancellation of the lateral phase factors. Assuming a large number of blinking emitters with similar intensity and blinking behavior is observed, we can think of the sum over all emitters as an integral over the emitter density function $\sum_j \varepsilon_j \mathrm{e}^{i(k_z - k_z')z_j} \rightarrow \int_{-\infty}^{\infty} \mathrm{d}z\, \rho(z) \cdot \varepsilon \mathrm{e}^{i(k_z - k_z')z}$. If the emitters are distributed uniformly along the axial direction, this becomes a δ-function, $\delta(k_z - k_z')$, times some constant pre-factor. In this case, above expression simplifies to

$$\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)}\rangle \propto \langle \delta s(t)\delta s(t + \tau)\rangle \int_{-\infty}^{\infty} \mathrm{d}k_z\, |\tilde{U}(k_x, k_y, k_z)|^2, \tag{2.98}$$

which is the axial projection of the absolute square of the OTF. As the axial distribution of emitters is not necessarily uniform (neither is their blinking behavior), this estimate might be bad in practice. However, if the sample is very flat or TIRF microscopy is used, we can assume that all emitters are very

close to the focal plane. Thus we set $z_j = 0 \; \forall j$ in Eq. (2.97) to get:

$$\left(\int\limits_{-\infty}^{\infty} \mathrm{d}k_z \, \tilde{U}(k_x, k_y, k_z) \right) \cdot \left(\int\limits_{-\infty}^{\infty} \mathrm{d}k'_z \, \overline{\tilde{U}(k_x, k_y, k'_z)} \right) \cdot \left(\sum_j \varepsilon_j^2 \langle \delta s_j(t) \delta s_j(t + \tau) \rangle \right) \qquad (2.99)$$

$$= \tilde{U}_{2D}(k_x, k_y) \cdot \overline{\tilde{U}_{2D}(k_x, k_y)} \cdot \left(\sum_j \varepsilon_j^2 \langle \delta s_j(t) \delta s_j(t + \tau) \rangle \right) \qquad (2.100)$$

$$= |\tilde{U}_{2D}(k_x, k_y)|^2 \cdot \sum_j \varepsilon_j^2 \langle \delta s_j(t) \delta s_j(t + \tau) \rangle. \qquad (2.101)$$

The axial projection of the OTF U_{2D} is the Fourier transform of the in-focus slice through the 3D-PSF according to the Fourier slice theorem. From this, the 2D-PSF can be reconstructed similar to the 3D case in Eq. (2.94).

2.6.3 OTF EXTRACTION IN PRACTICE
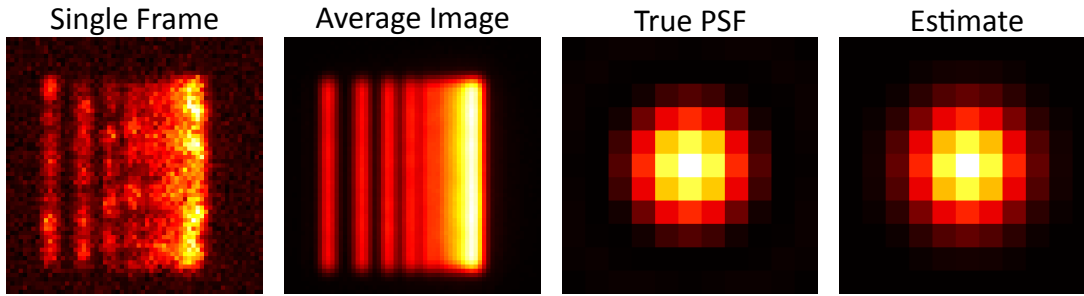
Figure 2.29: Extraction of the PSF from a simulated movie with 2000 frames, where all emitters are in the same focal position. The blinking rates are $k_+ = 0.1, k_- = 0.2$. A time-lag of $\tau = 1$ frame was used for extraction of the OTF magnitude and its shape extracted by fitting a Gaussian model to the data. The Fourier transform yields a Gaussian PSF, which is very close to the true Airy-disc shape, but is 5.7% wider. The difference vanishes with increasing signal-to-noise ratio.

For OTF-estimation the movie must be Fourier-transformed first and then the covariance between the Fourier transform at time t and its complex conjugate at frame $t + \tau$ has to be computed. Similar to second order SOFI this processing should be done for individual intervals of N frames to counter the effects of bleaching (see Sec. 2.2). Due to shot noise in the original signal, the resulting image should be fitted with a suitable model to extract the OTF even for nonzero time-lag. For TIRF / thin-sample movies this can be done simply using a 2D Gaussian with a background offset. Although the OTFs magnitude is expected to fall to zero at high frequencies, this offset arises when applying the Fourier transform to a signal subject to Poisson noise [49].

It was observed using simulations that the estimated OTF magnitude is narrower than the true OTF when the signal-to-noise ratio is low due to either strong background or very dim emitters. Figure 2.29 shows an example for the recovery of the PSF from the estimated OTF, where its estimated shape is 5.7% wider then the best fitting Gaussian. This difference gets smaller as the signal to noise ratio increases. So far there is no explanation for this bias, a better understanding of the effects of noise in real

space on the calculation in Fourier space is desirable to improve the estimation.

2.6.4 REMARKS ON RECOVERY OF THE PHASE INFORMATION

The 2nd order cumulant only yields magnitude information, but if the third order cumulant can be reliably computed for a movie, the phase of the OTF can be extracted as well. This allows a complete reconstruction of the microscope's PSF including aberrations. One way for performing this reconstruction is very similar to the "bispectrum" method in astronomy. There, the bispectrum method is commonly employed to reconstruct a high-resolution image from a movie of short-exposure images acquired by a ground telescope, which are blurred by atmospheric disturbances. If the exposure time is so short that the atmosphere is essentially "frozen", the recorded images still carry the high frequency information which is lost in long exposure images. In this scenario the OTF is fluctuating randomly while the observed object (i.e. the stars) is assumed to be constant. The method is well described in [50].

In our case, the OTF is assumed to be constant, while the object fluctuates randomly. Here we briefly sketch how the bispectrum method works following the description in [50] but adapted to our specific scenario. The bispectrum of our SOFI movie is given by

$$
\begin{aligned}
B(\mathbf{u}, \mathbf{v}) &= \langle \delta\tilde{F}(\mathbf{u}, t) \cdot \delta\tilde{F}(\mathbf{v}, t + \tau_1) \cdot \overline{\delta\tilde{F}(\mathbf{v} + \mathbf{u}, t + \tau_2)} \rangle \\
&= \sum_j \tilde{U}(\mathbf{u})\tilde{U}(\mathbf{v})\overline{\tilde{U}(\mathbf{u} + \mathbf{v})} \cdot e^{i\mathbf{v}\mathbf{r}_j}e^{i\mathbf{u}\mathbf{r}_j}e^{-i(\mathbf{u}+\mathbf{v})\mathbf{r}_j} \cdot \langle \delta s_j(t)\delta s_j(t + \tau_1)\delta s_j(t + \tau_2) \rangle \quad (2.102)
\end{aligned}
$$

$$
= \tilde{U}(\mathbf{u})\tilde{U}(\mathbf{v})\overline{\tilde{U}(\mathbf{u} + \mathbf{v})} \cdot \sum_j \langle \delta s_j(t)\delta s_j(t + \tau_1)\delta s_j(t + \tau_2) \rangle \quad (2.103)
$$

$$
= \tilde{U}(\mathbf{u})\tilde{U}(\mathbf{v})\overline{\tilde{U}(\mathbf{u} + \mathbf{v})} \cdot \text{const.} \quad (2.104)
$$

The vectors $\mathbf{u}$, $\mathbf{v}$ are two- or three-dimensional Fourier space coordinates, depending on the dimensionality of the recorded movie. Here we again exploited the fact that cross-terms cancel due to the cumulant calculation. The constant factor is not relevant to the problem and will be omitted for clarity of notation. Rewriting the equation above in the magnitude-phase representation of complex numbers, we see a connection between the bispectrum phase β and the phase of the OTF φ:

$$
|B(\mathbf{u}, \mathbf{v})|e^{i\beta(\mathbf{u},\mathbf{v})} = |\tilde{U}(\mathbf{u})||\tilde{U}(\mathbf{v})||\tilde{U}(\mathbf{u} + \mathbf{v})| \cdot e^{i\varphi(\mathbf{u})}e^{i\varphi(\mathbf{v})}e^{-i\varphi(\mathbf{u}+\mathbf{v})} \quad (2.105)
$$

$$
\Rightarrow \beta(\mathbf{u}, \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v}) - \varphi(\mathbf{u} + \mathbf{v}) \quad (2.106)
$$

$$
\Leftrightarrow \varphi(\mathbf{u} + \mathbf{v}) = \varphi(\mathbf{u}) + \varphi(\mathbf{v}) - \beta(\mathbf{u}, \mathbf{v}) \quad (2.107)
$$

This relationship can be used for a recursive reconstruction of the PSF phase. For simplicity of notation we assume the problem is two-dimensional here. Because the PSF $U(\mathbf{r})$ is a real function, its Fourier transform is Hermitian $\tilde{U}(\mathbf{k}) = \overline{\tilde{U}(-\mathbf{k})}$ and thus the phase at the origin must be zero $\varphi(\{0, 0\}) = 0$. Multiplication of $\tilde{U}(\mathbf{k})$ with an arbitrary global phase factor $\exp(ik_x x_0) \cdot \exp(ik_y y_0)$ only changes the position of the reconstructed PSF. As we are not interested in the absolute position, we can choose a phase factor that sets $\varphi(\{1, 0\}) = 0, \varphi(\{0, 1\}) = 0$. Knowing these three starting values we can

reconstruct all values using a recursion formula like:

$$\varphi(\{1,1\}) = \varphi(\{1,0\}) + \varphi(\{0,1\}) - \beta(\{1,0\}, \{0,1\})$$
$$\varphi(\{0,2\}) = \varphi(\{0,1\}) + \varphi(\{0,1\}) - \beta(\{0,1\}, \{0,1\})$$
$$\vdots$$

Note that the bispectrum has twice as many dimensions as the original data, giving a 6D-bispectrum for 3D-recordings. A 3D movie with 512x512x8 pixel frames has roughly 2 million pixels per frame (approx. 16 Megabyte using 64 bit real numbers), leading to a bispectrum with over 4 trillion pixels (approx. 70 terabyte using 64+64 bit complex numbers). For a two-dimensional 512x512 image the complete bispectrum would still take about 512 Gigabytes of memory, which is much more than common computers provide today, making the computation of the full bispectrum impractical. The bispectrum can be compressed without loss by a factor of 12 using symmetry relations [51], but it is still too large for most applications. For this reason, calculation of bispectrum and phase values might be done using only a restricted subset of **u**, **v** vectors [51] and there has been some work on further optimizing and parallelizing the computation [52]. Completely different (non-recursive) algorithms using the bispectrum are also used, for example based on least-squares optimization [53].

As phase recovery is outside the focus of this work, the involved algorithms are complex and take a long time to implement correctly, an implementation of this method is not presented here. However, using these ideas phase recovery is expected to work well given that the quality of the recorded data allows a reliable computation of the third order cumulant.

To sum up this section about OTF-estimation,

- The 2nd order joint cumulant of the Fourier transform of one movie frame with the complex conjugate of another (or the same) can be used to extract a magnitude-based estimate of the OTF of the microscope and thus the PSF.

- The 3rd order joint cumulant can be used to also extract the phase, but requires more sophisticated algorithms.

- For usual 2D recordings, the reconstruction of the PSF will only yield meaningful results if a TIRF configuration is used and all emitters are approximately in the same plane.

- For 3D-recordings, the complete 3D-PSF can be reconstructed.

2.7 ADAPTIVE RESOLUTION IMPROVEMENT OF SOFI IMAGES

As described in Section 2.1, SOFI works on movies of blinking fluorescent emitters

$$F(\mathbf{r}, t) = \sum_j U(\mathbf{r} - \mathbf{r}_j) \cdot \varepsilon_j \cdot s_j(t) \tag{2.108}$$

as input and uses the n-th order cumulant to compute the high resolution image

$$S_n(\mathbf{r}) = \sum_j U^n(\mathbf{r} - \mathbf{r}_j) \cdot \varepsilon_j^n \cdot \kappa_n\left(s_j(t)\right) \tag{2.109}$$

where $U(\mathbf{r})$ is the point spread function, ε_j the brightness of the j-th emitter and $s_j(t)$ its time dependent fluctuation. While the exponentiated brightness skews the local image intensity (see Fig. 2.5), the power of the point spread function yields a higher spatial resolution. If the in-focus PSF is described by a Gaussian, the resolution is increased by factor $\sqrt{n}$ (Eq. (2.4)). However, the theoretically possible resolution increase is linear in n, as the Fourier transform of $U^n(\mathbf{r})$

$$\mathcal{F}\left(U(\mathbf{r})^n\right) = \underbrace{\tilde{U}(\mathbf{k}) \circledast \cdots \circledast \tilde{U}(\mathbf{k})}_{(n-1) \text{ convolutions}} \tag{2.110}$$

has an n times broader support than the original function. The lower-than-expected resolution stems from the (relative) attenuation of higher frequencies. This is clearly visible in Figure 2.30, which compares the shape of the OTF for different orders. To ease the comparison, all OTFs are individually compressed to the same range on the frequency axis, making their maximum frequencies equal to one.
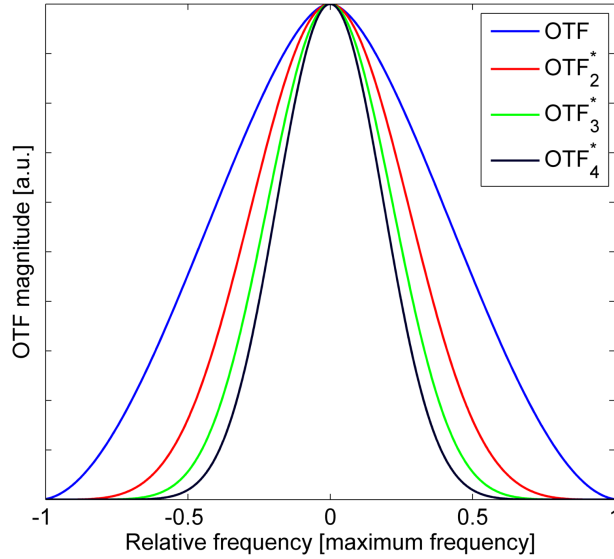
Figure 2.30: Comparison of the original OTF shape to different SOFI orders. To ease the comparison the OTFs are individually compressed to map their support to the range $[-1, 1]$. Note that this compresses higher order more, as the support increases linearly with the order. Optimally, all OTFs should be identical in this representation. However, it is clear that higher frequencies are strongly attenuated for higher SOFI orders. For a comparison of the uncompressed OTFs see Figure 2.2.

Correcting for inhomogeneous brightness is a demanding task. From Figure 2.5 it is immediately clear that multiple cumulant orders are necessary to linearize the brightness, as brightness changes in SOFI images can be caused by different density, emitter brightness, or blinking behavior. Thus, at least three different orders are necessary for the correction (and only if the photoblinking is simple enough to be modeled by one parameter). The time-averaged image (1st order cumulant) also carries background components which increases the number of unknowns further. Taking the lowest possible orders, the 2nd, 3rd and 4th order cumulant are necessary for the reconstruction. We know from Section 2.3 that reliable estimates for the higher order cumulants require too many frames to be used in many practical applications, making linearization of brightness difficult. Even if they are available, the situation is further complicated by the different optical resolution of the images. Geissbühler *et al.* presented a calculation of the parameters assuming a spatially uniform distribution of the fluorescent molecules [39]. However, they do not use the calculated brightness for linearization, but rely on a deconvolution scheme. Assuming single emitters are perfectly separated by the deconvolution, brightness is linearized using the *n*-th root. The author state this is done in order to not compromise the resolution, as the calculated parameter maps only carry information on the scale of the lowest cumulant (the 2nd order). The approximations taken here are more or less valid depending on the sample that is imaged. In practice, no reliable scheme for intensity linearization has been established so far.

It is desirable to extract the maximum possible resolution from each SOFI order. As already mentioned, the work of Geissbühler *et al.* proposes to use deconvolution. They rely on Richardson-Lucy deconvolution [54, 55], a well known optimization algorithm to estimate the underlying sample distribution for a known PSF and Poisson statistics. Although the noise statistics in SOFI images is not Poissonian, the authors report this still works well. Our own results, presented in Section 2.7.1, indicate that, while these deconvolution schemes can improve the visual quality, they do not perform well in improving the resolution of SOFI images. A different approach was suggested by Dertinger *et al* [40]. They proposed to perform a reweighting of the available spatial frequencies of the final image in Fourier space to achieve the desired PSF in real space. The approach works well for high SNR, but is liable to producing artifacts. A user definable parameter selects a trade-off between noise amplification and resolution and must be determined by trial and error.

This section deals with enhancing the resolution of SOFI images. First, we review conventional deconvolution schemes and point out some associated problems. This is important, as we feel that optimization based deconvolution is often not properly discussed. Then, we revisit the Fourier reweighting approach from Dertinger *et al* [40]. We take up the core concept of this method and present an improved algorithm to increase the resolution as well as the SNR of the final SOFI image. Furthermore we show one approach how to automatically adapt the method to the quality of the data, which improves the resolution only so far as to not produce artifacts.

2.7.1 CONVENTIONAL DECONVOLUTION

A lot of deconvolution/deblurring algorithms exist in the field of image processing. However, when processing data from scientific experiments the goal is not to simply achieve a crisper, sharper looking image, but to extract meaningful information. Thus, these image processing algorithms are best derived from physical models of the underlying process, and most are constructed using a similar scheme. To better understand how this works, we briefly derive the Richardson-Lucy algorithm as used by Geiss-bühler *et al.*[39]. Assuming the measured data is subject to Poisson noise. The chance to measure k photons in a pixel with average photon flux λ is

$$\text{Poiss}_\lambda = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \tag{2.111}$$

In a measured image, the average photon flux varies spatially $\lambda(\mathbf{r})$. Because all pixels are statistically independent, the chance to measure a specific image $i(\mathbf{r})$ is given by their product. This yields the *likelihood functional*

$$L\{\lambda(\mathbf{r})\} = \prod_r \frac{\lambda(\mathbf{r})^{i(\mathbf{r})}}{i(\mathbf{r})!} \cdot e^{-\lambda(\mathbf{r})} \tag{2.112}$$

In practice the measured image is known and the goal is to find the most likely distribution for $\lambda(\mathbf{r})$. This maximization of a physically sound probability is often called *Maximum Likelihood Estimation*. Instead of maximizing $L\{i(\mathbf{r})\}$, it is often easier to minimize the *negative log-likelihood functional*

$$l\{\lambda(\mathbf{r})\} = -\log(L\{\lambda(\mathbf{r})\}) = -\sum_r i(\mathbf{r}) \log(\lambda(\mathbf{r})) - \lambda(\mathbf{r}) - \log(i(\mathbf{r})!) \tag{2.113}$$

The last term is constant with respect to $\lambda(\mathbf{r})$ and can be neglected for the optimization. Now further information about the imaging process that produces $\lambda(\mathbf{r})$ can be integrated. For example, with the spatially homogeneous PSF $U(\mathbf{r})$, the sample density $\rho(\mathbf{r})$ and a uniform background photon flux b, we get

$$\lambda(\mathbf{r}) = U(\mathbf{r}) \circledast \rho(\mathbf{r}) + b \tag{2.114}$$

$$\rightarrow l\{\lambda(\mathbf{r})\} = -\sum_r i(\mathbf{r}) \log(U(\mathbf{r}) \circledast \rho(\mathbf{r}) + b) - U(\mathbf{r}) \circledast \rho(\mathbf{r}) - b, \tag{2.115}$$

where we omitted the previously mentioned constant term. The function of interest is the sample density $\rho(\mathbf{r})$. For known PSF and background, we can find its optimal value by setting the functional derivative to zero:

$$\frac{\delta l}{\delta \rho}(\mathbf{r}) = \left(\frac{-i(\mathbf{r})}{U(\mathbf{r}) \circledast \rho(\mathbf{r}) + b}\right) \circledast U(-\mathbf{r}) - 1 \stackrel{!}{=} 0 \tag{2.116}$$

$$\Rightarrow \underbrace{\left(\frac{-i(\mathbf{r})}{U(\mathbf{r}) \circledast \rho(\mathbf{r}) + b}\right) \circledast U(-\mathbf{r})}_{c} = 1 \tag{2.117}$$

Classical Richardson-Lucy uses the left side of this equation as a multiplicative factor to refine the current estimate of the sample density

$$\rho_{n+1} = c_n \cdot \rho_n \qquad (2.118)$$

If ρ is too big, c will be lower than 1 (as the denominator of the fraction is too big) and its value drops in the next iteration. If ρ is too small, it will increase in the same fashion[3].

For carrying out the minimization, a zoo of optimization algorithms exist, all with their strength and weaknesses. Here, we use a self-written implementation of the *nonlinear conjugate gradient* algorithm, a well established method for convex optimization problems [56]. The function to optimize (in our case the negative log likelihood function) is often also called the *cost function*. Many inverse problems are not well posed, which causes instabilities in the deconvolution process and is also the case for images from fluorescence microscopy. For example, because the OTF has a finite support, higher frequencies of the sample density are completely unconstrained and infinitely many solutions with different spectra exist. It is common to add terms to the physically derived cost that are supposed to select certain solutions and stabilize the deconvolution; a process usually called *regularization*.

One well-known method using this is *Least-Squares Deconvolution with Total Variance Regularization* with the cost function

$$C_{\mathrm{LSQ}}\{\rho(\mathbf{r})\} = \sum_r \left[(U(\mathbf{r}) \circledast \rho(\mathbf{r}) - i(\mathbf{r}))^2 + a_{\mathrm{TV}} \cdot |\nabla\rho(\mathbf{r})| \right] \qquad (2.119)$$

and the functional derivative

$$\frac{\delta C}{\delta \rho}(\mathbf{r}) = 2\left(U(\mathbf{r}) \circledast \rho(\mathbf{r}) - i(\mathbf{r})\right) \circledast U(-\mathbf{r}) - a_{\mathrm{TV}} \cdot \mathrm{div}\left(\frac{\nabla\rho(\mathbf{r})}{|\nabla\rho(\mathbf{r})|}\right) \qquad (2.120)$$

The first term of the cost function is the log-likelihood function for a signal with Gaussian-distributed noise and uniform standard error across the entire image. The second term punishes high gradients and thus promotes solutions which are "smoother". The degree of smoothness is varied by the user definable parameter a_{TV}, which is often determined by trial and error on the data of interest.

Optimization based deconvolution is a complex topic and has some fundamental problems when applied to problems in physics, which unfortunately seem to be not properly discussed in many publications. In many cases, this is connected to missing knowledge or misunderstanding of the limitations of the applied method; a topic that is often very hazy or simply unknown. For example, in the best case maximum likelihood estimation yields the "most likely" sample density distribution, but what does this actually mean? Although a maximum might exist, there might be arbitrarily many distributions which are almost as likely as the result. Since the measurement is a probabilistic process, only probabilistic statements on possible sample distributions are possible, which are encoded in the full likelihood func-

---

[3]The formal derivation of the algorithm uses a framework of mathematical statistics called *Expectation-Maximization*, but this is out of scope of this treatment.

tion. Thus, each proposed sample distribution has a certain probability to be true, but it is crucial to not mistake the result of some optimization algorithm for the true distribution. Unfortunately, the error of the estimation, which is approximated by the width of the likelihood function around the proposed value, is usually not given or it is even impossible to calculate. Many publications simply show the final image, pointing out that it "looks sharper" and show that some structure decreases in width by plotting a cross-section through it. This is often not even performed on simulated data with ground truth, but only using experimental data with unknown ground truth. It is important to point out, that this does not prove truly improved optical resolution. For instance, if fine structures are present in fluorescent images, they will also appear sharper by taking the pixel-wise power of the image pixels, as structures are always more intense in their center compared to the border, although the optical resolution of course remains unchanged.

As most deconvolution algorithms suppress noise and sharpen the image, the result is often more visually appealing, which makes it tempting to accept this as an improvement. But without quantifying the likeliness of the result and its error, it should always be viewed with appropriate skepticism. While giving error bars is common practice almost everywhere in physics, this is interestingly often neglected in scientific image processing. In some cases bootstrapping[4] may be applied to estimate the error if the imaging model fits very well to the data.

To further complicate the issue, minimization algorithm are only guaranteed to find the global minimum if the cost function is convex, i.e. has only one minimum. Most cost functions in practice form a complex landscape, causing the optimization to get stuck in local minima. Finding the global minimum can only be achieved by sampling the likelihood function for all reasonable values, which is computationally prohibitive. Thus, most results of iterative optimization will not be optimal. The shape of this landscape is often also too complicated to describe the error of some estimate simply by the standard deviation (i.e. a Gaussian approximation of the likelihood function around the value). This makes it even harder to visualize, describe, and discuss the errors of the estimation.

One more problem is the application of the regularization terms. Often, these are designed simply by intuition and knowledge of the desired properties of the solution. Using this approach, the cost function cannot be interpreted in terms of a physically grounded likelihood anymore. The parameters that vary the strength of a regularization term can heavily influence the look of the final result and are usually chosen by the user who tries to achieve the "best looking" result. However, methodologies exist that derive regularization terms from physical assumptions. For example, prior information on the sample can be included utilizing a Bayesian framework, which strongly improves the reconstruction if such knowledge can be safely assumed. For example, if it is known that the sample consists of multiple overlapping but continuous line structures, this could be enforced in the reconstruction. In this case it is even possible to fit curves to the lines, which essentially yields almost unlimited spatial resolution, completely bypassing the classical resolution limit. Regularization must be used with utmost care, however, as the final result

---

[4]The imaging model is repeatedly applied to the estimated values of the sampling distribution, e.g. blurring with the PSF and adding noise. This generates new data which is deconvolved again to produce a set of results. This is supposed to produce a random sampling of the likelihood function around the proposed sample distribution. The standard deviation of this set gives an estimate of the error.
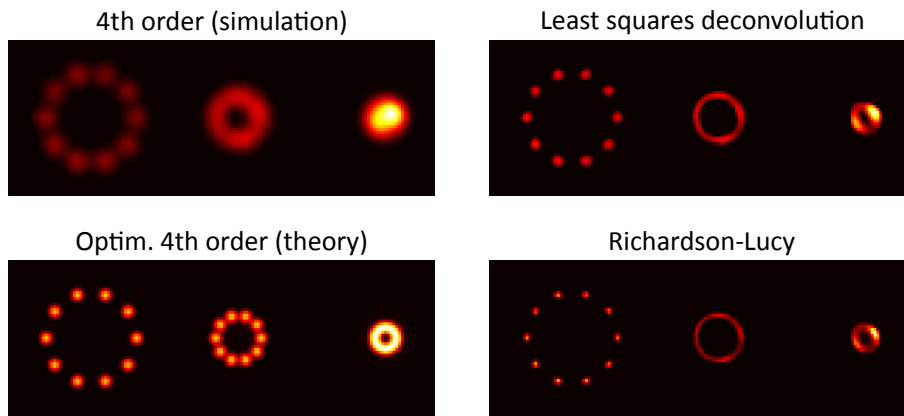
Figure 2.31: Examples for optimization-based deconvolution on 4th order SOFI images calculated from a noise-less simulation of 20000 frames compared to the theoretically optimal result. The original sample distribution is given by three rings of ten uniformly distributed emitters with decreasing radius. While the last structure can be clearly identified as ring-like after deconvolution, the central structure is not correctly recovered to show discrete emitters but appears as a very fine ring.

will always show the properties defined by the associated terms. So a reconstruction that expects lines will always tend to yield lines, even if the sample does not show this in reality.

In general it is very hard to prove what deconvolution can and cannot do and how well it performs. A complete mathematical description of the process and its errors is extremely difficult for the often nonlinear and ill-posed problems. Some publications studied the simplest possible configuration of two point sources in a defined distance [57–59]. In this case, the achievable resolving power is only determined by the SNR of the recorded image. The relationship between resolution and the measured number of photons can be quantified as well as the error of the estimation. However, even in this simple case, the analysis quickly gets complicated. For unknown sample structure it is generally hard to say what the achievable resolution is and if deconvolution can be trusted.

To summarize: Deconvolution algorithms should be used with extreme care and skepticism and their results should not be rated by simple visual examination. Furthermore, one should always keep the probabilistic nature of the result in mind. It is important to perform many simulations with known ground truth to prove the effectiveness of a certain approach for a specific imaging model. The application to experimental data should only be carried out after rigorous testing.

For SOFI images it is clear from the underlying theory that resolution improvement is possible based on the sub-optimal shape of the OTF (Figure 2.30). We used optimization based deconvolution on simulations to improve the resolution of the final image. A typical example of the results is presented in Figure 2.31, which shows the deconvolution of a 4th order SOFI image calculated from a noise-less simulation of 20000 frames compared to the theoretically possible resolution. While the last structure can be clearly identified as ringlike after deconvolution, the central structure is not correctly recovered to show discrete emitters but wrongly appears to be a very fine ring. This type of behavior was very common in our experiments using different data sets and cost functions. While the resolution is seemingly increased, this is often not really the case and the result reflects the true structure incorrectly. In fact, it is not even be possible to define a single resolution of the optimized image, at least not using the

classical definitions of optics. The improvement also varied for different structures within the same image. In practice it is uncertain whether the result really shows an improvement or not, which effectively prohibits the method from being useful. We also tried a state-of-the art cost function proposed for fluorescence microscopy by Arigovindan *et a.* [60]. Their regularization promotes high derivatives in the result only in regions of high intensity. Unfortunately structurally identical sample distributions that differ in brightness appeared with very different levels of sharpening after the deconvolution, although they do not differ in resolution according to the theory of SOFI.

We also tried to engineer various algorithms and cost functions of our own. Although the results in many cases appeared to "look better", we could not prove that any of the tested approaches reliably improves the resolution of SOFI images when comparing it to the ground truth of the simulation. Most algorithms and cost functions work with images in real space, where it is expected that the deconvolution also gets more difficult with higher SOFI orders. This is because the possible resolution improvement from $\sqrt{n}$ to n rises with the order (e.g. 2nd order $1.41 \rightarrow 2$, 4th order $2 \rightarrow 4$), while the initial resolution of the different orders is less and less distinct. Thus, the algorithm has to extract a large improvement based on minuscule differences in the intensity of the original image. So far it seems that iterative optimization-based deconvolution is not a good tool to extract the theoretically possible resolution from SOFI images.

2.7.2 Fourier Reweighting

Dertinger *et al* suggested a more direct approach for resolution improvement [40]. As explained in the beginning of this section, the support of the PSF of n-th order SOFI $U^n(\mathbf{r})$ in Fourier space increases linear with the order, but the resolution improves only by a factor of $\sqrt{n}$. This is reflected in a changed shape of the OTF where high frequencies experience increasing attenuation (Figure 2.30). Intuitively, the PSF with the desired resolution would be a linearly shrinked version of the one present in the raw data. Its Fourier transform

$$\mathcal{F}\left(U(n \cdot \mathbf{r})\right) = \tilde{U}\left(\frac{\mathbf{k}}{n}\right) \tag{2.121}$$

has the same support as the OTF of the SOFI image. Due to this, it is possible to increase the resolution of the SOFI image by manipulating the existing frequency information. Fourier reweighting is a one-step approach replacing the PSF of the SOFI image $S_n(\mathbf{r})$ in Fourier space:

$$
\begin{array}{l}
S_n(\mathbf{r}) \xrightarrow{\mathcal{F}} \tilde{S}_n(\mathbf{k}) \\[2ex]
\Big\downarrow \quad \cdot \dfrac{\tilde{U}\left(\frac{\mathbf{k}}{n}\right)}{\underbrace{\tilde{U}(\mathbf{k}) \circledast \tilde{U}(\mathbf{k}) \circledast \cdots \circledast \tilde{U}(\mathbf{k})}_{(n-1)\ \text{convolutions}} + a} \\[3ex]
S'_n(\mathbf{r}) \xleftarrow{\mathcal{F}^{-1}} \tilde{S}'_n(\mathbf{k})
\end{array}
\tag{2.122}
$$

Here $a > 0$, $a \in \mathbb{R}$ is a small positive constant which prevents division by zero and limits the maximally possible amplification. In practical applications the achievable resolution enhancement depends on the

signal-to-noise ratio of the SOFI image – too low values for a result in noise amplification and artifacts within the deconvolved image. The approach is only exact when three dimensional data is processed. For two dimensional movies $\tilde{U}(\mathbf{k})$ should be interpreted as the axial projection of the full OTF (see Section 2.6.2) and the procedure is strictly correct only for emitters within the focal plane.
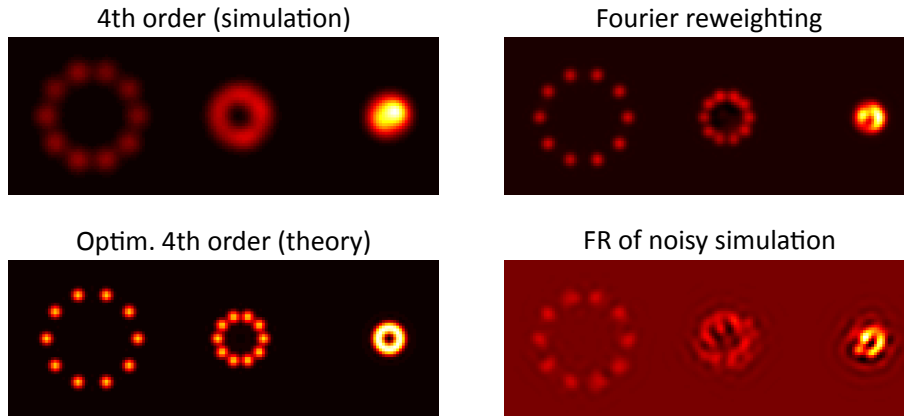
Figure 2.32: Example for Fourier reweighting on 4th order SOFI images calculated from a noise-less simulation of 20000 frames and the same data subject to Poisson noise. The original sample distribution is given by three rings with decreasing radius comprising ten uniformly distributed emitters. Fourier reweighting of the noise-less simulation correctly reproduces the sample structure of the second ring, in contrast to optimization-based deconvolution. Reweighting of the noisy data shows "ringing" artifacts which occur for too low values of a. Artifacts of this kind also appear if the estimated PSF is too different from the true one.

Figure 2.32 shows results and problems of Fourier Reweighting on a simulation similar to the one used in Figure 2.31 for optimization based deconvolution. In contrast to the optimization based approaches, Fourier reweighting correctly recovers the discrete emitters in the second ring. Because of the simple nature of the algorithm, it is guaranteed that it recovers the desired resolution in the limit of processing infinitely long image sequences. As the simulation shown here is finite, the third structure is not recovered to its ring-like form (compare to theoretical image in Fig. 2.31). The bottom image shows Fourier reweighting performed on the same raw data subject to Poisson noise, to highlight some typical problems of this approach. The recovered image shows strong artifacts, which are commonly described as "ringing" because of their wave-like appearance around the original intensity. The major cause for this are too low values of a in the presence of noise. Ringing also occurs if the estimated PSF is too different from the true one. It is important to note that artifacts are also visible for noise-less data if the processed movie sequence is too short, and the cumulant values are not sufficiently converged to their true values. The non-homogeneous convergence can cause non-PSF-like structures in the image and the imaging model used to derive Fourier reweighting does not apply anymore. This can cause artifacts that are often accompanied by unphysical changes in the sign of values within the image.

2.7.3   FOURIER PREWEIGHTING

The problems of Fourier reweighting can be reduced by a variation of the original concept. As was mentioned, changing the OTF in the final SOFI image causes artifacts in parts of the image where the cumulant values are not converged well enough. Instead of this post-processing, we can also manipulate

the raw data in a way that the cumulant calculation directly yields an image with the desired OTF. That is, for n-th order SOFI, we want to reweight the data to a new OTF $\tilde{U}_n(\mathbf{k})$ with the property

$$\underbrace{\tilde{U}_n(\mathbf{k}) \circledast \tilde{U}_n(\mathbf{k}) \circledast \cdots \circledast \tilde{U}_n(\mathbf{k})}_{(n-1) \text{ convolutions}} = \tilde{U}\left(\frac{\mathbf{k}}{n}\right) \tag{2.123}$$

or the equivalent property of the corresponding PSF

$$\left(U_n(\mathbf{r})\right)^n = U(n \cdot \mathbf{r}) \tag{2.124}$$

with the constraint that $U_n(\mathbf{r})$ and $U(\mathbf{r})$ must have the same support in Fourier space in order for the reweighting to be possible[5]. Unfortunately, this constraint is not guaranteed by simply taking the n-th root of $U(\mathbf{r})$. In fact, it is not even guaranteed that such a function exists. The in-focus PSF is given by the Airy pattern [17, 18]

$$U_{\text{airy}}(\mathbf{r}) = \left[\frac{2 \cdot J_1\left(\frac{2\pi}{\lambda_{\text{em}}}\text{NA} \cdot |\mathbf{r}|\right)}{\frac{2\pi}{\lambda_{\text{em}}}\text{NA} \cdot |\mathbf{r}|}\right]^2 \tag{2.125}$$

where J_1 is the Bessel function of the first kind of order one, λ_{em} the emission wavelength and NA the numerical aperture of the microscope. This can be used to compute the exact preweighting OTF for 2nd order SOFI

$$\tilde{U}_2(\mathbf{k}) = \mathcal{F}\left(\sqrt{U_{\text{airy}}}\right) = \text{circ}\left(\mathbf{k}; \frac{\text{NA}}{\lambda_{\text{em}}}\right) = \begin{cases} 1/\left(\pi\left(\frac{\text{NA}}{\lambda_{\text{em}}}\right)^2\right) & |\mathbf{k}| \leq \frac{\text{NA}}{\lambda_{\text{em}}} \\ 0 & |\mathbf{k}| > \frac{\text{NA}}{\lambda_{\text{em}}} \end{cases} \tag{2.126}$$

which reflects the circular aperture of the microscope. The support of this function is exactly half the support of the original OTF. Thus, 2nd order preweighting can be performed by dividing the Fourier transform of the raw data by the original OTF, which produces a circular OTF with uniform magnitude, and setting all frequencies above the support to zero.

Figure 2.33 illustrates the preweighting scheme for the second order. Each frame of the raw movie is "preweighted" by dividing its Fourier transform by the original OTF, multiplication with a new one and back-transforming. Although the resulting PSF produces visible artifacts in the individual frames, it is designed to produce a compressed wide-field PSF shape after the cumulant computation. For noise-less data, this boosts the resolution improvement of n-th order SOFI from $\sqrt{n}$ to the theoretical maximum of linear improvement. For noisy data the new OTF should be scaled to a smaller frequency support (k_{max}), depending on the data quality, to avoid artifacts. We discuss one idea how to choose this support automatically further below.

---

[5]Although having a smaller support is also possible, it is desirable to use all frequencies available in the original data to get the best possible SNR.
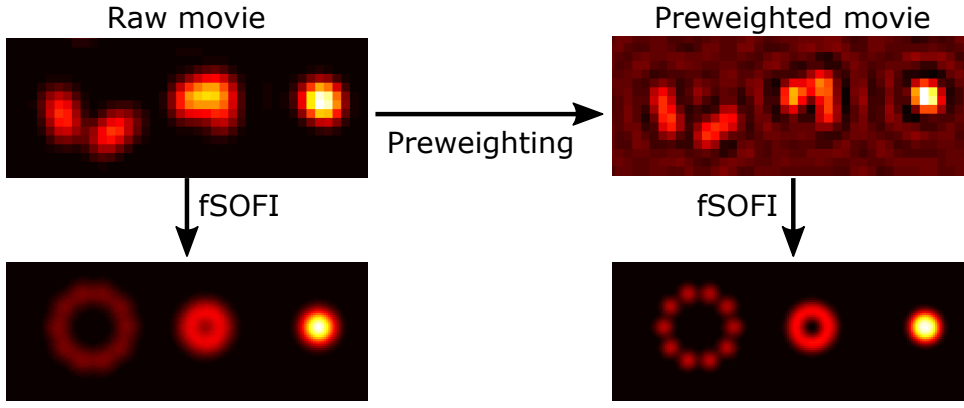
Figure 2.33: Schematic of second order Fourier preweighting illustrated on a noise-less simulation. Each frame of the raw movie is "preweighted" by dividing its Fourier transform by the original OTF, multiplication with a new one and back-transforming. Although the resulting PSF produces visible artifacts in the individual frames, it is designed to produce a compressed wide-field PSF shape after the cumulant computation. For noise-less data, this boosts the resolution improvement of n-th order SOFI from $\sqrt{n}$ to the theoretical maximum of linear improvement. For noisy data, the new OTF should be scaled to a smaller frequency support, depending on the data quality (see text), to avoid artifacts.

For higher orders the situation is more complicated, as taking higher order roots does not yield functions whose Fourier transforms are within the required support, but show even increased support compared to the wide-field OTF. This is even true when the ambiguity in the sign of the n-th root is solved by filling in the known sign of $\sqrt{U_{\text{airy}}}$ to construct the smoothest function. For 4th order SOFI it is clear that no purely real function with finite support exists, that produces the desired shape: The final OTF is created by the auto-convolution of the auto-convolution of this function

$$\tilde{U}_4(\mathbf{k}) \circledast \tilde{U}_4(\mathbf{k}) \circledast \tilde{U}_4(\mathbf{k}) \circledast \tilde{U}_4(\mathbf{k}) = \left[\tilde{U}_4(\mathbf{k}) \circledast \tilde{U}_4(\mathbf{k}) \right] \circledast \left[\tilde{U}_4(\mathbf{k}) \circledast \tilde{U}_4(\mathbf{k}) \right] \qquad (2.127)$$

As a circular OTF is the exact solution for the second order, one auto-convolution of the fourth order OTF must yield a circular function for the auto-convolution of the auto-convolution to yield the exact wide-field PSF. However, no real function with finite support exists that has a constant auto-convolution up to a certain radius and immediately drops to zero. Even taking the complex root to construct $U_n(\mathbf{r})$ (which produces a complex PSF) does not yield the desired support in Fourier space. Thus, an exact solution might not exist for orders higher than the second.
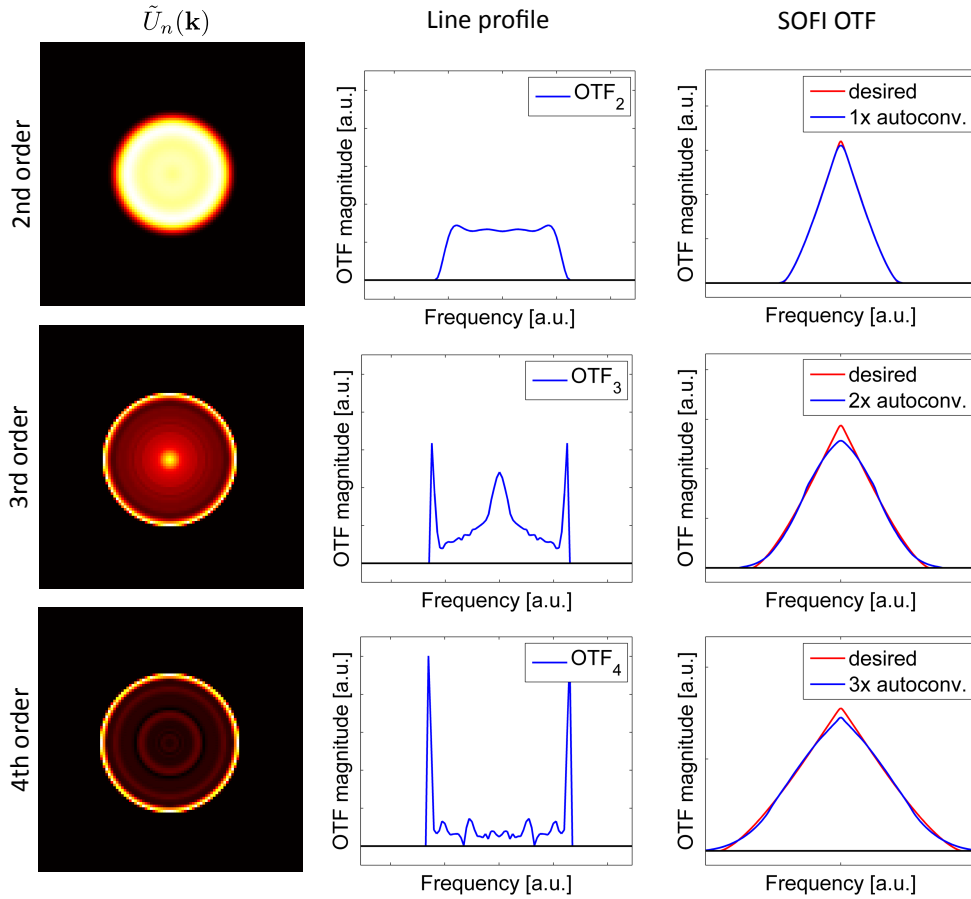
Although an exact solution does not exist, it is possible to construct a good approximation. We used least-squares optimization with a smoothness-enforcing regularization to calculate possible shapes for preweighting OTFs of higher order SOFI, as defined by the cost function

$$C_n \left\{ U_n(\mathbf{k}) \right\} = \sum_r \left[\left(\underbrace{\tilde{U}_n(\mathbf{k}) \circledast \cdots \circledast \tilde{U}_n(\mathbf{k})}_{(n-1) \text{ convolutions}} - \tilde{U}_{\text{airy}} \left(\frac{\mathbf{k}}{n} \right) \right)^2 + a_{\text{smooth}} \cdot \left(\nabla \tilde{U}_n(\mathbf{k}) \right)^2 \right] \qquad (2.128)$$

with the regularization parameter $a_{\text{smooth}} = 15$. Additionally, we enforce radial symmetry as well as the desired support of the solution. We don't enforce the exact support but allow it to be slightly bigger, as this improved the stability of the iterative optimization which is performed using the nonlinear

conjugate gradient algorithm[56] . Figure 2.34 shows the results of the reconstruction. Although the analytical result of the second order is known, we also show its reconstruction to asses the quality of the optimization.

Figure 2.34: OTFs for Fourier preweighting of different orders acquired by least squares optimization. The 2nd order is shown to assess the quality of the optimization only, as the exact shape is known to be a circle of constant magnitude. The high frequencies of the original movie need stronger amplification with increasing cumulant order to yield the desired shape after the SOFI calculation. Although an exact solution for the required OTF might not exist, these results show that it can be reasonably well approximated.

It is evident that a good approximation for the preweighting OTF is possible which very closely matches the shape of the wide-field microscopes OTF after $(n - 1)$ auto-convolutions. The shape of the reconstructed OTF was observed to vary noticeably with the regularization parameter and the used initial values, although the result of the auto-convolutions appeared very similar. However, a common characteristic of all solutions is the strong amplification of high frequencies with increasing cumulant order, leading to visible spikes at the rim of the OTF. This is easy to understand, as the rim of the auto-convolution of a finite support function is mainly produced by the rim of the original function. It also demonstrates, that increasingly high signal-to-noise ratios on the high frequencies are necessary to achieve the desired linear resolution increase for SOFI. Strong irregular ripples in the OTF, as present for the 4th order in Figure 2.34, were observed to produce noticeable artifacts in the SOFI images. Thus, it is desirable to find smoother solutions, which could potentially be achieved by a better algorithm to construct the desired OTF. For now, we smoothed the illustrated OTFs with a small running mean
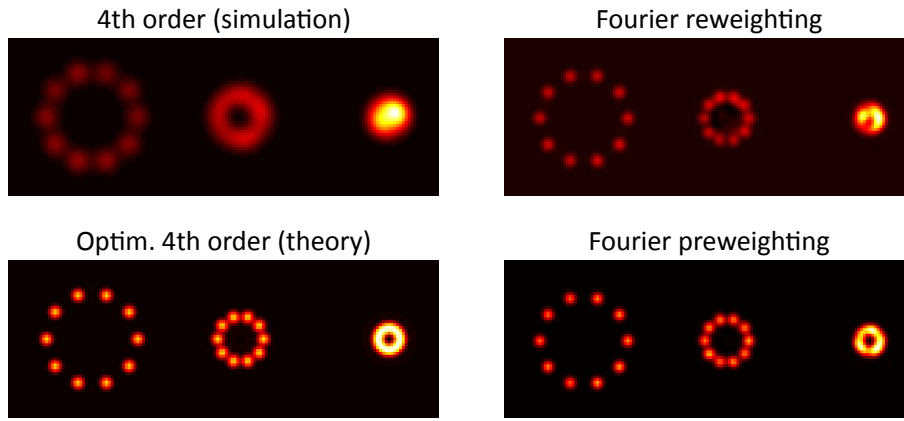
Figure 2.35: Comparison of Fourier preweighting and reweighting on 4th order SOFI images calculated from a noise-less simulation of 20000 frames. The original sample distribution is given by three rings of ten uniformly distributed emitters with decreasing radius. Unlike reweighting, Fourier preweighting is able to resolve the smallest ring and yields results close to the theoretical optimum. Thus it outperforms both reweighting and optimization based deconvolution (compare Fig. 2.31)

filter, which was effective in reducing artifacts without compromising resolution too much.

For noisy data, the preweighting OTF is scaled down to a smaller size, which is effectively a low-pass filter and limits the maximum frequency k_{max} used for the reconstruction. Currently, we select this threshold by requiring a minimum magnitude m of the original OTF $|\tilde{U}(\mathbf{k}_{max})| = m$. A method how to choose this parameter automatically, based on the data quality, is suggested further below. If the threshold is chosen too low, this results in artifacts in the SOFI result that usually show up as switches in sign of the final image values. Similar to reweighting, artifacts also appear if the assumed PSF/OTF is very different from the true one.

Figure 2.35 shows the results of preweighting for the same scenario presented in Fig. 2.32 and Fig. 2.31 for reweighting and optimization based deconvolution, respectively. Fourier preweighting yields a result very close to the theoretical optimum and correctly resolves the single emitters in the central ring as well as the smallest ring structure. Thus it outperforms both reweighting and optimization-based deconvolution and is the only method that comes close to the desired result. The difference to reweighting is strongest for the smallest ring, where the emitter density is the highest.

The higher performance of preweighting in regions of high emitter density is a general property of this method. Surprisingly, processing the preweighted data also improves the homogeneity of the image compared to normal SOFI, as is shown in Fig. 2.36. The underlying sample distribution consists of 11 lines spaced at decreasing distances from left to right. While normal SOFI as well as the reweighted result show artifacts in the high density region, preweighting leads to a more homogeneous image and correctly resolves the structure. This is even more apparent when the number of frames is reduced by almost a factor of 10. Preweighting the data seems to speed up the convergence of cumulant values in the high density regions, which drastically reduces the required number of frames for a proper SOFI image. This is a very important property, as it might allow the calculation of higher order images for samples where the required recordings times were infeasible before. As shown in Section 2.3, higher orders become more and more sensitive to the emitter density. For very long simulations (over 100k

frames) the results of pre- and reweighting become similar to one another.
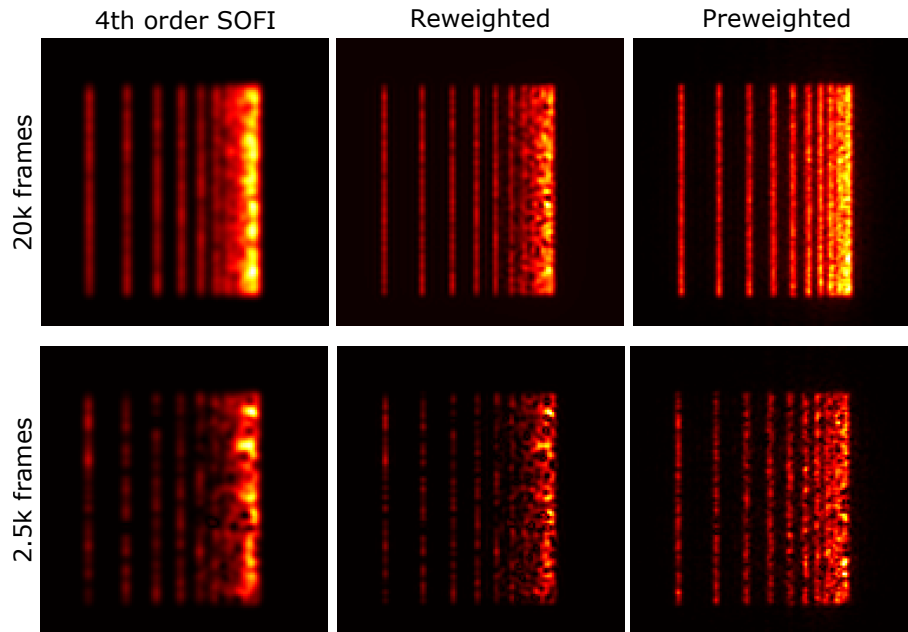
Figure 2.36: Comparison of Fourier reweighting and preweighting on a structure of 11 lines spaced at decreasing distances from right to left. Rows show the achieved results when processing a different number of frames from the same movie. The absolute values are shown in each image, as small parts become negative when fewer frames are processed, which disturbs the visualization. Preweighting results in a more homogeneous image when compared to reweighting.
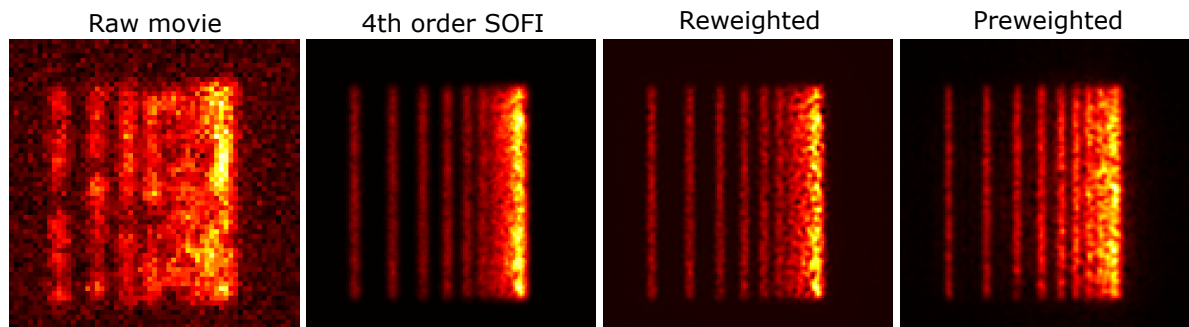
Figure 2.37: Comparison of Fourier reweighting and preweighting for the same simulation as in 2.36 subject to Poisson noise. The brightness of each emitter is 100 photons/frame with 5 photons/pixel background signal. The parameters of both algorithms (*a* and *m* respectively) were individually adjusted to yield the best visual result without artifacts. Although the achievable resolution is lowered by the noise in both cases, preweighting yields an overall smoother result and reconstructs the underlying structure better, especially in the high-density regions.

The same result is true if the data is noisy. Figure 2.37 shows a comparison between normal SOFI, reweighting and preweighting for the same scenario subject to Poisson noise. The parameters of both algorithms (*a* and *m* respectively) were individually adjusted to yield the best visual result without artifacts. Although the achievable resolution is lowered by the noise in both cases, the preweighting result is overall smoother and represents the structure better in the high-density regions.
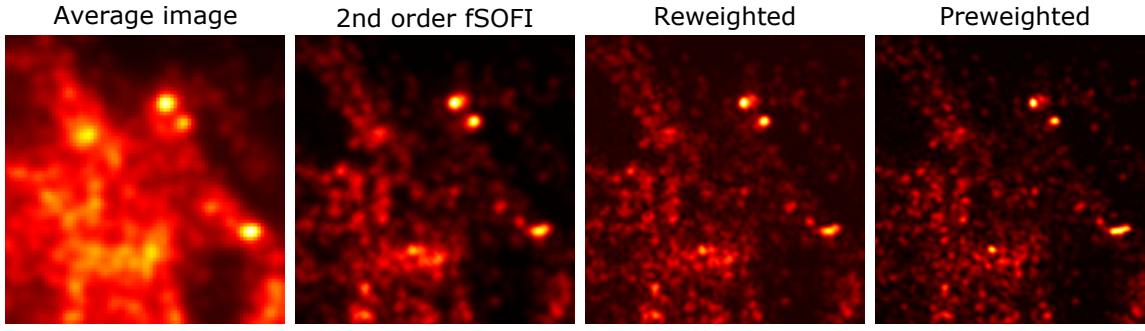
Figure 2.38: Comparison of Fourier reweighting and preweighting for QDot 585 spincoated on a coverslip. The parameters of both algorithms (*a* and *m* respectively) were individually adjusted to yield the best visual result without artifacts. The data was recorded by *Soheil Mojiri*.

Figure 2.38 shows a comparison of the best possible results of both methods on an experimental data of quantum dots (QDot 585) spincoated on a coverslip. The presented results are characteristic for what was observed for other experimental datasets as well. Preweighting achieves a slightly better resolution improvement than reweighting in all cases. A close look at the reweighted image reveals that it seems disturbed in some places, showing structures that appear somewhat artificial. The preweighted result looks more natural in comparison and should be preferred.

While it is apparent that preweighting is a useful tool for increasing the resolution of SOFI images and lowering the recording time of higher orders, especially for high density samples, the parameter *m* must be manually adjusted to the data quality to achieve the best result. We now present an approach that automatically selects a value for the parameter based on the data quality without user intervention. The idea is to calculate the reliability of the spatial frequencies and restrict the preweighting OTF to this reliable range. As will be demonstrated, the method yields good results for simulations as well as experimental data, but is for the most part empirically derived. Also, it is best suited for second oder SOFI; for higher orders fine-tuning might still be required. This suggests that the framework can possibly still be optimized and improved through further research.

To estimate the reliability of the frequencies, we use a process similar to the one introduced to estimate the OTF from the recorded data in Section 2.6.2. According to Eq. (2.97), computing the second order cumulant of the spatial Fourier transform of the data with its complex conjugate yields a function that is independent of the sample distribution (i.e. 2D emitter positions) and depends only on the blinking behavior and the shape of the original OTF:

$$\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t+\tau)} \rangle =$$

$$\varepsilon_j^2 \langle \delta s_j(t) \delta s_j(t+\tau) \rangle \cdot \sum_j \int \int\limits_{-\infty}^{\infty} dk_z dk_z' \, \tilde{U}(k_x, k_y, k_z) \cdot \overline{\tilde{U}(k_x, k_y, k_z')} \cdot e^{i(k_z - k_z')z_j}, \qquad (2.129)$$

where the sum runs over all emitters. Although it is in general difficult to get an accurate estimate of the OTF from this, as the axial position of the molecules is unknown, it still reflects how much signal is collected for the different frequencies. The raw data is cut up into small non-overlapping chunks of h_p frames and this estimate is computed for each individual chunk. It is necessary to only take the real part of this for further processing, as the noise causes small imaginary components in the estimate. In the

next step, the ratio between the mean and the standard deviation of the set of estimates, which we dub the *frequency signal-to-noise ratio*

$$\text{fSNR}(\mathbf{k}) = \frac{\underset{l}{\text{Mean}}\left(\text{Re}\left\{\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)}\rangle_{[l]}\right\}\right)}{\underset{l}{\text{Std}}\left(\text{Re}\left\{\langle \delta\tilde{F}(\mathbf{k}, t) \cdot \overline{\delta\tilde{F}(\mathbf{k}, t + \tau)}\rangle_{[l]}\right\}\right)} \tag{2.130}$$

is computed, where the index l denotes the interval the covariance is calculate for.
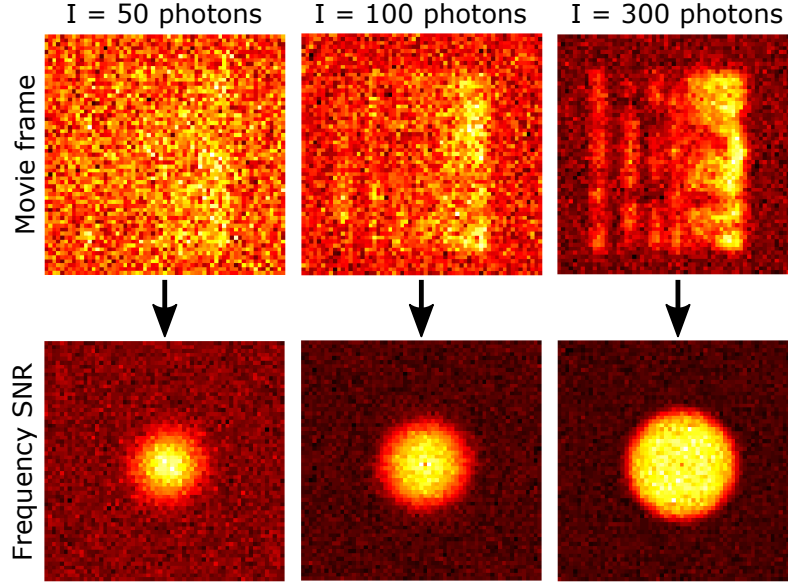
Figure 2.39: Examples of the frequency SNR image (time-lag $\tau = 1$) for different signal-to-noise ratios of the raw movies. As the SNR of the raw movie is increased, the region of frequencies with fSNR values that are significantly different from the noise grows. Within this region, the values are relatively constant and only drop toward the edge of the visible plateau. The interval size h_p used for computing the estimate has only negligible effect for normal photoblinking behaviors.

Examples for the calculated fSNR images for movies with different signal-to-noise ratios are shown in Figure 2.39. A plateau with constant values can be observed which drops off at higher frequencies in which only noise is observed. As the quality of the raw movie increases, the support of the plateau extends to higher frequencies. Smaller intervals or fewer frames used for computing the fSNR only lead to a more noisy appearance with less distinction between the plateau and the surrounding noise., but we did not observe any influence on the plateau size for normal photoblinking behavior. However, for quantum dots, which often show photoblinking where the occupation time of fluorescent states follows a power law distribution and spreads over multiple orders of magnitude, it was observed that the support grows slightly with the interval size.

The estimated frequency support hints towards the optimal size of the transfer function $\tilde{U}_n(\mathbf{k})/\tilde{U}(\mathbf{k})$ which switches from the original OTF to the preweighting OTF by multiplication with the Fourier transforms of the raw frames. The size of the transfer function depends on the size of the preweighting OTF $\tilde{U}_n(\mathbf{k})$ which is scaled to a maximum frequency k_{max} as described above and determines the frequency support. Figure 2.40 shows SOFI images calculated using different transfer function sizes
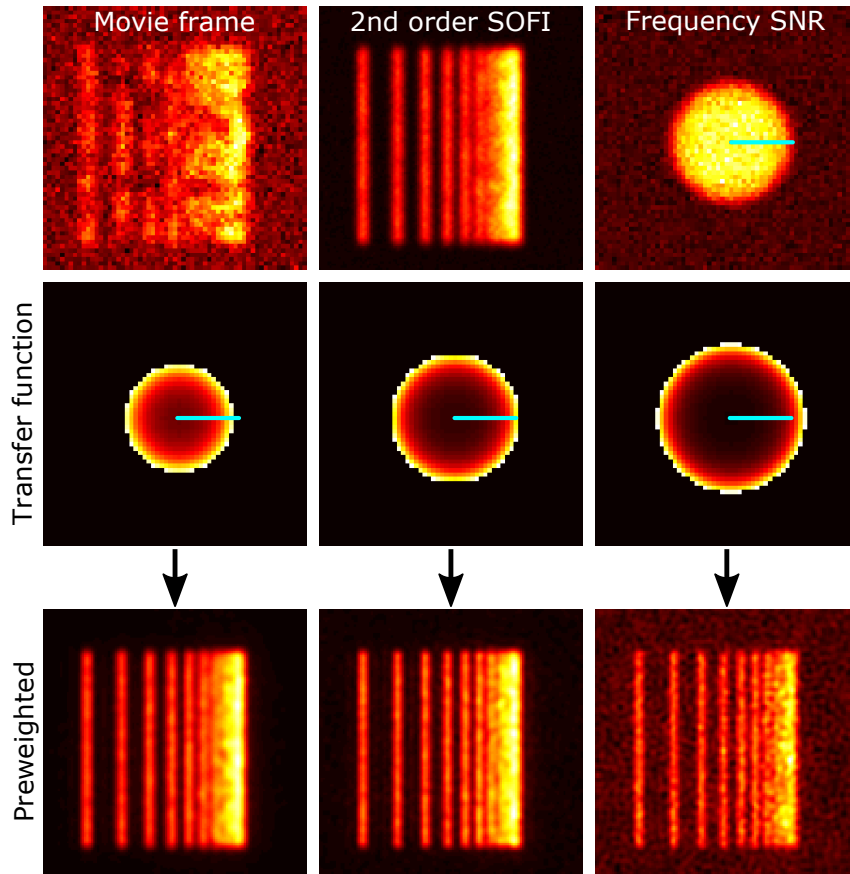
Figure 2.40: Quality of preweighted SOFI for different sizes of the transfer function $\tilde{U}_n(\mathbf{k})/\tilde{U}(\mathbf{k})$ compared to the fSNR image. The cyan line represents the radius of the plateau with significant fSNR values. While too small transfer function do not achieve the optimal resolution enhancement, as they act as a strong low-pass filter, too large transfer functions amplify noise and cause artifacts. If the size is matched to the support of the fSNR image, the same resolution is acquired as for the larger transfer function without these artifacts. This is is an indication that the fSNR image can be used as a hint to determine the optimal transfer function size.

compared to the corresponding fSNR image. Too small transfer functions act as a strong low-pass filter. This does not cause artifacts, but does not extract the maximally possible resolution. It was observed that transfer function sizes that exceed the support of the fSNR image lead to strong artifacts in the final SOFI image, which are caused by the amplification of noise. A close look reveals that the larger transfer function does not resolve more features than the one that matches the fSNR support, but the latter does not produce visible artifacts in the image. This behavior was observed in simulations as well as experimental data. Although this finding is empirical, we assume for now that the optimal transfer function size can be deduced from the fSNR image, which turned out to work well in practice for second order images. For higher orders it can be necessary to reduce the estimated support. An estimator similar to Eq. (2.130) can also be calculated using the fourth order joint cumulant between two times the Fourier transformation of the signal and two times its complex conjugate, where the phase-factors encoding the lateral emitter positions cancel in the same way as for the co-variance used here. In this case we observed a structurally similar image, which shows a smaller support. Thus, more investigations are necessary to adapt the estimator for higher orders.

We use this assumption to construct an algorithm that automatically adjusts the transfer function for
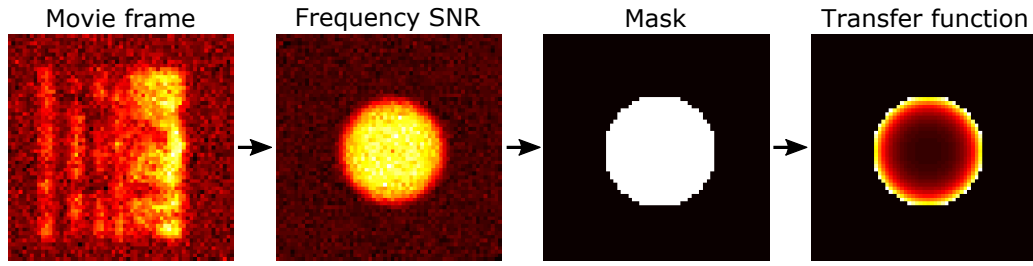
Figure 2.41: Automated generation of the preweighting transfer function based on data quality.

a given data set with known PSF. This is schematically illustrated in Figure 2.41: First, the fSNR image of the movie is calculated. In the second step a mask is generated that reflects the frequency support. A simple threshold selects the region with fSNR values that are greater than $m_{bg} + a_{fsnr} \cdot std_{bg}$, where m_{bg} is the average of all values along the fSNR image border and std_{bg} their standard deviation. All results presented here are computed with the threshold parameter set to $a_{fsnr} = 2$. The resulting binary mask is filtered to smooth outliers at the edge as well as fill eventual holes, using a process of binary erosion, dilation and again erosion. Then, the maximum frequency k_{max} of the mask is extracted. In the third step the transfer function is calculated using the preweighting OTF and scaled to match the extracted k_{max}. It is important to bear in mind that the Fourier transform of non-square images leads to differently scaled frequency axes, requiring a proper scaling of the transfer function when processing such data. The algorithm is very basic, but works reliably in most cases.
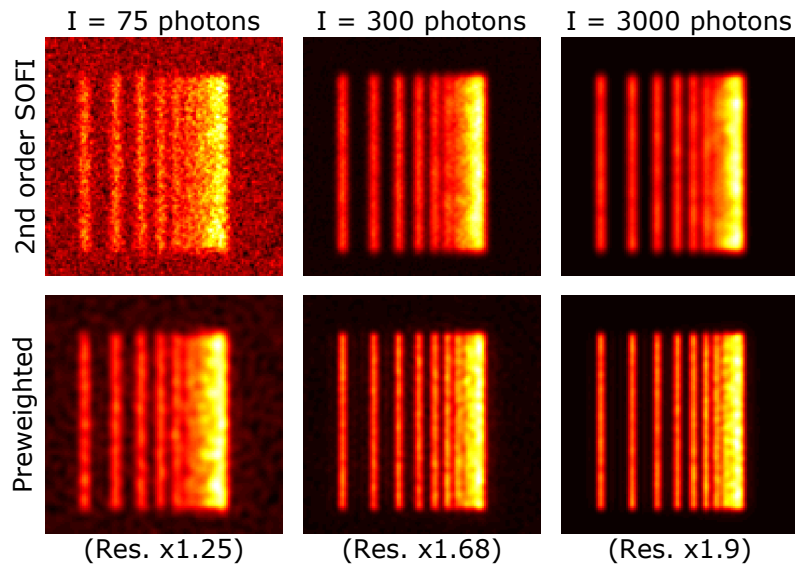
Figure 2.42: Results of preweighted SOFI with adaptive transfer function for different data quality. Also shown is the calculated resolution improvement which is inferred from the fSNR support mask (see text). The achieved resolution improvement increases with the SNR of the raw movie and converges toward an n-fold improvement for n-th order SOFI, as theoretically expected. If the fSNR support is too low to support the resolution improvement of factor $\sqrt{n}$, which is expected for SOFI, the standard SOFI images become very noisy. In this case the preweighted result still yields a good image with improved resolution and z-sectioning. The artifacts seen here are caused by the discretization (see text).

As an interesting byproduct of the preweighting process, the resolution of the final n-th order SOFI image can be easily inferred from the estimated mask, as it is simply the inverse of the cutoff frequency

$n \cdot k_{max}$ of the final OTF[6]. For most super-resolution methods, estimating the achieved resolution is very difficult, especially on experimental data. The accuracy of the estimated resolution depends mainly on the accuracy of the measured PSF. Figure 2.42 shows a comparison of standard 2nd order fSOFI to results based on preweighting using the automatically adjusted transfer function for different data qualities. The estimated resolution increase compared to the wide-field resolution is also shown. For low SNR, the standard SOFI image is full of noise. Most of these artifacts are suppressed in the preweighted result. The residual artifacts can be attributed to the discretization of the 61x61 pixel movie, as the fSNR support at this pixel resolution is only a few pixels in radius. Thus, the mask is either a bit too small or too big depending on the threshold parameter a_{fsnr}. The calculated resolution improvement is a factor of 1.25 which is lower than the $\sqrt{2} \approx 1.41$ expected for second order SOFI. We observed, that normal SOFI images always show strong noise if the fSNR support is too small to support the theoretically expected resolution improvement. In this case, preweighting still yields an improvement and possible z-sectioning, although at lower resolution improvement than $\sqrt{n}$. The achieved resolution improvement increases with the SNR of the raw movie and converges toward an n-fold improvement for n-th order SOFI, as theoretically expected.
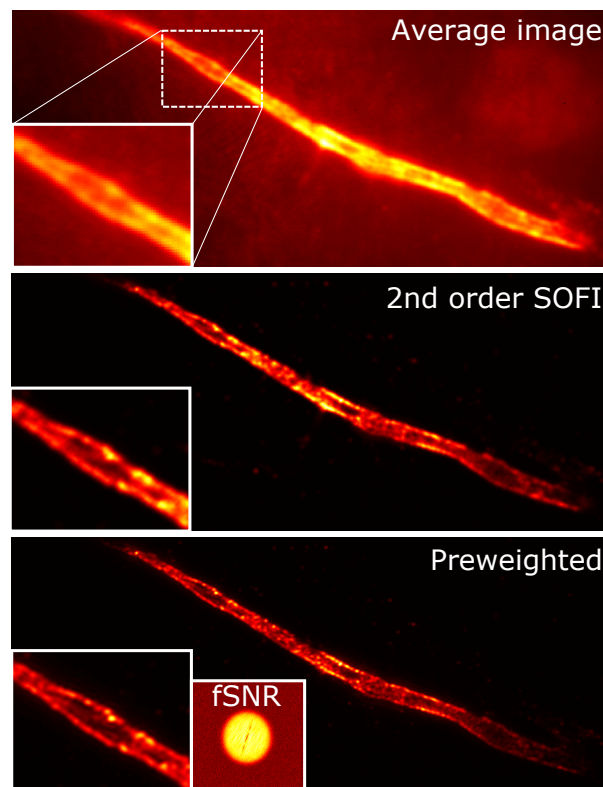
Figure 2.43: Example result of adaptive preweighted SOFI on experimental data of a rat hippocampal neuron axon initial segment with βIV-spectrin in the cytoskeleton labeled with Alexa647. The calculated resolution of the reweighted image is improved by a factor of 1.83 compared to the wide-field PSF which was determined from single particles in the movie. The improvement compared to normal SOFI is clearly visible in the inset and the preweighted image is without noticeable artifacts. Data recorded by *Anja Huss*.

Figure 2.43 shows results of adaptive preweighting on experimental data of a rat hippocampal neu-

---

[6]As it is essentially shaped like the optimal wide-field OTF.

ron axon initial segment with βIV-spectrin in the cytoskeleton labeled with Alexa647. The PSF was determined from single particles in the raw movie. The calculated resolution of the reweighted image is improved by a factor of 1.83 compared to the original wide-field data. A look at the inset also reveals the clear improvement compared to normal SOFI.
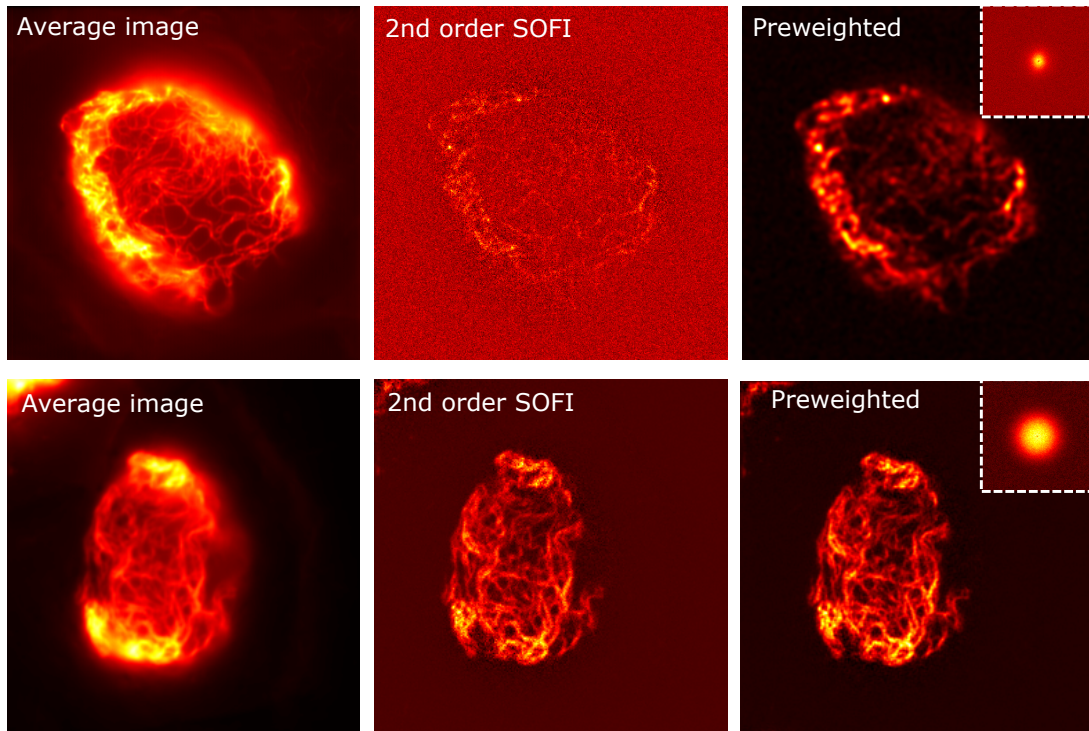
Figure 2.44: Example result of adaptive preweighted SOFI on low quality experimental data showing a HeLa cell with vimentin labeled by Dreiklang (top row) and rsEGFP (bottom row). Data recorded by *Anja Huss*.

Figure 2.44 shows results of adaptive preweighting on low quality experimental data showing a HeLa cell with vimentin labeled by Dreiklang (top row) and rsEGFP (bottom row). In both cases, artifacts are visible in the normal SOFI image. This is most prominent for the Dreiklang movie, which also has a much lower fSNR support. These artifacts are removed in the preweighted results. The calculated improvement factor is 0.47 for the Dreiklang data and 1.37 for rsEGFP. In the case of Dreiklang the resolution is lower than the original wide-field resolution, but still with improved z-sectioning due to SOFI. Note that Fourier reweighting cannot achieve this result as it is only applied to the very noisy SOFI image, which causes artifacts. For rsEGFP, the estimated resolution improvement is slightly worse than the theoretically expected $\sqrt{2}$, which explains the similarity between the normal and preweighted SOFI results. Still, preweighting is to be preferred as it filters the existing artifacts. This shows that preweighting works even in low quality scenarios for which normal SOFI does not yield a useful result and truly adapts to the data.

Fourier preweighting can be combined with Fourier interpolation as introduced in Section 2.4 without any alterations. We also tried using spatial cross-cumulants on the preweighted data, but this results in lots of artifacts in the final image. Combination with the noise correction in Section 2.5 requires some additional work. Here we briefly derive the procedure for 2nd order SOFI:

Let us take the imaging model

$$F(\mathbf{r}, t) = F_S(\mathbf{r}, t) + n(\mathbf{r}, t), \tag{2.131}$$

where $F_S(\mathbf{r}, t)$ represents the noise-less signal of interest and $n(\mathbf{r}, t)$ contains all noise components (e.g. background noise and zero-mean signal-dependent noise). Preweighting in Fourier space is performed by multiplication with the transfer function $\tilde{T}(\mathbf{k})$

$$\begin{aligned}
\tilde{F}^*(\mathbf{k}, t) &:= \tilde{T}(\mathbf{k}) \cdot \tilde{F}(\mathbf{k}, t) = \tilde{T}(\mathbf{k}) \cdot \tilde{F}_S(\mathbf{k}, t) + \tilde{T}(\mathbf{k}) \cdot n(\mathbf{k}, t) \\
&= \tilde{F}_S^*(\mathbf{k}, t) + \tilde{T}(\mathbf{k}) \cdot n(\mathbf{k}, t) \tag{2.132}
\end{aligned}$$

Back-transforming into real space gives

$$F^*(\mathbf{r}, t) = F_S^*(\mathbf{r}, t) + T(\mathbf{r}) \circledast n(\mathbf{r}, t). \tag{2.133}$$

When calculating the 2nd order cumulant, cross-terms between the signal and the noise drop out, as they are statistically uncorrelated. This leaves us with

$$\langle \delta F^*(\mathbf{r}, t) \rangle = \langle \delta F_S^*(\mathbf{r}, t) \rangle + \left\langle (T(\mathbf{r}) \circledast \delta n(\mathbf{r}, t))^2 \right\rangle \tag{2.134}$$

$$= \langle \delta F_S^*(\mathbf{r}, t) \rangle + \int \int T(\mathbf{r} - \mathbf{r}') T(\mathbf{r} - \mathbf{r}'') \left\langle \delta n(\mathbf{r}', t) \delta n(\mathbf{r}'', t) \right\rangle, \tag{2.135}$$

as the noise is spatially uncorrelated this gives

$$\langle \delta F^*(\mathbf{r}, t) \rangle = \langle \delta F_S^*(\mathbf{r}, t) \rangle + T^2(\mathbf{r}) \circledast \left\langle \delta n^2(\mathbf{r}, t) \right\rangle. \tag{2.136}$$

The expression $\langle \delta n^2(\mathbf{r}, t) \rangle$ is the second central moment of the noise distribution. For a Poisson distribution, this is just the time average image of the movie, which is also the correction term introduced in Section 2.5. To correct the preweighted image, this average image has to be convolved with the square of the Fourier transform of the transfer function $T^2(\mathbf{r})$ before subtraction.

Although the presented algorithm is very useful in practice, it can probably be improved by further research. As it is based on empirical observations, a more theoretical approach is needed to check what can be improved (e.g. how to set the empirical thresholding parameter for mask generation) or if an entirely different approach might be favorable. The current implementation also works only globally on the whole image, rather then improving the resolution locally based on data quality. This could be useful for higher orders, where the results depend more on density. More investigations are also needed how the length of movies influences the result. For very long movies it is expected that a linear resolution increase is possible and thus should be achieved by every method that tries to improve the SOFI resolution. Another interesting topic is to analyze the imaging speed in the context of different sizes of the transfer function. It seems that the convergence of SOFI values might be faster for smaller transfer functions (stronger low-pass filtering), which would enable to trade spatial resolution against time resolution. This is especially interesting when recording 3D movies with wide-field illumination [28], where

the gained *z*-sectioning is often more important than the lateral resolution increase, possibly allowing 3D structures to be imaged at high speeds.

To sum up this section about improving the resolution of SOFI images,

- Theoretically, the resolution improvement of *n*-th order SOFI can be boosted from a factor of $\sqrt{n}$ to *n*.

- Optimization-based deconvolution did not yield useful results so far. The resulting images should be interpreted with care and while keeping their probabilistic meaning in mind. In many cases, the resolution of these images can not be defined well, at least in the context of the optical definition of resolution.

- Fourier reweighting [40], which post-processes the SOFI image, yields better results, but often suffers from artifacts and needs a user adjusted parameter.

- Fourier preweighting changes the OTF in the raw movie by multiplication with a transfer function in Fourier space, to achieve a PSF in the SOFI image which is shaped like the original in-focus PSF shrunken in size.

- This approach always yields better performance than Fourier reweighting, especially in regions of high emitter density. In those regions, the preweighted image shows, surprisingly, much better results and quicker convergence of the cumulant values than normal SOFI images. This can drastically decrease the required recording time at high densities, especially for higher SOFI orders.

- The cutoff frequency of the transfer function is varied by scaling of the function and needs to match the quality of the input data to avoid artifacts / noise amplification. A method is proposed how to automate this scaling based on the data quality. The cutoff frequency can be used to easily compute the actually achieved resolution of the final image. This is an advantage over many other super-resolution methods, as estimating the resolution is usually very difficult.

- The estimated cutoff frequency works best for second order SOFI; for higher orders a reduction of the transfer function size might still be necessary to avoid artifacts. For cumulants of even order a similar estimator to the one proposed here can be constructed. This might adapt better to the quality that is achievable with specific cumulant orders.

2.8 SOFI WITH CARBON NANODOTS

Luminescent carbon nanodots (CNDs) have emerged in the last decade as novel probes for fluorescence microscopy. They exhibit many properties of classical semiconductor based quantum dots, like size and wavelength dependent emission and greater resistance to photobleaching, but without their inherent toxicity and are cheap to make as well as easy to produce [61]. This combination of properties is very interesting for imaging applications and thus has drawn increased attention in the imaging community. Super-resolution imaging using CNDs has been demonstrated in a recent publication using STED microscopy [62]. Some types of carbon nanodots show reversible photoswitching [63], which inspires ideas to use CNDs also for stochastic super-resolution methods. However, localization based methods require high stability and control over the photoswitching behavior to achieve the desired spatio-temporal separation of dyes. While this is nowadays possible for many conventional dyes and fluorescent proteins after years of research, it is currently difficult to control the behavior of carbon nanodots.

In contrast to SMLM methods, the demands on sample density and photoswitching behavior are more relaxed in SOFI [27]. Because SOFI microscopy requires only a wide-field microscope and a fast recording camera, it is essentially a low-cost method. This is a perfect fit for the cost-effective nanodots which in combination allows super-resolution imaging at very affordable prices and enables more labs to access this technology.

We synthesized CNDs, examined their photoswitching, and demonstrated the first application of SOFI to cells labeled with dual-color CNDs. We also characterized various other properties, such as size, charge and quantum yield of the probes and give a hypothesis on the photoswitching mechanism. This work was published in the article [64]:

Chizhik, A.M.; **Stein, S.C.**; Dekaliuk, M.O.; Battle, C.; Li, W.; Huss, A.; Platen, M., Schaap, I.A.T.; Gregor, I.; Demchenko, A.P.; Schmidt, C.F., Enderlein, J.; Chizhik, A. I.; "Super-Resolution Optical Fluctuation Bio-Imaging with Dual-Color Carbon Nanodots", Nano Letters, 16, 237-242, 2016.

The analysis of the photoswitching behavior, as well as SOFI related data evaluation were performed by me and will be briefly described here. Parts of this section are taken from the publication. All experiments as well as further analysis illustrated in the publication were performed by the co-authors.

The carbon nanodots were manufactured by a temperature-treatment-based synthesis [64, supplementary information], which yields two species with different emission wavelength at 420 nm (blue) and 515 nm (green) respectively. Laser-scanning confocal microscopy using an azimuthally polarized excitation laser revealed a linear excitation dipole moment as well as fluorescence intermittency for both species. To analyze the photoswitching statistics, we spincoated a 10 μl droplet of aqueous solution of CNDs on the surface of a fused silica coverglass at 6000 rpm rotational speed. The glass was verified to be free of contamination by imaging it under identical excitation conditions, before the single particle study was conducted. We used an epi-fluorescence microscope with a 473 nm continuous wave

laser for illumination of a field of view of $80 \times 80 \, \mu m^2$ with an irradiance of $500 \, W/cm^2$ through a water-immersion objective (NA=1.2, 60X). The emitted light from the sample is collected by the same objective, travels through a FF01-550/88 (Semrock) bandpass filter and is imaged onto an EMCCD camera.
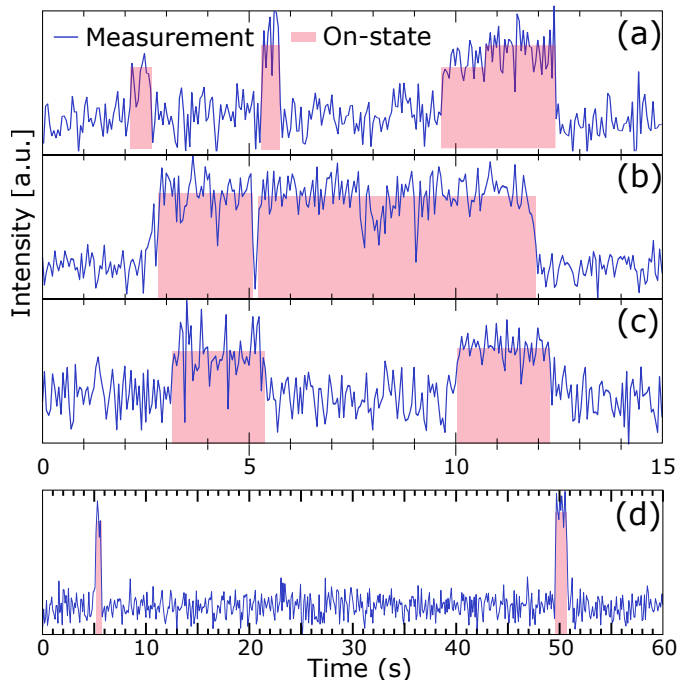
Figure 2.45: Time traces of single carbon nanodots spincoated on a cover glass surface. While most emitters show only a single fluorescent state, in some cases multiple emission states were observed (a). The visible single-step transition together with experiments using defocused imaging (see text) suggest that these are indeed single particles. The occupation time of the on- as well as the dark state cover a broad range of values from 50 ms (the exposure time) to multiple tens of seconds (d).

To extract the intensity time traces of individual CNDs, emitter candidates were identified as local maxima in the time average of the movie. A threshold was applied to select only candidates which could be clearly identified as particles instead of noise. Intensity estimation was performed by integration of an 7x7 window centered around each detected particle and subtracting the background, which is assumed as constant within the window and estimated by the average along the window border. Different fluorescent states were extracted using a recently published step-detection algorithm [65], which tries to distinguish steps from local noise and is completely parameter free. Figure 2.45 shows a few examples of states recovered by the method. As the first time trace shows, some of the CNDs exhibit several different fluorescent states. The single-step behavior suggests that this is indeed a single particle and not an aggregation of multiple emitters. However, most emitters exhibited only one intensity level. The occupation time of the fluorescent states is subject to broad variations from 50 ms (the exposure time) to tens of seconds. The length of dark state segments show similar variations. Figure 2.45 (d) shows two emission states which are separated by about 45 seconds. To check if we indeed image single particles, a similar experiment was performed recording defocused images of the sample. The observed patterns match the patterns expected for single dipoles and also showed no reorientation upon photoswitching, which suggests the absence of aggregates.
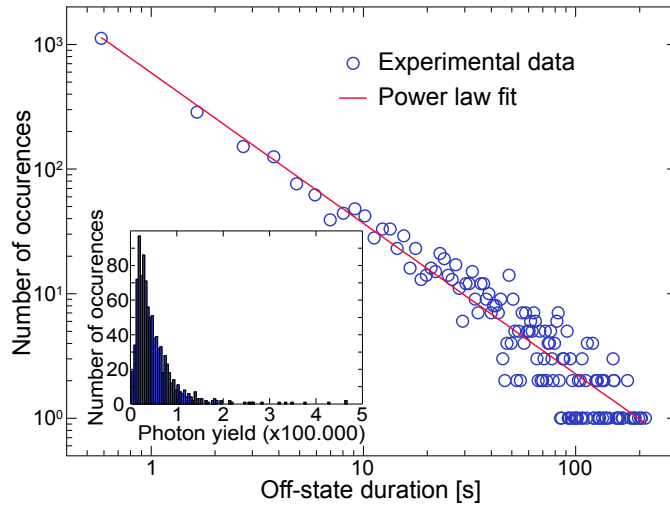
Figure 2.46: Distribution of the off-state occupation time from 964 single particles. The distribution can be well fitted with a power law $a \cdot t^b$ with $a = (568.4 \pm 4.2) \frac{1}{s}$ and the time constant $b = -1.247 \pm 0.012$. Inset: The distribution of the overall photon yield of the carbon nanodots. The average number of detected photons is 51253. [This image has been published in the article [64].]

The statistics of the photoswitching allow speculations about the mechanism that causes the phenomenon. Conventional dyes typically show exponential distributions for the occupation time of the optically dark triplet-state and the active singlet-state [9]. In contrast, semiconductor based quantum dots, where photoswitching is mediated by the trapping and release of (surface) charges, show power law distributions which can span multiple orders of magnitude [66]. To exclude the influence of the multi-state fluorescence on the analysis, we only look at duration of the off-time between clearly fluorescent states. Figure 2.46 shows the resulting distribution from 964 single particles together with their overall photon yield. The distribution could be well fitted with a power law distribution $a \cdot t^b$ with time constant $b = -1.247 \pm 0.012$ and an average dark state occupation time of 13.625 s within the observation window. Exponential statistics can be clearly ruled out and the quality of the fit suggests that a single dark state is enough to describe the system. This suggests that an electron transfer mechanism might play a larger role for the fluorescence intermittency of CNDs. For further discussion of this mechanism, please refer to [64].

To test the suitability of carbon nanodots for SOFI microscopy, Sarcoma osteogenic (Saos-2) and Madin-Darby Canine Kidney Epithelial (MDCK-II) type cells were incubated with a mixture of the dual-color CNDs at 37° for 1 h. Both cell lines took the nanodots in spontaneously without further manipulation. Imaging was performed with a wide-field microscopewith 395 nm and 480 nm LED epi-illumination and an 1.4 NA 100x immersion oil objective. Post magnification of 1.6x resulted in a pixel size of 100 nm per EMCCD pixel. We excited both colors of CNDs at the same time and recorded a movie of 783 frames with 100 ms exposure time before applying SOFI.

Figure 2.47 shows a comparison between standard wide-field imaging, 2nd order SOFI and a linearized SOFI image that was recovered in post-processing to attenuate the intensity skewing introduced by SOFI and facilitate comparison. The linearization is based on the (strong) assumption that the intensity differences in the image are only caused by local brightness changes of the CNDs, while density
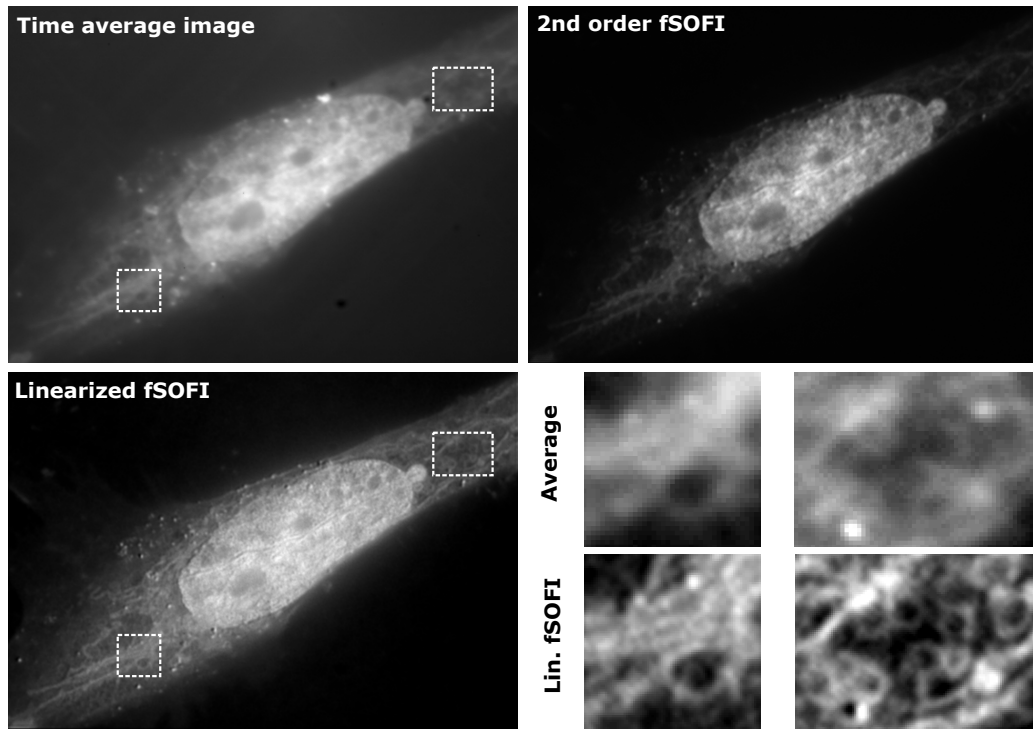
Figure 2.47: Comparison of wide-field image and second order SOFI. Additionally, a linearized SOFI image was computed to attenuate the intensity skewing of SOFI using a least squares optimization (see text). Bottom left: two zoomed in views corresponding to the white boxes in the left images. The contrast is individually adjusted for each zoomed view, but with identical settings for the wide-field and the linearized SOFI cutout to allow comparison. It is easy to see that the SOFI image looks clearer and exhibits more detail. This is partly due to the higher resolution and partly due to the background removal inherent to SOFI.

and blinking behavior are assumed to be homogeneous within the sample. Then, the least-squares cost function

$$C_{\mathrm{LSQ,CND}} \left\{ \varepsilon(\mathbf{r}) \right\} = \sum_r \left(U^2(\mathbf{r}) \circledast \varepsilon^2(\mathbf{r}) - i(\mathbf{r}) \right)^2 \qquad (2.137)$$

with the wide-field PSF $U(\mathbf{r})$, the local brightness $\varepsilon(\mathbf{r})$ and the second order SOFI image $i(\mathbf{r})$ sufficiently describes the imaging process and can be used to recover a linear intensity. The optimization problem is solved using the nonlinear conjugate gradient algorithm [56]. Furthermore, in every optimization step, the Fourier-transformed image is projected onto the finite support of the optical transfer function of the SOFI image, so that no non-physical frequencies are recovered and the image resolution is untouched. Any perceived difference in sharpness between the original SOFI and the linearized image are due to the changed intensity and can be locally achieved by contrast adjustments. Although the used assumptions are very strong, the recovered image shows a brightness distribution that closely matches the wide-field image. As the linearized image is locally identical to a contrast adjusted SOFI image and showed the exact same geometry without artifacts, it can be trusted for further analysis.

The SOFI image shows clearly finer detail and better contrast than the wide-field image, which is partly due to the increase in resolution and partly due to the background reduction which is inherent to SOFI. This is especially true for the network-like structures surrounding the nucleus, as depicted in the bottom right of the figure. Cross sections through dot-like structures in the vicinity of the cell suggest a

resolution increase of factor 1.40, which matches 2nd order SOFI. This indicates that the resolution was not corrupted by other effects, such as bleaching or drift of the sample. As for many types of semiconductor quantum dots, the power law statistics prevented the calculation of useful higher order images, which exhibited local changes in the sign of the computed cumulant. This is a consequence of the (non-ergodic) broad distribution of occupation times which hinder sufficient convergence of the cumulants within the recording time.

In conclusion, this demonstrates that carbon nanodots are suitable probes for SOFI microscopy, which allows super-resolution imaging at very low cost. However, the utilized nanodots can only be used for non-specific labeling right now and further research is needed to enable more specific applications as well as control of the photoswitching behavior.

To sum up this section about SOFI with carbon nanodots:

- Carbon nanodots are cheap to make fluorophores which share many properties of semiconductor quantum dots without their inherent toxicity.

- We synthesized dual-color CNDs and observed a power law distribution for the occupation of the dark state. This suggests that an electron transfer mechanism might play a large role for the fluorescence intermittency of CNDs, instead of the triplet-state photoblinking with exponential statistics that is observed for conventional dyes.

- Using the fluorescence intermittency, we demonstrated the first application of SOFI with CNDs on unspecifically labeled cells, which enables super resolution imaging at very low cost.

<div style="text-align: right; font-size: 3em;">3</div>

Super-resolution cryo-fluorescence microscopy

Tʜᴇ first light microscopes designed to image samples at very low temperatures were developed almost half a century ago, to study freezing and thawing processes in cells [67]. Investigations at cryogenic temperatures are especially interesting for biological applications, as rapid freezing induces vitrification, allowing the fixation of samples without the structural changes commonly caused by chemical techniques [68]. Proper fixation is crucial for very high resolution methods like cryo-electron microscopy, which resolves structures down to the sub-nanometer scale at optimal conditions [69]. Very low temperatures also increase the sensitivity of spectroscopic measurements, as the absorption and emission band of molecules is considerably narrowed, making the detection of weak effects possible [70]. With the advent of super-resolution methods in recent years, research in the field of cryo-fluorescence microscopy has seen increased interest [71]. One of the most exciting topics is correlative microscopy, the combination of electron and light microscopy. While cryo-EM excels in terms of resolution, it is often difficult to identify areas of particular interest. Fluorescence microscopy is highly specific, but exhibits a resolution that is two orders of magnitude lower. Super-resolution cryo-fluorescence microscopy could diminish this resolution gap and holds the potential for correlative microscopy that reveals ultra-structural details with the specificity of fluorescence labeling [72]. Currently, appropriate immersion objectives for low temperatures are missing and commonly air objectives with a numerical aperture of 0.8 are used, which collect only about half the photons of a well adjusted oil immersion objective [71]. However, this disadvantage is somewhat counteracted by the changed photophysics of the fluorophores: At low temperatures photobleaching is heavily reduced, leading to considerably longer observation times and much higher overall photon yield [70, 73].

The first section of this chapter demonstrates sub-nanometer localization of single Atto647N mol-

ecules using a custom build state-of-the-art cryo-fluorescence microscope, which was developed by *Weixing Li* as part of his Phd thesis [74]. In the second section, a concept for accurate super-resolution co-localization of up to three fluoropohres in close proximity to each other is presented and discussed.

## 3.1 Single molecule localization with sub-nanometer precision

Imaging at cryogenic temperature allows the collection of millions of photons from a single fluorophore. Each photon that hits the detector is essentially a measurement of the emitters position. In principle, for a freely rotating fluorophore in the focal plane of an aberration-free imaging system, the accurate position of the fluorophore could be determined by simply averaging the positions of all detected photons. As is known from basic statistics, the standard error of the arithmetic mean is proportional to $\frac{1}{\sqrt{N}}$, where N is the number of measurements. In practice, positions are usually extracted by fitting a suitable model of the PSF to the data, but this proportionality still holds [26]. Due to this, the increased photon yield at low temperatures enables a localization precision that easily surpasses room temperature measurements.

We demonstrate sub-nanometer localization precision using Atto647N with liquid nitrogen cooling, which was published in the article [75]:

Weixing, L.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Enderlein, J.; "Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy", Optics Express, 23, 3770-3783, 2015.

All experiments were conducted by *Weixing Li*, while I developed most of the data evaluation, which is briefly discussed here. Our experiments show that cryo-fluorescence microscopy has the potential to achieve a localization precision close to the physically reasonable limits. Although infinite resolution could be achieved in principle, the finite length of the linker that attaches the fluorophore to the structure of interest is usually already larger than a few nanometers, which restricts more accurate measurements. Similar precision was demonstrated by Weisenburger *et al* using liquid helium cooling in 2013 [76]. Despite the less expensive coolant, we use a different drift correction algorithm that does not require tracking beads in the sample. Furthermore, their approach is based on sub-pixel shifts of the raw images to a steady reference frame using interpolation before the positions are localized from multiple summed-up images. This can introduce additional bias to the position estimation. We also do not require a correction factor for the measured number of photons to match our results to the theoretically expected precision, which is necessary in their case due to aberrations of the PSF.

For the study, an aqueous solution of Atto647N molecules was spin-coated on a clean fused silica coverslip. Cleaning was performed by flaming with a Bunsen burner and the surface hydrophilized in a plasma cleaner. The sample was excited by a 637 nm laser with an irradiance of about $300\,\mathrm{W\,cm^{-2}}$ and movies were recorded using an *Andor iXon Ultra 897DU* EMCCD camera at room temperature (294 K) and cryogenic temperature (89 K) with frame rates of 0.5 Hz and 1 Hz respectively. Single-molecule localization was performed using the open source software *rapidSTORM* 3.2 [77]. To reduce noise in
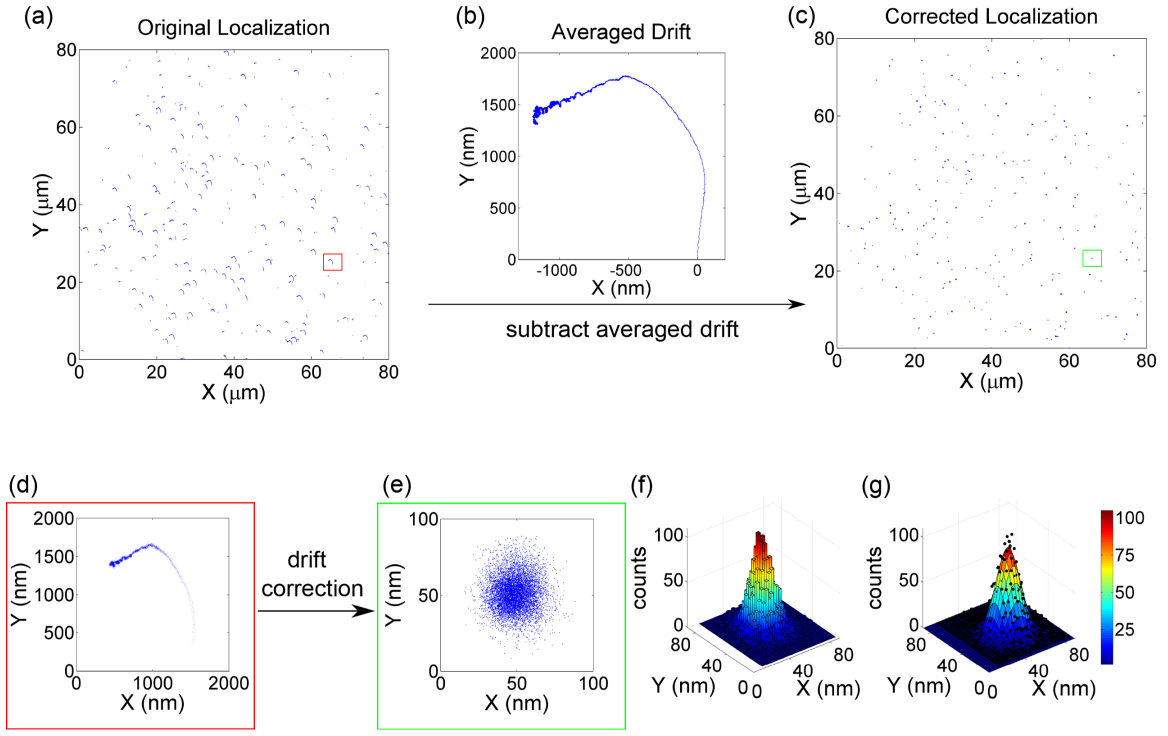
Figure 3.1: Scheme of localization and drift correction. (a) Localizations of single Atto647N molecules acquired by rapidSTORM. (b) After nearest-neighbor tracking, the drift is calculated as the average displacement of molecules. (c) Drift-corrected localizations. (d) Localizations of a single molecule before and (e) after drift correction. (f) The two-dimensional histogram of the localizations is well-fitted by a (g) Gaussian distribution (black dots). [This image has been published in the article [75].]

the acquired data, a statistical filter removes outliers by computing the distance d_p of every point p to its K closest neighbors and excluding points with

$$|d_p - \overline{d}_{\text{dataset}}| > c_{\text{cutoff}} \cdot \sigma_d, \tag{3.1}$$

where $\overline{d}_{\text{dataset}}$ is the mean of the K-neighbors distances and σ_d their standard deviation. We used $K = 50$ and the cutoff parameter $c_{\text{cutoff}} = 1$ for this step in all experiments. Next, nearest-neighbor particle tracking was used to identify the movement of all particles and the average drift is calculated from all tracks with a minimum length of 300 frames. After subtracting the drift from the data, it is again processed by the statistical filter with $c_{\text{cutoff}} = 0.1$ to remove remaining outliers in the vicinity of clusters of localizations that correspond to single molecules. In the last step, a euclidean clustering algorithm groups all points that are closer than 200 nm, which was much smaller than the inter-molecular distance but much larger than the distance between localizations of the same emitter, yielding a cluster of localizations for each fluorophore. The scheme for drift correction is illustrated in Fig. 3.1.

This sophisticated data processing was necessary, as localizations returned by rapidSTORM were either too noisy or removed non-negligible parts of the desired data, depending on the software's settings. The reason for this is the simple local-maximum-based algorithm for detection of molecule candidates, which is very fast but prone to noise. While the resulting errors are usually negligible when generating STORM images, they had to be taken into account for accurate drift correction and clustering. Similar
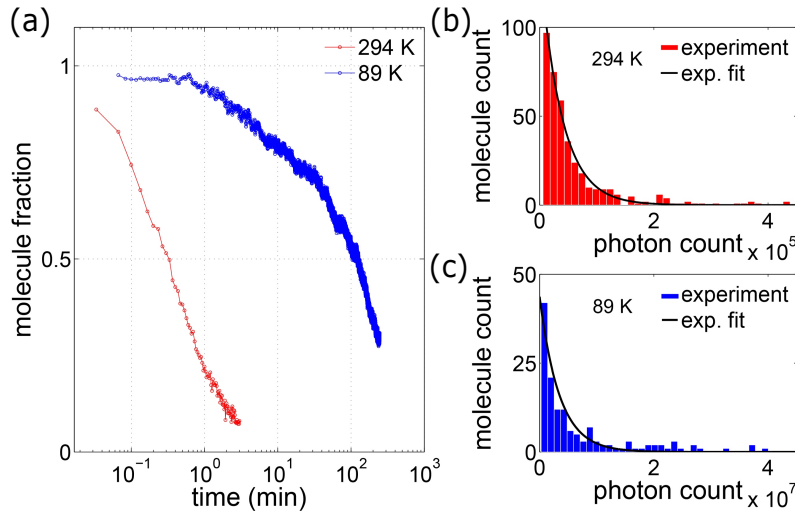
Figure 3.2: Comparison of survival time and photon yield for Atto647N at cryogenic and room temperature. (a) At 294 K, the number of detectable molecules drops below 25% of its initial value within the first minute, while at 89 K over 50% of molecules are still fluorescing after 100 minutes. The photon yield at room temperature (b) and cryogenic temperature (c) is well fitted by an exponential decay. The fits gave an average photon count of $3.8 \cdot 10^4$ and $3.5 \cdot 10^6$ for 294 K and 89 K respectively; an increase over two orders of magnitude. [This image has been published in the article [75].]

results can be achieved by using a more sophisticated candidate detection method, for example cross-correlation with the theoretical in-focus PSF in *TrackNTrace* (Chap. 4).

Figure 3.2 shows the enhanced photostability at cryogenic temperatures compared to room temperature. At 294 K, the number of detectable molecules drops below 25% of its initial value within the first minute, while at 89 K over 50% of molecules are still fluorescing after 100 minutes. The amplitude estimate A from rapidSTORMs was converted from analog/digital counts to the expected number of detected photons using the formula

$$N_{\text{photons}} = \frac{(A - \text{bias}) \cdot \text{sensitivity}}{\text{EM-gain}} \tag{3.2}$$

by the camera manufacturer[1]. The sensitivity depends on the specific imaging settings and was taken from the performance specification sheet of the camera; the bias offset was determined from dark frames (recorded with closed shutter), and the EM-gain is simply set by the experimenter. The total photon output of each molecule was determined by summing up all localizations of the corresponding cluster. The histograms of the total photon yield could be well fitted with an exponential decay, which gave an average photon output of $3.8 \cdot 10^4$ and $3.5 \cdot 10^6$ for 294 K and 89 K respectively; an increase over two orders of magnitude.

The position of each molecule is calculated as the average over its cluster of localizations. The spread of the cluster gives an estimate of the precision of a single localization. A principal component analysis

---

[1] A motivation for this formula as well as an explanation of the inner workings of an EMCCD camera can be found in this publication by Hirsch *et al* [48].
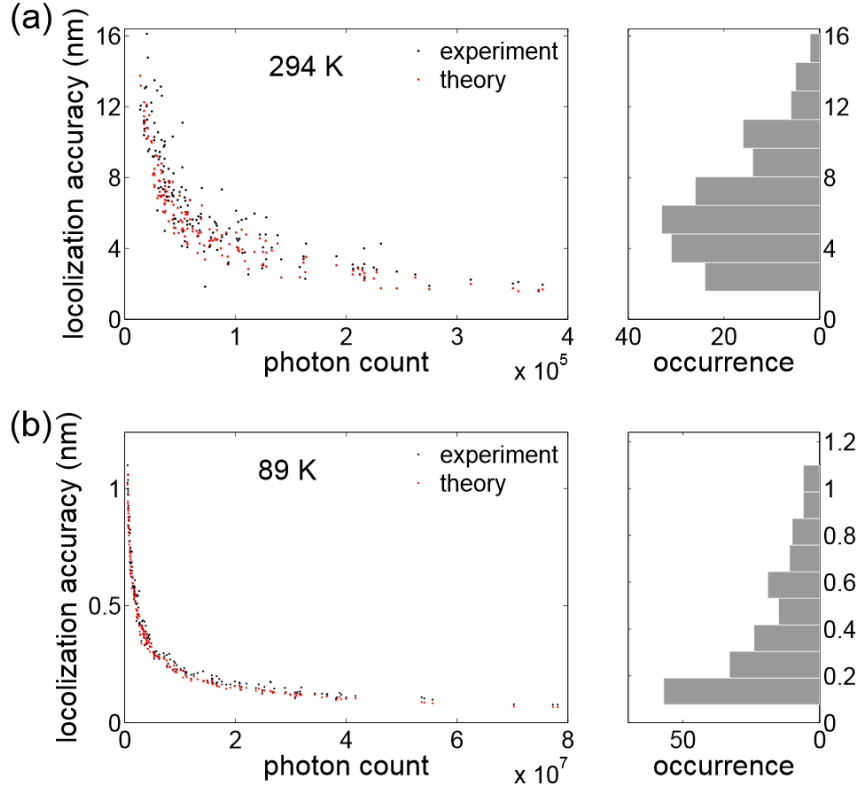
Figure 3.3: Experimental and theoretical localization precision of Atto647N at (a) 294 K and (b) 89 K. The theoretical prediction was calculated for each molecule individually to capture the influence of the local background. The localization precision at cryogenic temperature is about one order of magnitude better, allowing a precision of below 1 nm for over 80% of molecules. [This image has been published in the article [75].]

of each molecules cluster was performed to calculate the standard deviations σ_1, σ_2 along the two orthogonal directions of largest extent of the cluster. The localization error of the mean position is then calculated as

$$\sigma = \frac{1}{\sqrt{N_{\mathrm{loc}}}} \cdot \frac{\sigma_1 + \sigma_2}{2}. \tag{3.3}$$

This measurement of the localization precision is compared to the theoretical prediction by Mortensen *et al* using the formula [26]

$$\sigma_{\mathrm{theo}} = \sqrt{\frac{\sigma_a^2}{N_{\mathrm{photons}}} \left(\frac{16}{9} + \frac{8\pi\sigma_a^2 b}{N_{\mathrm{photons}}a^2} \right)} \tag{3.4}$$

where a^2 is the pixel area, $\sigma_a^2 = \sigma_{\mathrm{PSF}} + a^2/12$ the pixel corrected spatial standard deviation of the PSF and b the number of background photons[2]. The theoretical prediction was calculated for each molecule individually to capture the influence of the local background. Figure 3.3 shows a comparison of the localization precision for room- and cryogenic temperature. The theoretical and experimental values are in excellent agreement. Over 80% of molecules can be localized with a precision below 1 nm at 89 K compared to about 10 nm at room temperature, which fits the observed 100-fold increase in photon

---

[2]In the original publication of Mortensen, the number of background photons is b^2, not b. This was changed, as we think using b is less confusing.

yield, as the dominant scaling of the precision in Eq. (3.4) is $1/\sqrt{N_{\text{photons}}}$. This shows, that the photostability of fluorophores at cryogenic temperatures allows the localization of isolated single molecules with sub-nanometer precision.

To sum up this section about sub-nanometer localization at cryogenic temperatures

- Experiments on single Atto647N molecules were performed by *Weixing Li* at liquid nitrogen temperatures.

- The overall photon yield of molecules increases over two orders of magnitude due to increased photostability.

- An evaluation scheme was developed that corrects drift in the data and allows a high precision localization of single emitters.

- Over 80% of molecules can be localized with a precision below 1 nm.

3.2   Co-localization using polarized excitation & detection

While the accurate localization of single fluorophores demonstrates the advantages of the increased photon yield at cryogenic temperatures, the co-localization of close-by emitters is of much greater interest. Distance measurements with sub-nanometer precision would allow the determination of molecular structures and further facilitate the combination of super-resolution fluorescence microscopy with cryo-electron microscopy. Nanometer-distance measurements based on fluorescence microscopy were first demonstrated by Needham *et al* in 2013 [78, 79]. Their approach is based on the identification of step wise bleaching or blinking of a construct labeled with two fluorophores. In a state where only one of the fluorophores is emissive, its position and intensity can be extracted using SMLM methods. A global two-emitter fit to each spot uses the information from the different fluorescent states to extract the distance. They verified a resolution of about 10 nm resolution using double helix DNA molecules labeled with Atto 647N and measured the separations between inactive epidermal growth factor receptor (EGFR) molecules down to 8 nm. However, there are some drawbacks to the method: The identification of fluorescent states is not fully automatic, but requires a lot of manual work to achieve the desired precision [79]. A major problem is the low yield: From 40.000 traces in the EGFR experiment, only 193 (0.5%) are selected for the distance measurement. Thus, the method is not very robust and a large effort is required for each distance measurement. A similar study based on step-like blinking at cryogenic temperatures has been published by Weisenburger *et al* in 2014 using a different algorithm for the analysis [80]. When the number of emissive fluorophores changes, the center of mass of the intensity distribution is slightly shifted, which allows the calculation of the distance of the two fluorophores. Again, the major drawback is the low yield: only 0.1% of the recorded constructs (12 out of 12000) could be used for co-localization.

These results and our own experience show that it is very difficult to tune fluorophore behavior to allow the reliable identification of the different states and extraction of the distance from the blinking,

even more so at cryogenic temperatures. For a wide and easy applicability, a more robust method that requires less manual intervention and gives a higher yield is desirable. To achieve this, we looked at methods to separate two fluorophores without the requirement of photoblinking. First, we tried measurements with two spectrally different fluorophores which can be imaged individually by different filters in the emission pathway. Super-resolution with two colors requires highly accurate characterization and correction of the chromatic aberrations of the imaging system, but has been reported to achieve nanometer resolution at room temperature in the best case [81]. However, when imaging multi-color fluorescent beads, we observed systematic but non-parameterizable offsets between the colors [74, Section 10.2], which could not be removed by applying a linear correction (translation, rotation, scaling). As the objective is located outside the cryo-chamber to prevent aberrations due to different thermal expansion of its components [71], the sample is imaged through a thick optical window. This seems to introduce additional and irregular chromatic aberration, which made precise co-localization impossible.

To co-localize two molecules, the utilized method must be able to accurately attribute the measured photons to the individual fluorophores. One physical property that can be used to achieve this is the direction of polarization of the emitted light, which is in general not identical between two different fluorophores. If SMLM conditions can be achieved, the localizations of a two-emitter-complex could in principle be separated by measuring their polarization. As photoblinking is still hard to tune at cryogenic temperature, we thought of a different concept: Here we will illustrate an idea how to co-localize two non-blinking fluorophores by using a setup with polarized emission and detection. Although a microscope for these measurements was constructed and calibrated [74], we unfortunately never got a sample that was reliably labeled with two fluorophores after sample preparation for the cryo-microscope. For this reason, we only briefly describe the method and demonstrate its performance on simulations.

Imagine imaging two very close dipole emitters with a $90°$ angle between their orientations. Dipoles are only excited by the electric field component parallel to their orientation, so for a given dipole with brightness a and angle Θ excited with a linearly polarized laser beam in an angle Φ (to the x-axis) the measured intensity is

$$I(\Phi) = a \cdot \cos^2(\Phi - \Theta),$$

which is known as *Malus' law*. For an angle of $90°$ between laser and dipole orientation, no excitation is possible and the emitter stays dark on your camera. If the dipole axis are orthogonal, one fluorophore is optimally excited when the other one is completely dark. By changing the polarization of the excitation, the two fluorophores can be individually imaged and localized (Fig.3.4).

The distance can simply be computed as the difference of their positions. Here, we assumed the absolute molecule orientations are known. In this example, the orientations are simple to measure, as the maximum of the $I(\Phi)$ curve gives the orientation of the brighter emitter and the minimum the orientation of the dimmer one ($90°$ offset) [3].

---

[3]If both emitters are equally bright the intensity is constant, independent of the excitation angle ($\cos^2(x) + \sin^2(x) = 1$) and their orientation can not be recovered
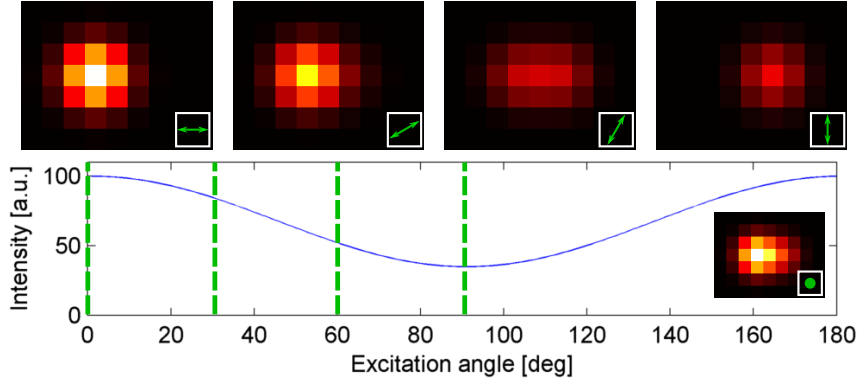
Figure 3.4: Measuring two dipole emitters with orthogonal orientations using various excitation polarization angles. The emitter orientations are axis aligned with the brighter one parallel to the x-axis ($0°$). The bottom curve shows the intensity obtained by spatial integration of the recorded frames. The distance shown is $d = 2$ pixel with $\sigma_{\text{PSF}} = 1$ pixel (FWHM ≈ 2.3548 pixel). Bottom right shows the unpolarized image.

The center of mass of the intensity distribution in Figure 3.4 wanders between the positions of both emitters as the excitation angle changes. For unknown brightness values a_1, a_2 and orientations Θ_1, Θ_2 the center of mass of the two-molecule image at a certain laser excitation angle Φ is given by

$$\mathbf{r}_{\text{CM}}(\Phi) = \frac{\mathbf{r}_1 \cdot I_1(\Phi|a_1, \Theta_1) + \mathbf{r}_2 \cdot I_2(\Phi|a_2, \Theta_2)}{I_1(\Phi|a_1, \Theta_1) + I_2(\Phi|a_2, \Theta_2)}, \tag{3.5}$$

where we assume a radially symmetric PSF for the emitters at positions $\mathbf{r}_1$ and $\mathbf{r}_2$. If the individual emitter intensities are known at two excitation angles, the corresponding system of equations can be solved for the two unknown positions, yielding their distance. Thus, the emitter orientations Θ_1, Θ_2 and brightnesses a_1, a_2 have to be recovered first.

The intensity of the two molecules at a fixed excitation angle is given by:

$$I(\Phi) = a_1 \cdot \cos^2(\Phi - \Theta_1) + a_2 \cdot \cos^2(\Phi - \Theta_2).$$

This can be rewritten in the form:

$$
\begin{aligned}
I(\Phi) &= a_1 \cdot \cos^2(\Phi - \Theta_1) + a_2 \cdot \cos^2(\Phi - \Theta_2), \quad \text{using } \cos^2(x) = 0.5 \cdot (1 + \cos(2x)) \\
&= \frac{a_1 + a_2}{2} + \frac{a_1}{2} \cos(2\Phi - 2\Theta_1) + \frac{a_2}{2} \cos(2\Phi - 2\Theta_2) \\
&= B + A \cdot \cos(2\Phi - \vartheta), \text{ with}
\end{aligned}
$$

$$A = \sqrt{\frac{a_1^2 + a_2^2}{4} + \frac{a_1 a_2}{2} \cos(2(\Theta_1 - \Theta_2))}, \ \tan\vartheta = \frac{a_1 \sin(2\Theta_1) + a_2 \sin(2\Theta_2)}{a_1 \cos(2\Theta_1) + a_2 \cos(2\Theta_2)}, \ B = \frac{a_1 + a_2}{2},$$

where the last identity can be proven by expressing the cosine using complex exponentials. From this equation it is apparent that a set of three parameters A, B, ϑ defines a unique curve $I(\Phi)$, which means that we can not acquire unique solutions for four unknowns $a_1, a_2, \Theta_1, \Theta_2$ from the measurement[4].

---

[4]For every parameter set, one indistinguishable solution yielding the same intensity-angle curve can be constructed using emitters with orthogonal polarization, setting one emitters brightness to the minimum of the curve ($a_2 = B - A$), the other one to the maximum ($a_1 = B + A$) and adjusting the phase to match the original curve $\vartheta = 2\Theta_1$
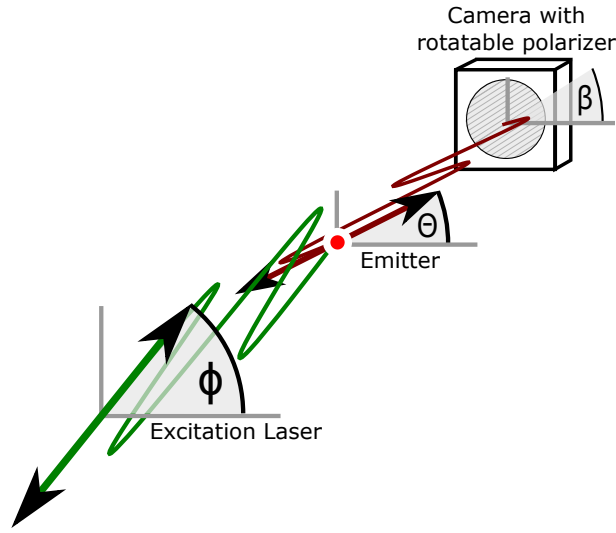
Figure 3.5: Measurement setup using polarized excitation and polarized detection via a rotateable polarizer.

To recover all parameters, the measurement needs to capture more information about the sample, which can be achieved by an additional polarizer in the emission pathway. Figure 3.5 shows an example of a setup with a rotating polarizer at angle β. For simplicity we choose $\beta = \Phi$ here, so that the polarization of the detection matches the excitation. If we excite/detect an emitter with $45°$ angle to its orientation, its signal will be suppressed stronger than before, as it is now excited **and** detected with lower efficiency. Thus it is somewhat intuitive that we should be able to separate emitters of different orientations better than before. The measured intensity is
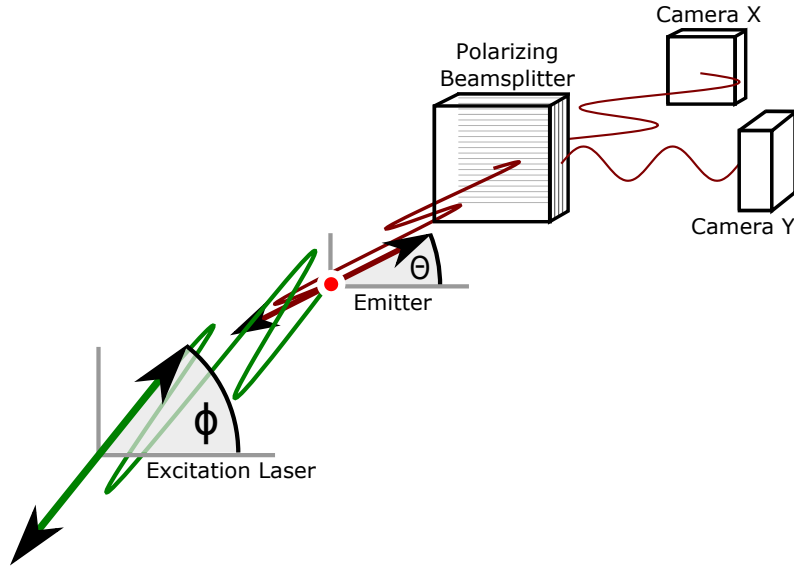
$$I(\Phi, \beta = \Phi) = a_1 \cdot \cos^2(\Phi - \Theta_1) \cdot \cos^2(\beta - \Theta_1) + a_2 \cdot \cos^2(\Phi - \Theta_2) \cdot \cos^2(\beta - \Theta_2)$$
$$= a_1 \cdot \cos^4(\Phi - \Theta_1) + a_2 \cdot \cos^4(\Phi - \Theta_2)$$
$$= \frac{3}{8}(a_1 + a_2) + \frac{a_1}{2}\cos(2\Phi - 2\Theta_1) + \frac{a_2}{2}\cos(2\Phi - 2\Theta_2)$$
$$+ \frac{a_1}{8}\cos(4\Phi - 4\Theta_1) + \frac{a_2}{8}\cos(4\Phi - 4\Theta_2)$$
$$= B + A_1\cos(2\Phi - \vartheta_1) + A_2\cos(4\Phi - \vartheta_2),$$

which is uniquely defined by five parameters. Therefore we can determine $a_1, a_2, \Theta_1, \Theta_2$ from the measured intensity curve.

A more practical setup uses a polarizing beam-splitter and two cameras to record the images of two orthogonal polarizations at the same time (see Fig. 3.6). This saves recording time and, more importantly, ensures the mechanical stability of the detection pathway, as no movable parts are present. Each excitation angle gives two independent measurements, uniquely defined by six parameters:

$$I_x(\Phi) = a_1 \cdot \cos^2(\Phi - \Theta_1)\cos^2(\Theta_1) + a_2 \cdot \cos^2(\Phi - \Theta_2)\cos^2(\Theta_2) \tag{3.6}$$
$$= B_x + A_x\cos(2\Phi - \vartheta_x),$$
$$I_y(\Phi) = a_1 \cdot \cos^2(\Phi - \Theta_1)\sin^2(\Theta_1) + a_2 \cdot \cos^2(\Phi - \Theta_2)\sin^2(\Theta_2) \tag{3.7}$$
$$= B_y + A_y\cos(2\Phi - \vartheta_y).$$

Figure 3.6: Measurement setup with polarized excitation and polarized detection using a polarizing beam-splitter and simultaneous recording on two cameras. This type of setup was built by *Weixing Li* in our lab.

Again, this is enough to uniquely determine the emitter brightness and orientations. This type of setup was built and calibrated by *Weixing Li* in our lab [74] and is thus used for further analysis. The parameters $a_1, a_2, \Theta_1, \Theta_2$ can be recovered by a global fit to the measured intensities using Equations (3.6) & (3.7). As each model is based on three parameters, at least three excitation angles have to be measured. To increase robustness, we measure six angles that are uniformly distributed in $[0, \pi)$. Note that the method can distinguish constructs labeled with one, two and three emitters, which allows an estimation of the labeling efficiency or stoichiometric measurements. Because only six parameters determine the curve, structures with more labels can not be identified.

Once the parameters are known, the distance between the emitters can be recovered by measuring the movement of the center of mass of the intensity distribution for different angles. Although this evaluation could be done for each of the two cameras individually, the visibility of individual constructs can be low for specific excitation angles. Simulations suggest that better results can be achieved if the beam-splitter is removed and all photons are concentrated on one camera. Again, the center of mass of the intensity distribution is given by

$$\mathbf{r}_{\mathrm{CM}}(\Phi) = \frac{\mathbf{r}_1 \cdot I_1(\Phi|a_1, \Theta_1) + \mathbf{r}_2 \cdot I_2(\Phi|a_2, \Theta_2)}{I_1(\Phi|a_1, \Theta_1) + I_2(\Phi|a_2, \Theta_2)} \tag{3.8}$$

$$= \begin{pmatrix} A_1(\Phi) & A_2(\Phi) \end{pmatrix} \cdot \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix}, \tag{3.9}$$

$$A_1(\Phi) := \frac{I_1(\Phi|a_1, \Theta_1)}{I_1(\Phi|a_1, \Theta_1) + I_2(\Phi|a_2, \Theta_2)}, A_2(\Phi) := \frac{I_2(\Phi|a_2, \Theta_2)}{I_1(\Phi|a_1, \Theta_1) + I_2(\Phi|a_2, \Theta_2)},$$

which requires at least two measurements of the center of mass to recover the emitter positions by inversion. Like before, more measurements increase the robustness of the estimation, yielding the overde-

termined system of equations

$$
\begin{pmatrix} \mathbf{r}_{CM}(\Phi_1) \\ \mathbf{r}_{CM}(\Phi_2) \\ \mathbf{r}_{CM}(\Phi_3) \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} A_1(\Phi_1) & A_2(\Phi_1) \\ A_1(\Phi_2) & A_2(\Phi_2) \\ A_1(\Phi_3) & A_2(\Phi_3) \\ & \vdots \end{pmatrix}}_{M} \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} ,
\tag{3.10}
$$

which is solved in a least-squares sense for $\mathbf{r}_1$ and $\mathbf{r}_2$ using the Moore–Penrose pseudoinverse of the matrix M. Note that the proposed method requires a fixed orientation of the fluorophores during the whole measurement. Fortunately, we never observed reorientation of single molecules at cryogenic temperatures in test measurements, which indicates a high stability of the dipole orientation at low temperatures.
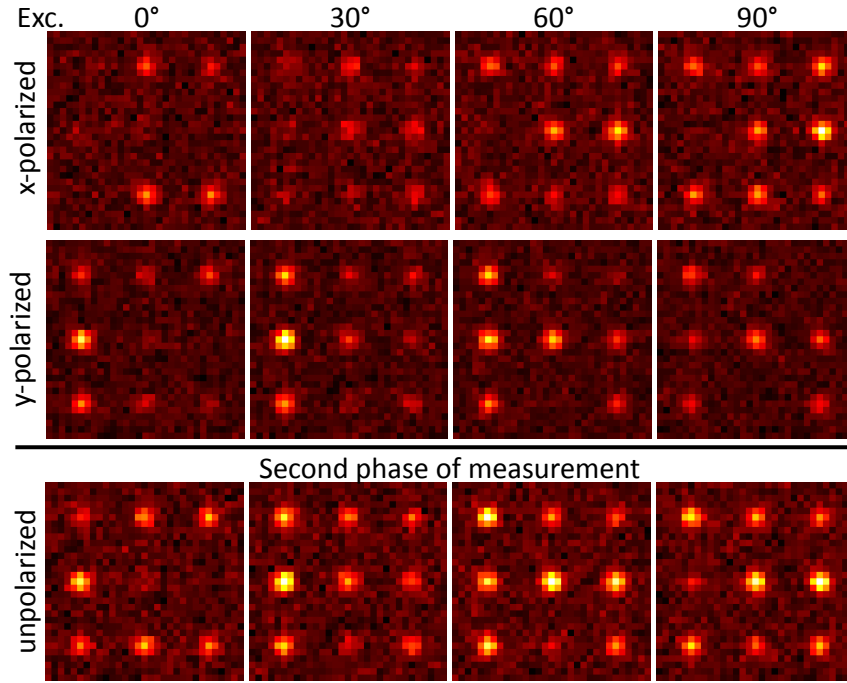
Figure 3.7: Example simulation of the data that must be recorded for the distance measurement. Emitter brightness and background were set to 3770 photons/emitter and 1200 photons/pixel per frame respectively (Fig. 3.7), based on single molecule experiments that were conducted with the setup. In **phase 1** the sample (here: emitters on a grid) is detected in two orthogonal polarization directions and recorded at different orientations of the linearly polarized excitation laser. In **phase 2** the sample is imaged at the same excitation angles on one camera without polarized detection.

In summary, the measurement is organized as follows: In **phase 1** the sample is successively excited with linearly polarized laser light with at least 3 different directions and simultaneously detected in two orthogonal polarizations with two cameras using a polarizing beam-splitter (Fig. 3.6). In **phase 2** the beam-splitter is removed and the measurement is repeated for the same excitation angles with unpolarized detection using one camera. Figure 3.7 shows the type of data that is acquired in the process.
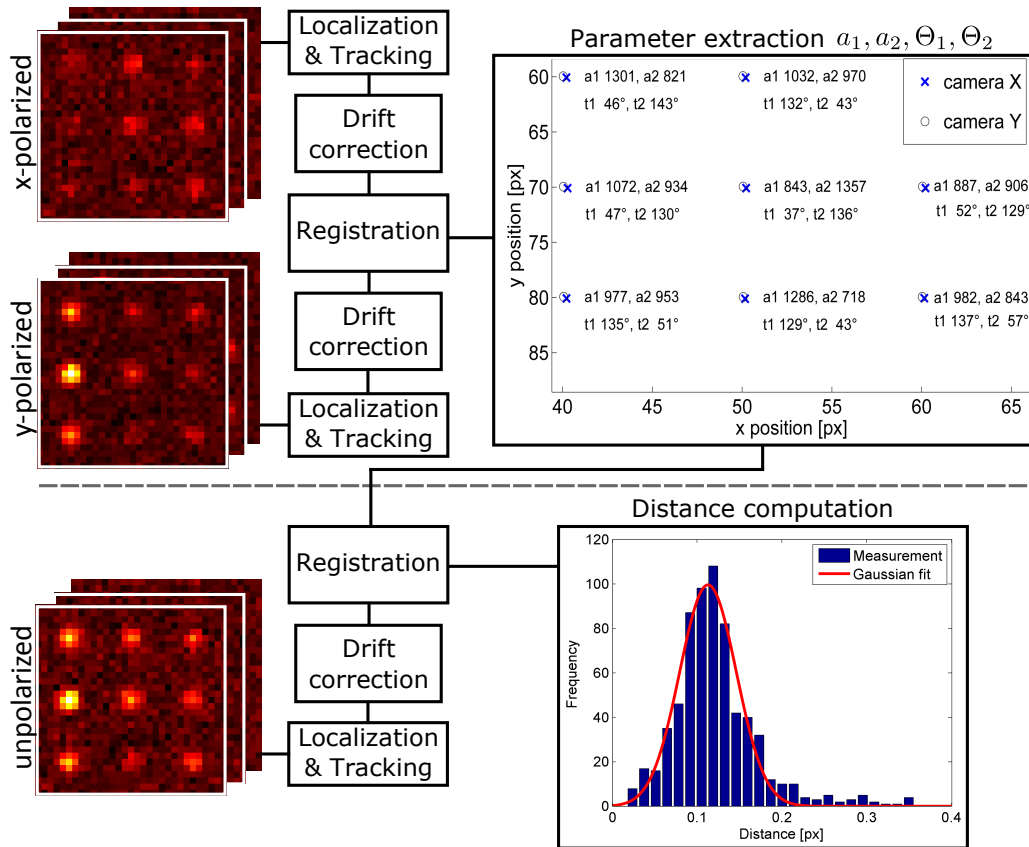
Figure 3.8: Scheme of distance evaluation: The two-fluorophore structures are first localized and tracked in all movies using TrackNTrace (Chap. 4), extracting their brightnesses and positions, followed by drift correction. The data from the two cameras in **phase 1** is registered to the same frame of reference using an algorithm for point-to-point registration [82]. Using the measured intensities, the individual emitter brightnesses and orientations are recovered by a global fit based on Equations (3.6) & (3.7). The recovered parameters are then registered to the unpolarized data from **phase 2**. The movement of the centroid in the unpolarized measurement together with the recovered parameters is used for computing the distance by Equation (3.10).

We developed a software suite for the evaluation, which implements the scheme depicted in Figure 3.8. The two-emitter constructs are localized and tracked, extracting their brightness for both polarization directions by Gaussian fits to the **phase 1** data using TrackNTrace (Chap. 4). Similar to Section 3.1, the drift is corrected from the tracking data using the average displacement of the particles. The data from both cameras is registered to the same frame of reference using a state-of-the-art algorithm published in 2013 by Cohen *et al* [82]. After the registration, tracks belonging to the same construct are identified by a nearest-neighbor search. Using all angles angles where the brightness could be successfully extracted for both cameras, the individual emitter brightnesses and orientations are recovered by a global fit based on Equations (3.6) & (3.7). If both emitters have similar angles ($\Delta\Theta < 10°$) or very different brightness estimates (more than factor 4 difference), the structure is not used for further evaluation. Localization and drift correction were performed on the **phase 2** data, giving the angle dependent center of mass of intensity of each structure. As removal of the beam-splitter can change the field of view, registration to the phase 1 data is performed, linking the recovered parameters to the correct structure. With these parameters, the distance is recovered for each construct using Eq. (3.10).

Simulations showed that assuming a single Gaussian intensity distribution for the two-emitter construct does not hurt the result, as long as the distance between the emitters is much smaller than the width of the PSF, which is the case of interest here. If larger distances, for example only a factor of two below the diffraction limit, are considered, a different model for the extraction of the brightnesses and center of mass of intensity must be used.
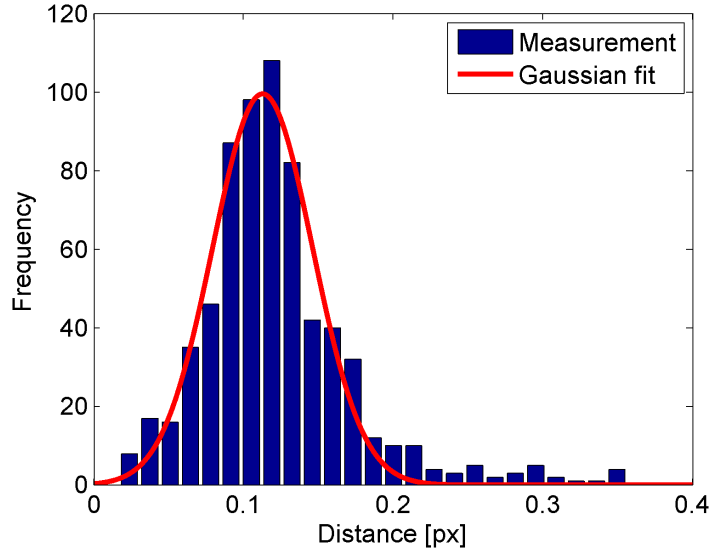
Figure 3.9: Histogram of measured distances from simulation of 900 randomly oriented emitters. Emitter brightness and background were set to 3770 photons/emitter and 1200 photons/pixel per frame respectively (Fig. 3.7). For both phases, six excitation angles $[0°, 30°, 60°, 90°, 120°, 150°]$ were simulated with 15 images per angle using a Poisson noise model, giving 180 images overall. The standard deviation of the PSF is $\sigma_{\mathrm{PSF}} = 1$ pixel, representing a well adjusted microscope setup. The recovered distance is $\bar{d} = 0.1127$ pixels compared to a true value of $d = 0.1154$ pixels.

To test the performance of the method, a simulation of 900 two-emitter constructs with random orientations of the fluorophores was performed. Emitter brightness and background were set to 3770 photons/emitter and 1200 photons/pixel per frame respectively (Fig. 3.7), based on single molecule test experiments that were conducted with the setup. For both phases, six excitation angles (uniformly distributed in $[0, \pi)$) were simulated with 15 images per angle using a Poisson noise model, giving 180 images overall. The standard deviation of the PSF is $\sigma_{\mathrm{PSF}} = 1$ pixel, representing a well adjusted microscope setup. The distance between the emitters was set to $d = 0.1154$ pixel. Figure 3.9 shows the recovered distance distribution[5]. Of the 900 structures, 680 ($\approx 76\%$) were used for the distance estimation. A Gaussian fit to the distance distribution yielded an estimate of $\bar{d} = 0.1127$ (2.4% difference to the true value) for the distance with an error of $\sigma_{\bar{d}} = 0.0334$ pixel for a single distance measurement[6]. For a well adjusted microscope with 100 nm pixel size, this translates to a recovered distance of about 11 nm with 3.4 nm error for a single measurement. This proves that the proposed method is theoretically able to achieve the desired performance in resolving two fluorophores with distances of only a few nanometers.

---

[5]Although not necessary for the simulation, a full evaluation including drift correction and registration was performed to keep results closer to the experimental situation.

[6]Corresponding to the width of the Gaussian.

We expect that it is easier to achieve stable behavior of the fluorophores at cryogenic temperatures, requiring less effort for sample preparation and allowing a higher percentage of samples to be used for the distance estimation than the $< 1\%$ yield common for blinking or bleaching based methods [78–80]. In the latter methods, precision and yield also depend on the occupation time of single fluorophores in specific fluorescent states which have to be identified first, while our method could potentially utilize the whole measurement time, using all available photons. The measurement could also be performed at room temperature if a fixed orientation of fluorophores can be guaranteed, for example by strong chemical fixation. The major remaining test is of course the experimental validation, which could not be performed due to the missing samples. We expect some additional challenges for the actual measurement: We observed polarization-dependent transmission of the excitation and emission light through some optical components in our setup. Polarization dependent background was also observed, which is partly due to fluorescent impurities in the objective [74]. The phase between s- and p- polarized components is also altered, introducing a very slight ellipticity to the linearly polarized beams. If the influence of these effects on the measurement is strong, they can be characterized and included in the model. Furthermore, if not all molecules are positioned in the same plane, or the microscope is a bit defocused, the PSF becomes asymmetric, introducing errors in the estimation of brightness and position of emitters via a simple two-dimensional Gaussian [83–85]. Although it has been shown that a more complex fitting model can compensate this when the axial range is restricted by using TIRF illumination [26], this is impossible to achieve with our low-NA objective. In the general case, where the z-position is largely undetermined, correcting this issue becomes harder [85], which poses a potential problem for our method. However, we performed calculations indicating that in a polarized detection scheme, the position of a dipole along a line orthogonal to the direction of the polarizer is virtually free of bias, which alleviates the problem. Surprisingly, the issue of orientation-dependent localization bias was never mentioned in previous co-localization approaches [78–80], including a very recent study which reports angstrom resolution of 3D protein structure data [86].

To sum up this section about polarization-based co-localization at cryogenic temperatures,

- Existing methods using fluorescence microscopy for nanometer distance measurements rely on blinking or bleaching, require time-consuming evaluations, and are only able to use less than 1% of the sample.

- A method is proposed how to distinguish constantly emitting fluorophores using linearly polarized excitation combined with polarized detection in two orthogonal directions.

- Simulations at conditions similar to our single molecule experiments show that distances far below the diffraction limit can be measured reliably, making the method promising for nanometer-distance measurements in the future.

- Although the necessary experimental setup was built and calibrated [74], the experimental verification is still pending due to missing samples that are reliably labeled with two fluorophores after preparation for cryogenic measurements.

4

TrackNTrace

Far-field fluorescence imaging of single molecules at room temperature was first achieved in the mid 1990's using Total Internal Fluorescence (TIRF) microscopy [87], shortly after it was demonstrated with near-field optical microscopy [16]. Now, only about 20 years later, single-molecule measurements are performed with relative ease on a daily basis in many research labs. This is mainly attributed to the steadily increasing sensitivity of cameras, which nowadays has reached a level that enables the detection of single photons. Single-molecule techniques avoid ensemble averaging, allowing the investigation of distributions of molecular properties and the study of their fluctuations. Since their advent, these methods have made important contributions to many fields, especially in biophysics [88]. In recent years, techniques such as Single-Molecule Localization Microscopy (SMLM) [10, 89], which reveals structures below the diffraction limit by localizing fluorophores that are spatio-temporally separated via carefully tuned photoblinking, and Single Particle Tracking (SPT) [90], where the acquired positions are subsequently linked in time to study the trajectories of single molecules, are becoming increasingly popular.

While a plethora of high-performance software suites for these techniques has been developed [91, 92], most of these programs are difficult to extend for specific needs and suffer from very restricted visualization capabilities, often limited to a display of the final result. This complicates parameter tuning or the identification of unsuitable settings and obfuscates processing errors. These issues are especially relevant when developing and testing new algorithms.

In labs which focus on method development, ideas for new kinds of measurements come up regularly, often requiring not only specialized experimental setups, but customized data evaluation as well.

For the work presented in this thesis, single-particle imaging was used in various ways: analyzing the blinking behavior of special fluorophores (Sec. 2.8); determining the photon yield and photostability of particles at different temperatures and localizing them with sub-nanometer precision (Sec. 3.1 & 3.2); recovering the orientation of fixed dipoles in defocused images (Sec. 5.2); and studying the intensity distribution of fluorophores in nanochannels (Sec. 5.2). As a product of trying to simplify these evaluations, their continuous improvement, and the design of entirely new methods, my colleague Jan Thiart and I developed the software TrackNTrace.

TrackNTrace (TNT) is the first framework specifically designed to facilitate the development of single-molecule localization and tracking applications. All processing stages – detecting single particles or patterns, refining their positions and extracting parameters, and subsequent tracking – are implemented as user-definable plugins. Existing algorithms can be integrated with minimal effort, including automatic incorporation into the graphical user interface (GUI). Each stage's output can be previewed and inspected in detail using a novel visualization interface, enabling the user to quickly notice oddities and trace the effect of changes in parameters or utilized algorithms, which is not possible with existing programs. TrackNTrace is implemented in MATLAB, runs on Linux and Windows and is freely distributed as open-source software via its GitHub repository https://github.com/scstein/TrackNTrace.

This chapter gives a brief overview of the TrackNTrace framework, the ideas behind it and its capabilities. Furthermore, a performance comparison to established SMLM programs is presented along some examples of experimental applications. I focused on the design of the core framework as well as plugins for cross-correlation candidate detection, nearest-neighbor tracking and processing of defocused images. Most plugins were implemented by Jan Thiart, who also performed the experimental and performance evaluation shown here. Together, we also developed the (to our knowledge) first freely available[1] high performance C++ software for 2D Gaussian fitting that offers a MATLAB interface, which became the default plugin for sub-pixel localization in TrackNTrace.

Parts of this section have been published in the article [93]:

Stein, S.C.; Thiart, J.; "TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms.", Scientific Reports, 6, 37937, 2016.

Additionally, some parts are available in similar form in the TrackNTrace manual. For anyone interested in using the software, the manual is a good place to start from, and reading it is definitely recommended.

---

[1]Available independent from TrackNTrace as "Fast Gaussian Point Spread Function Fitting (MEX)" on the *MATLAB File Exchange.* URL: http://www.mathworks.com/matlabcentral/fileexchange/52417-fast-gaussian-point-spread-function-fitting--mex-.

4.1 Concept

Instead of concentrating on a single task, TrackNTrace was designed with the intent to enhance the typical workflow of data evaluation in research labs. Ideas for new measurement methods or improvements of existing ones arise on a regular basis, usually requiring new and adapted algorithms to process the data. Furthermore, especially in the field of localization microscopy, it is often cumbersome to compare existing methods and optimize their parameter settings for the situation at hand. Rich and fast feedback is very important for proper parameter adjustment, even more so when developing new algorithms. Consequently, TrackNTrace focuses on the following design philosophies:

- **Extensibility:** New and existing methods should be easy to integrate. The system is build around user-definable plugins to achieve this. Plugins can be shared easily by copying the corresponding file to the `plugin` sub-folder.

- **Simple coding:** To place the emphasis on the scientific questions, creating and changing plugins should be easy. TrackNTrace is written in the MATLAB programming language, which is common in the life sciences and comparably easy to learn. The plugin interface is kept as simple as possible. Plugins are automatically integrated into the main GUI and parallelized for simultaneous execution on multi-core processors to improve performance.

- **Quick & rich feedback:** Fast feedback on the quality of results and their quantitative aspects should be available. TrackNTrace implements a "preview" feature and a powerful visualizer. The preview computes the output for a defined interval of frames of the input data using the current settings and displays them in the TNT visualizer. The visualizer shows the outcomes of all steps of the data processing, including the output of every involved plugin (e.g. fit parameters for each molecule).

TrackNTrace was designed to support a wide range of fluorescence imaging applications. Although the currently available plugins mostly focus on single-molecule localization and tracking, every task that requires the detection of arbitrary intensity patterns is straightforward to integrate. For example, the system could be extended for cell or microtubule-tip tracking. The TrackNTrace framework subdivides the data analysis into four distinct stages which we identified as common in different applications:

1. **Correcting raw data:** Measured movies are read and corrected for dark currents and camera artifacts. Analog-digital-converter counts can be converted to photons if a quantitative signal analysis (e.g., maximum-likelihood estimation) is needed.

2. **Detecting candidates:** Potential sources of signal above the background noise are identified in each frame to obtain rough estimates of emitter positions. This stage usually relies on fast, approximate algorithms.

3. **Refinement:** A more demanding evaluation is performed for every candidate identified in the previous stage. Commonly, this involves fitting a representation of the microscope's PSF to a

subsection of the frame, extracting arbitrary parameters. In SMLM this is used to refine each candidate's position estimate to higher accuracy.

4. **Tracking:** Positions separated in time are connected frame-by-frame to form trajectories. High particle density, intersecting tracks, and re-appearing emitters are the main obstacles to overcome during this stage.

While stage 1 is performed by the TrackNTrace core program, stage 2-4 are executed by the user-definable plugins. The different steps are reflected in the structure of the TNT main GUI, as illustrated in Figure 4.1.
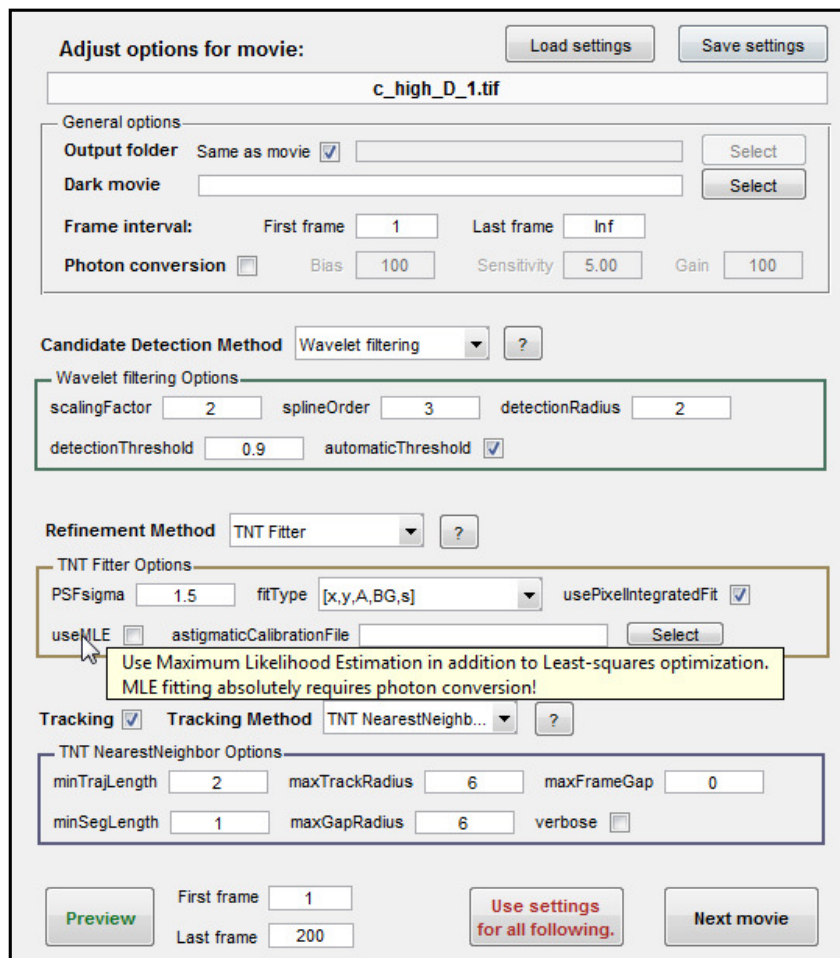

Figure 4.1: Main GUI of TrackNTrace. The settings for the four processing stages are set in the corresponding panels. A graphical representation of user plugins is created automatically and integrated into the GUI. [This image has been published in [93, Supplementary Information].]

As mentioned above, graphical representations for all user plugins are automatically created and integrated into the GUI. An explanation of each plugin can be displayed by pressing the "?" button. For additional help, all control elements (even inside user plugins) provide a mandatory tooltip, that is available by hovering over it with the mouse. This makes it easy to get information about the influence of individual parameters.

The typical workflow with TrackNTrace comprises the following steps: Multiple movies are selected

for data evaluation. Suitable plugins and their settings or chosen for the individual movies. This is done by choosing a certain parameter set and checking the results with the preview function, which invokes the visualizer (Figure 4.2). Here, the output of all plugins can be checked for qualitative or quantitative errors. Once the visualizer is closed, new settings can be chosen and re-checked if necessary. After optimizing the parameters for all movies, they are processed without the need for additional user intervention. Alternatively, the same parameter set can be used for all movies. When the computation is finished, the individual results are collected in a single file for convenient post-processing or visualization.
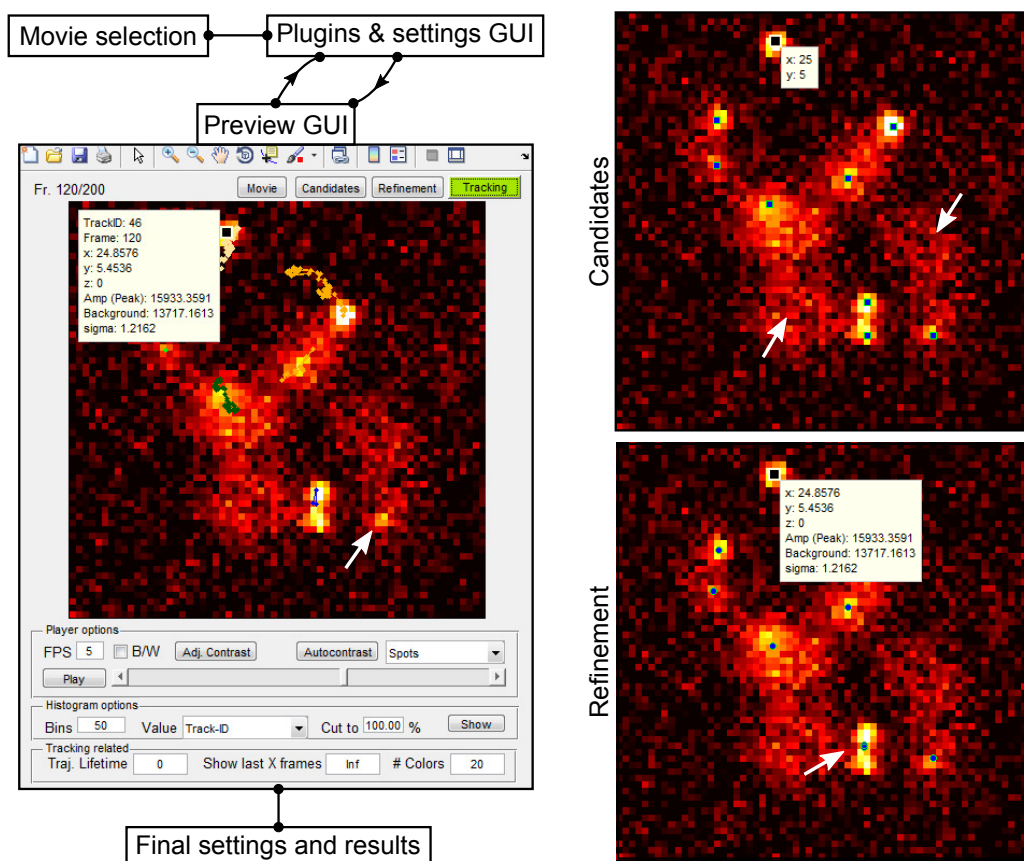
Figure 4.2: The visualizer of TrackNTrace displays the output of the plugins for all processing stages. Selecting a detected molecule reveals all information that is computed by the corresponding plugin. White arrows indicate typical issues such as undetected candidates, badly refined positions or prematurely ending tracks, which can be identified and resolved by choosing different settings. [This image has been published in [93].]

Figure 4.2 shows the TrackNTrace visualizer. Selecting a detected molecule with the mouse reveals all information computed by the corresponding plugin. White arrows indicate typical issues such as undetected candidates, badly refined positions or prematurely ending tracks, which can be easily identified with the visualizer and corrected by choosing different settings. Furthermore, it is possible to display and check the distributions of all computed parameters, which is often helpful. The visualizer code is heavily optimized and relies on intelligent caching as well as low-level plotting functions to allow the display of thousands of single molecule tracks at typical video frame rates[2].

---

[2]Performance is lower for MATLAB versions above 2014b due to the new MATLAB graphics engine.
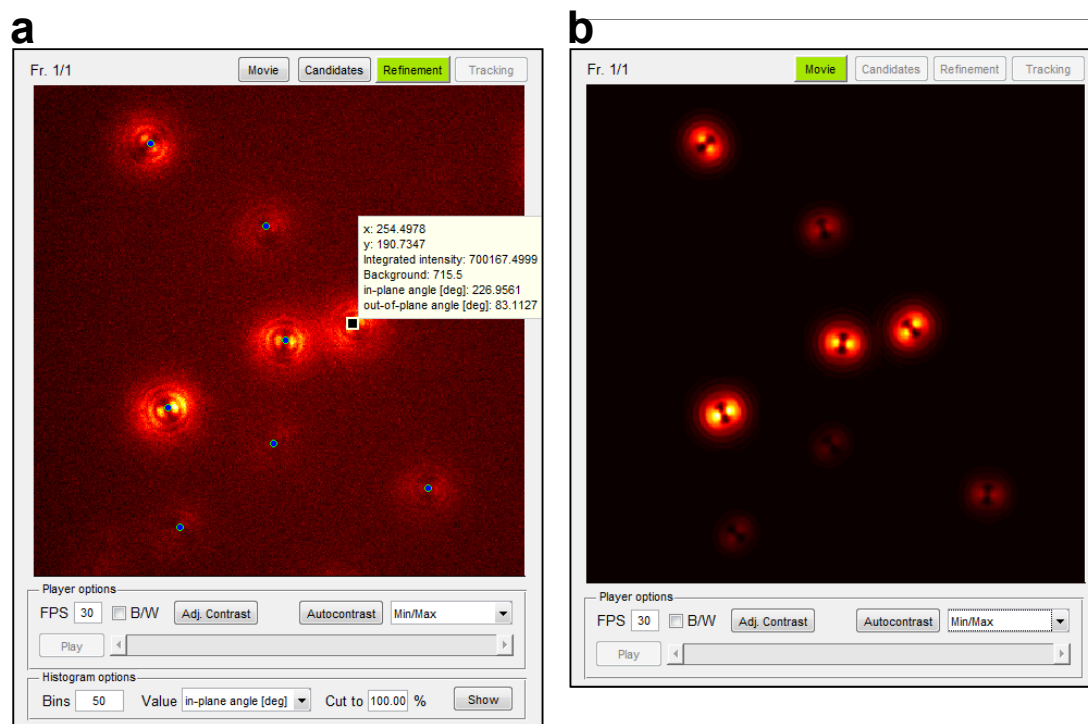
Figure 4.3: Preview output of the `Defocused refinement` plugin. (a) shows the original data (Atto655 molecules spincoated on a glass coverslip) with the detected candidates and calculated parameters and (b) a synthetic image visualizing the final results. Data recorded by *Narain Karedla*. [This image has been published in [93, Supplementary information].]

Strong visualization is especially helpful when using or developing applications that go beyond typical single-molecule localization experiments. Figure 4.3 shows one such task, the extraction of the positions and 3D orientations of single molecules, which is possible with TrackNTrace via the `Defocused patterns` and `Defocused refinement` plugins[3]. With the visualizer, it is easy to check the extracted parameters, like the in-plane and out-of-plane angles. To our knowledge, the evaluation of defocused images is currently not possible with any other SMLM software. With TrackNTrace, even subsequent tracking of the patterns is easily to achieve.

Currently, the following plugins are part of the TrackNTrace framework:

Candidate detection plugins

Name	Description
Cross-correlation	Fast normalized cross-correlation [94] with a theoretical Gaussian PSF.
Defocused patterns	Fast normalized cross-correlation [94] with a set of wave-optical images of defocused dipole emitters for varying azimuthal and polar angles [95, 96].
Image filtering	Various image filters (running mean/median, difference of averages, etc.) followed by local maximum extraction.
p-value filtering	Hypothesis-test-based filter (Gaussian vs. flat background).
Wavelet filter	Wavelet-based filter as described in [97].

---

[3]The plugins implement the methods described in Section 5.2.

<div align="center">**Refinement plugins**</div>

Name	Description
Defocused refinement	Sub-pixel refinement of `Defocused patterns` output. Based on least-squares minimization via a conjugate-gradient algorithm[56].
Gauss-2d-Circ	Gaussian fit by matrix inversion [98].
GPU-Gauss MLE	Gaussian fit implemented in C++ and CUDA (for graphics cards) [99].
Radial symmetry	Centroid calculation suitable for radially symmetric PSFs [100].
TNT fitter	MLE & least-squares Gaussian fit for pixel-integrated, asymmetric and rotated PSFs. Implemented with Google's *ceres* optimization library [101].
TNT *z*-Calibration	Calibration for astigmatic 3D imaging with the *TNT fitter* plugin.

<div align="center">**Tracking plugins**</div>

Name	Description
TNT NearestNeighbor	Fast C++ nearest-neighbor tracker using the *nanoflann* library [102].
u-Track	Powerful, high-density particle tracker using global optimization [103].

4.2 DESIGN OF THE PLUGIN SYSTEM

The plugin system of TrackNTrace is designed to minimize effort by the user and require no in-depth programming knowledge. To give an idea of how the system is used, we look at the (slightly simplified) *Radial symmetry* refinement plugin, which implements the "radial symmetry centers" algorithm by Parthasarathy [100] to extract sub-pixel positions from images of isolated emitters in the focal plane.

A plugin is defined as a function in a single file named "plugin_NAME.m" inside the `plugins` subfolder of the TrackNTrace root directory. Plugin functions return a `TNTplugin` object, which gathers all information needed for its execution. On startup, all "plugin_XX.m" files are loaded by the main program and integrated into the GUI. This is the content of "plugin_RadialSymmetry.m" file:

```matlab
function [plugin] = plugin_RadialSymmetry()
name = 'Radial symmetry';
type = 2;
mainFunc =  @refinePositions_radialSymmetry;
outParamDescription = {'x';'y';'z';'Amp (Peak)'; 'Background'; '
   width'};
% Create the plugin
plugin = TNTplugin(name, type, mainFunc, outParamDescription);
% Description of plugin
plugin.info = ['Localization by radial symmetry centers.', ...
               'Published in Parthasarathy, NatMet 2012(9).'];
```

```
% Add parameters
plugin.add_param('PSFSigma',...
    'float',...
    {1.3,0,inf},...
    'PSF standard deviation in [pixel].');
plugin.add_param('estimateWidth',...
    'bool',...
    true,...
    'Estimate particle width.');
end
```

Parameters that should be adjustable by the GUI are defined by calling the add_param(name, type, settings, tooltip) function of the newly created plugin instance. The definition above produces the graphical representation shown in Figure 4.4.
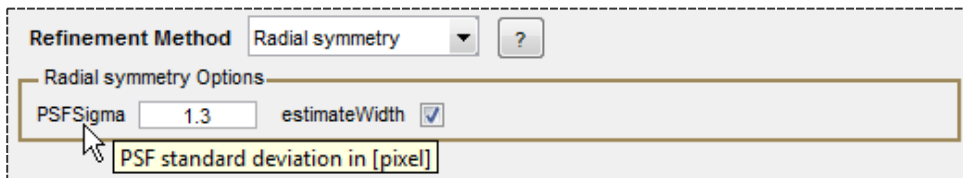
Figure 4.4: GUI of the *Radial symmetry* plugin as created from the code in the main text.

Each parameter has a name, a type, type-dependent settings and a tooltip. The tooltip should explain its meaning and is displayed when hovering over the parameter with the mouse (as shown in the figure). The parameter name is not only used for display in the GUI, but also to create a corresponding variable for use in the main function of the plugin. Parameters are gathered in an "options" struct and passed to the main function as an input parameter. For example, PSFSigma can later be used inside the function as options.PSFSigma. Different parameters need different settings as the third input. The complete list of current parameter types and settings is:

par_type	par_settings	Example
'float'	Double, {defaultValue, lowerBound, upperBound}	{1.3,0,inf}
'int'	Integer, {defaultValue, lowerBound, upperBound}	{4,-10,10}
'bool'	Boolean, defaultValue	true
'string'	String, 'defaultValue'	'@exp'
'list'	Interactive list box, {'defaultEntry', 'Entry2',...}	{'x','[x,y]'}
'filechooser'	Interactive file chooser dialog, {'defaultDir','fileEnding'}	{'C:/Sci/','csv'}

To better structure the GUI, users can call the newRow() function of a plugin instance between

`add_param(...)` calls, which starts a new row in the graphical representation of the plugin. Furthermore, headings or descriptions can be added with `add_text(text, horizontalAlignment, fontWeight, fontAngle, tooltip)`.

While the four parameters used in its constructor are mandatory for every plugin, there are additional parameters that can be set optionally after the object is constructed. Here is a complete list:

<div align="center">TNTplugin properties</div>

Property	Description
`name`	Plugin name as displayed by the GUI.
`type`	Type of plugin. 1: Candidate detection, 2: Refinement, 3: Tracking
`mainFunc`	Handle to main function the plugin implementing the actual algorithm. For type 1 and 2 this is called by TrackNTrace in a loop for each individual frame of the input movie. The interface (output/input) of this function depends on the plugin type and must match the TrackNTrace specification (see manual).
`outParamDescription`	Cell array of strings with description/name of all output parameters (columns) of the plugin
`info`	Description of the plugin itself. Should describe the method and the general way how to use it.
`initFunc`	Initialization function called before the main one is first executed.
`postFunc`	Post-processing is called after the last main function execution.
`useParallelProcessing`	Boolean. If false, TrackNTrace does not parallelize this plugins main function on a frame-by-frame basis for type 1/2 plugins. Useful if the main function itself is parallelized (e.g. a compiled multi-threaded *MEX* file) or if global information is needed (e.g. access to multiple frames of the movie).

Candidate detection and refinement plugins are automatically parallelized, speeding up data processing without requiring a single line of code by the user. If parallel processing is not possible, the framework falls back to single-threaded execution. This feature can be deliberately turned off by the user, which is useful if, for example, the plugin code itself is parallelized.

The main function of the plugin implements the actual algorithm. It is executed on each frame of the movie for candidate detection and refinement plugins, and called once to process the refinement output for tracking plugins. The functions must follow a type-dependent interface, which is specified in the TrackNTrace manual. Main functions of refinement plugins, such as our *Radial symmetry* example, receive one image of the movie and the corresponding detected single molecules from the candidate detection step at a time. The main function of the *Radial symmetry* plugin looks as follows:

```matlab
function [refinementData] = refinePositions_radialSymmetry(img,
    candidatePos,options,currentFrameNr)

% Parse input
halfWindowSize = ceil(3*options.PSFSigma);
refinementData = zeros(size(candidatePos,1),5+ ...
                            options.estimateWidth);

for iCand = 1:size(candidatePos,1)
    % Define subimage
    posXY = candidatePos(iCand,1:2);
    idxX = posXY(1)-halfWindowSize:posXY(1)+halfWindowSize;
    idxY = posXY(2)-halfWindowSize:posXY(2)+halfWindowSize;

    % Call Radial Symmetry function
    [xc,yc,sigma] = radialcenter(img(idxY,idxX), ...
                                    options.estimateWidth);

    % Transform relative coordinates back to image coordinates
    xc = xc+posXY(1)-(halfWindowSize+1);
    yc = yc+posXY(2)-(halfWindowSize+1);
    zc = 0;

    % Save data
    refinementData(iCand,:) = [xc,yc,zc, ...
                    img(round([yc,xc])), sigma];
end

end
```

For every particle detected in the first stage, a subimage of a few pixels around its position is created and passed to the function radialcenter which is available from the original publication by Parthasarathy [100]. After the local coordinates are converted to image coordinates, the output is saved in the refinementData matrix for further processing. GUI parameters are simply accessed by their name via the options struct (options.PSFSigma and options.estimateWidth), as mentioned before. The generated output can be used with arbitrary tracking plugins in the next processing stage.

This example shows how simple it is to integrate new or existing algorithms into the framework. For existing code, the necessary modifications are often incorporated within minutes.

4.3 PERFORMANCE EVALUATION AND EXPERIMENTAL EXAMPLES

The main advantage of TrackNTrace compared to existing software is its information-rich visualization, extensibility and simplicity of coding, rather than being optimized for a specific task. Nonetheless, we demonstrate its competitive performance by comparison with two state-of-the-art programs for single-molecule localization microscopy: *rapidSTORM* [77], an efficient an popular stand-alone software, and *ThunderSTORM* [104], an *ImageJ* plugin that showed the highest performance in a recent meta-study [91]. Furthermore, we present some typical experimental applications that can be conducted with the program out-of-the-box. The evaluation and experimental examples were performed by Jan Thiart.
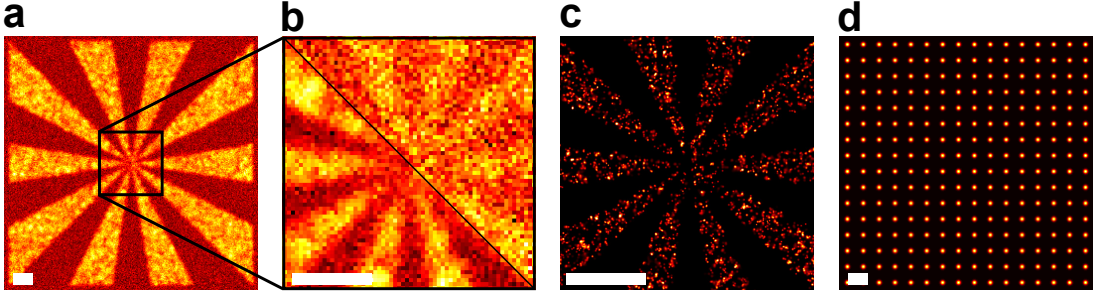
Figure 4.5: Simulations for the performance evaluation. (a) Average intensity of Siemens star pattern at SNR = 3. (b) Central region for an average SNR of 1 (upper right) and 3 (lower left). (c) SMLM reconstruction of zoomed-in region at SNR 3 using TrackNTrace with cross-correlation candidate detection. (d) Single frame of SNR = 10 simulation grid for execution speed evaluation. [This image has been published in [93, Supplementary information].]

To test the detection efficiency, accuracy and precision, we prepared simulations of emitters distributed on a Siemens star pattern at different signal-to-noise ratios under STORM conditions. Furthermore, a high-SNR movie of non-blinking emitters on a grid is prepared to compare the execution time of different programs. The high SNR ensures that all programs detect the same number of emitters. Figure 4.5 shows both scenarios.

As outlined in the publication [93, Methods]: Simulations were performed as previously described by Smith *et al* [105]. Briefly, a Siemens star with $n = 10$ arms was simulated on a $256 \times 256 \, \mathrm{px}^2$ grid with a pixel size of $a = 100 \, \mathrm{nm}$. The particle density was fixed to $\varrho = 7.5 \, \mathrm{px}^{-1}$ and dSTORM-like conditions were emulated by drawing active state durations from a Poisson distribution ($1/k_{\mathrm{active}} = 1/k_{\mathrm{deactivate}} + 1/k_{\mathrm{on}}$ and $k_{\mathrm{deactivate}} = k_{\mathrm{off}} + k_{\mathrm{bleach}}$) combined with photo-beaching. The time until bleaching was drawn from a geometric distribution with a bleaching probability $p_{\mathrm{bleach}} = k_{\mathrm{bleach}}/k_{\mathrm{deactivate}}$. The rates were chosen as $k_{\mathrm{off}} = 1 \, \mathrm{frame}^{-1}$, $k_{\mathrm{bleach}} = 0.15 \, \mathrm{frame}^{-1}$, and $k_{\mathrm{on}} = k_{\mathrm{off}}/(5\varrho) \, \mathrm{px}^{-1}$. The procedure yielded a list of positions (μ_x, μ_y) and active state durations per frame $A \in [0, 1]$. The simulation uses an integrated Gaussian Point Spread Function model,

$$g = \int_{-0.5}^{+0.5} \mathrm{d}x \int_{-0.5}^{+0.5} \mathrm{d}y \, \mathcal{N}\left(\mu_x, \mu_y, \sigma_{\mathrm{PSF}}\right), \tag{4.1}$$

with a size of $\sigma_{\mathrm{PSF}} = \lambda/(4\sqrt{\log 2} \, \mathrm{NA})$. Here, $\lambda = 670 \, \mathrm{nm}$ and $\mathrm{NA} = 1.4$ were chosen. For all active emitters, $A \times Ng$ was added to the image where we assumed a fixed photon yield of $N = 50$.

An additional background of 10 photons was added. The end result was distorted by Poisson noise and normalized to average signal-to-noise ratios of 1 to 5. Each movie had a length of 1500 frames. For execution time measurements, 3500 frames were used. The execution time is corrected for reading the movies into memory.

For quality assessment, we calculated the Jaccard index (JAC), the root-mean-square distance of localizations $\mathcal{L}_{\text{obs}}$ from the ground truth $\mathcal{L}_{\text{obs}}$ (RMSE) and the Fourier ring correlation (FRC) between the data and the ground truth as described by Nieuwenhuizen *et al* [106]. The Jaccard index is simply the number of correct detections (true positives, TP) divided by the all detections (TP + false positives, FP), plus the number of emitters the algorithm failed to detect (false negatives, FN):

$$\text{JAC} = \frac{|\mathcal{L}_{\text{obs}} \bigcap \mathcal{L}_{\text{ref}}|}{|\mathcal{L}_{\text{obs}} \bigcup \mathcal{L}_{\text{ref}}|} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (4.2)$$

Therefore, JAC = 1 reflects a perfect result. The Fourier ring correlation determines how fast the spatial correlation between the localizations and the ground truth declines. The reported value is the inverse of the spatial frequency where the correlation drops to $1/7$ of its maximum.
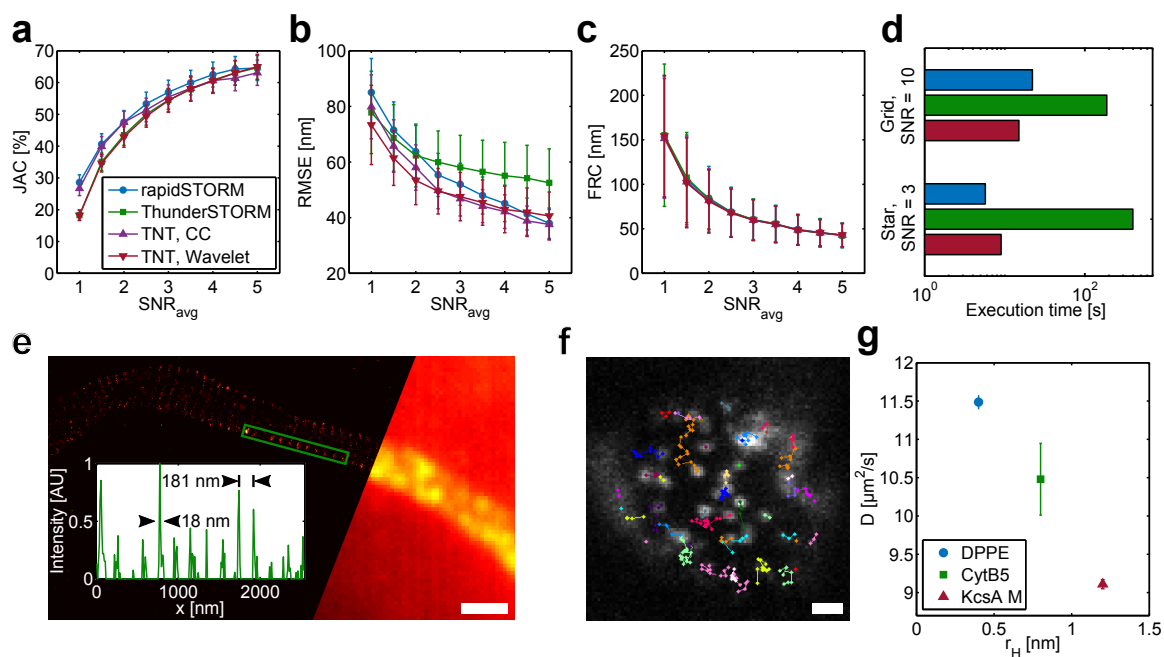
Figure 4.6: Performance evaluation and experimental applications. (a) Jaccard index. (b) Root-mean-square error. (c) Fourier ring correlation. (d) Measurement of execution time. (e) dSTORM image of an axon initial segment from a fixed mouse hippocampal neuron labeled with an Alexa647-tagged antibody against the N-terminus of βIV-spectrin (data recorded by *Melanie Dannemeyer* [107]). The inset shows the normalized 1D intensity projection along the green rectangle's long axis. (f) Atto655-labeled DPPE diffusing in a BLM. (g) Diffusion coefficients obtained for lipid bilayer experiments. Scale bars are 1 μm. [This image has been published in [93].]

Figure 4.6 (a)–(d) shows the result of the evaluation. TrackNTrace was used with the *Cross correlation* (CC) and *Wavelet filtering* (Wavelet) candidate plugins for the first stage and the *TNT fitter* for the second. ThunderSTORM and TrackNTrace, both using wavelet filtering, underperform in the correct detection of emitters (JAC) when compared to rapidSTORM and TrackTrace cross-correlation at low

SNR. However, the RMSE, which depends on the fitting after detection, is similar for all programs at low SNR, indicating a similar spatial resolution. For high SNR, ThunderSTORM shows the lowest precision of localization, while TrackNTrace (CC) achieves the best result, slightly ahead of rapidSTORM. The overall representation of the structure is similar in all cases, as shown by the FRC values. In terms of execution time, TrackNTrace and rapidSTORM show similar performance, while ThunderSTORM is one order of magnitude slower. As rapidSTORM is a highly optimized C++ based software, while TrackNTrace is not optimized for one specific task and runs in MATLAB, we consider the framework's computational performance as high. In summary, this shows that TrackNTrace achieves competitive performance for SMLM applications.

Figure 4.6 (e)–(f) displays the results acquired with TrackNTrace in two typical experimental scenarios: single-molecule localization microscopy and single particle tracking. In both cases the preview feature helped a lot with the necessary optimization of parameters. Panel (e) shows a dSTORM image of an axon initial segment from a fixed mouse hippocampal neuron labeled with an Alexa647-tagged antibody against the N-terminus of βIV-spectrin (data recorded by *Melanie Dannemeyer* [107]). The super-resolution image reveals the periodic structure of the spectrin-actin cytoskeleton, which is unresolved in conventional wide-field imaging. Fourier analysis of a line profile along the structure yields a period of $180 - 190$ nm for the structure, the same as previously reported [108]. We used the tracking feature of TrackNTrace to remove non-blinking emitters, improving image clarity.

For particle tracking, lipids and membrane proteins labeled with Atto655 were incorporated into an artificial black lipid membrane (BLM) spanned over a polytetrafluoroethylene pore 120 μm in diameter. The labeled probes – either DPPE, Cytochrome B5, or a monomeric subunit of the ion channel protein KcsA – were recorded while diffusing through the POPC/POPE BLM. The *TNT NearestNeighbor* plugin was used in conjunction with the previewer for a quick optimization of the tracking parameters, while the u-Track plugin was chosen for the final tracking (Figure 4.6 (f)–(g)). The diffusion coefficients obtained by mean-squared displacement analysis range from $9.1(1)$ $\mu m^2\,s^{-1}$ for the KcsA monomer to $11.5(1)$ $\mu m^2\,s^{-1}$ for DPPE in excellent agreement with earlier fluorescence correlation spectroscopy measurements [109].

Summing up this chapter about TrackNTrace:

- TrackNTrace is an open-source MATLAB framework to support the development of fluorescence imaging applications like single-molecule localization microscopy or particle tracking.
- Its design is focused on straightforward extensibility through plugins, simplicity of coding, and quick and rich feedback. The output of all stages of the processing pipeline can be visualized, facilitating parameter optimization and the recognition of errors.
- Many state-of-the-art algorithms are available out-of-the-box.
- We demonstrate competitive performance and execution speed in SMLM applications compared to established software. In addition, TrackNTrace allows analyzing and tracking of defocused emitters, which is (to our knowledge) currently not possible with other programs.

5

Other contributions

Through my Phd I have contributed in various other pieces of work. In this chapter, I will briefly describe two of these projects, which are interesting for single-molecule imaging and thus thematically very close to the rest of the thesis.

In section 5.1 the brightness distribution of molecules flowing through nanochannels is characterized, quantifying the different phenomena that play a role in the measurement. A better understanding of all contributions is helpful for one of our current projects, where we develop a method to compute the ratio of differently labeled species (different fluorophore or number of fluorophores) within a mixture, using only a camera-based measurement of the intensity distribution. This is of great interest for different biological applications, for example, to determine the stoichiometry of molecular complexes in biological environments.

Section 5.2 shows an imaging-based method to determine the 3D-orientation of the excitation and emission dipole axis of single emitters. In general these axes are not parallel, which is known, for example, from fluorescence anisotropy measurements[110]. To our knowledge, we presented the first method which measures both orientations simultaneously in three dimensions [111].

5.1 Modeling the brightness distribution of molecules in nanochannels

As mentioned, we want to measure the intensity distribution of a mixture of molecules and disentangle the contributions of differently labeled species. Similar measurements were already demonstrated in 1997 by Machara *et al* using a single photon counting avalanche photodiode to detect molecules elut-

ing off an optically trapped microsphere in a well defined sheath flow through a microcapillary [112]. However, the required optical setup is very sophisticated, sensitive and hard to align, which prohibited a broader application of the method. In contrast, we use a single-photon-sensitive EMCCD camera to image molecules flowing through a nanochannel with a cross section of approximately $(45 \times 45)\,\mathrm{nm}^2$ (Figure 5.1), restricting their movement to mostly one dimension. This also produces a well defined flow and keeps molecules from diffusing out of the focal plane. The camera-based measurement is much simpler to perform and allows a higher throughput of molecules, which reduces the necessary measurement time. A manuscript is currently in preparation.

The location of each emitter is extracted using the TrackNTrace software (Chapter 4) and its intensity extracted by integration of its pixel neighborhood. The measurement is background-corrected by estimating the background from the border of the integration window and subtracting it.
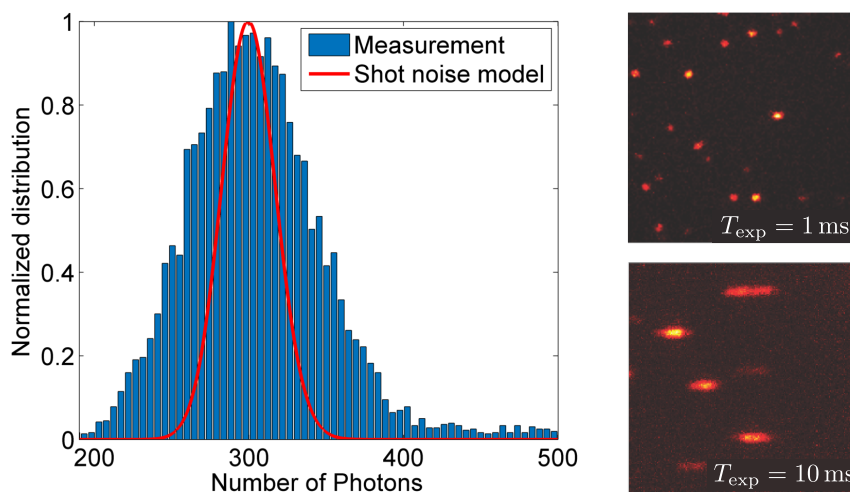
Figure 5.1: Left: Intensity distribution of Cy5 molecules inside nanochannels with a cross section of $(45 \times 45)\,\mathrm{nm}^2$. The distribution is significantly broader than expected for a Poisson distribution. Right: Camera images at high exposure time show the channel structure. For low exposure time the molecules are essentially static, which is preferred for the intensity measurements. Measurements performed by *Hao Cheng*.

As illustrated in the figure, the measured intensity distribution is much broader than expected for a shot-noise-limited measurement. In this section we identify the different phenomena that influence the distribution and build a model that describes the measurement process. An accurate model allows disentangling the contributions of individual species in the distribution of a mixture if the model parameters can be determined with sufficient precision.

Shot noise and EMCCD measurement

When measuring the brightness of fluorescent molecules with a perfect sensor, the resulting distribution of collected photons is the well-known Poisson distribution

$$p_{\mathrm{Poiss}}(k; \lambda) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \tag{5.1}$$

and the measurement is called shot-noise-limited. The Poisson distributions mean λ is also its variance. It is noteworthy that adding multiple random variables with a Poisson distribution yields again a Poisson distribution with a mean equal to the mean of all summed up variables. Thus, integrating a pixel area still yields a Poisson distribution for the overall intensity and it is still Poisson if fluorescent background is present.

An EMCCD camera enables low light imaging by amplifying the signal generated by photons hitting the detector using an electron cascade (EM-gain) while leaving the sensor read-out noise at a constant level. This electron multiplication process leads to a more complex distribution for the number of measured input counts n_{ic} [48]:

$$p_{n_{ic}}(n_{ic}; \lambda, g, s, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\lambda - \frac{(s \cdot n_{ic})^2}{2\sigma^2}\right) + \frac{2}{g}F_\chi(2\lambda; 4, 2s \cdot n_{ic}/g) & n_{ic} > 0 \\ \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\lambda - \frac{(s \cdot n_{ic})^2}{2\sigma^2}\right) & n_{ic} \leq 0 \end{cases} \tag{5.2}$$

Here $F_\chi(2\lambda; 4, 2s \cdot n_{ic}/g)$ is the non-central χ^2 distribution with 4 degrees of freedom and the non-centrality parameter $2s \cdot n_{ic}/g$, λ is the mean number of detected photons, g the EM-gain setting of the camera, s the sensor sensitivity and σ the standard deviation of the readout noise. The distribution can be expressed in terms of measured photons k rather than input counts, using a simple rescaling and normalization.

$$p_k(k; \lambda, g, s, \sigma) = p_{ic}\left(k \cdot \frac{g}{s}; \lambda, g, s, \sigma\right) \cdot \frac{g}{s} \tag{5.3}$$

When enough photons are collected by integration of multiple pixel values following this distribution, the result can be modeled by a scaled and shifted Poisson distribution

$$p_{\text{EMCCD}}(k; \lambda) = p_{\text{Poiss}}(k + \lambda; 2\lambda) = \frac{(2\lambda)^{k+\lambda}}{(k+\lambda)!}e^{-2\lambda}. \tag{5.4}$$

Even for intensities as low as 100 photons the true distribution is approximated with negligible error. The variance of this distribution is

$$\text{Var}_{\text{EMCCD}}(\lambda) = 2\lambda, \tag{5.5}$$

a factor of 2 higher than the shot-noise limited case. It is common to further simplify this to a Gaussian distribution of the same variance and mean, again with negligible error.

Background photons

The number of background photons b collected in the considered area is assumed to be independent from the process that generates the signal ε and follows the distribution of the EMCCD given above. The variance of independent processes is additive, which yields

$$\text{Var}(\varepsilon + b) = Var(\varepsilon) + \text{Var}(b) \tag{5.6}$$

for the overall variance. The distribution of background photons is influenced by shot noise, the EM-CCD measurement and the excitation profile, which is discussed below.

As the brightness of a fluorophore depends on the strength of its excitation, the average emitter brightness varies spatially if the excitation profile is inhomogeneous. This variation leads to a broadening of the intensity distribution of fluorophores in the field of view. In our experiments we use a typical Gaussian excitation profile. Assuming the fluorophores are homogeneously distributed and that the excitation-emission relation is linear (no saturation), the expected brightness $\tilde{\lambda}(r) \in [0, \lambda_{\mathrm{max}}]$ of the fluorophores at position r can be modeled as

$$\tilde{\lambda}(r) = \lambda_{\mathrm{max}} \cdot \exp\left(-\frac{\mathbf{r}^2}{2\sigma_{\mathrm{exc}}^2}\right) \tag{5.7}$$

For a given brightness, the corresponding radius is given by:

$$r(\tilde{\lambda}) = \sqrt{-2\sigma_{\mathrm{exc}}^2 \ln\left(\frac{\tilde{\lambda}}{\lambda_{\mathrm{max}}}\right)} \tag{5.8}$$

$$\Rightarrow r(c) = \sigma_{\mathrm{exc}} \cdot \sqrt{-2\ln(c)}, c \in [0, 1] \tag{5.9}$$

with $c = \tilde{\lambda}/\lambda_{\mathrm{max}} = I_{\mathrm{exc}}(r)/I_{\mathrm{exc}}(r = 0)$. being the ratio between the excitation power at r to its maximum in the focus. In the analysis, only molecules up to a maximum radius r_{max} determined by the cutoff parameter c are taken into account. Limiting the analysis to a smaller area produces a more narrow histogram, but reduces the number of available samples. The probability that a randomly picked point within a disc of radius r_{max} has radius r is

$$p(r) = \frac{2\pi r}{\int\limits_0^{r_{\mathrm{max}}} 2\pi r\, \mathrm{d}r} = \frac{2r}{r_{\mathrm{max}}^2} \tag{5.10}$$

Random variable transformation yields the distribution of the expected average brightness

$$p_{\mathrm{exc}}(\tilde{\lambda}; \lambda_{max}, c) = p(r(\tilde{\lambda})) \cdot \left|\frac{\mathrm{d}r(\tilde{\lambda})}{\mathrm{d}h}\right| = \frac{2\sigma_{\mathrm{exc}}^2}{\tilde{\lambda} r_{\mathrm{max}}^2} = \begin{cases} \frac{-1}{\lambda \cdot \ln(c)} & \lambda_{\mathrm{max}} \cdot c < \lambda < \lambda_{\mathrm{max}} \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Figure 5.2 shows the resulting distribution for different cutoff parameters. With increasing radius (lower c) the distribution shifts to lower intensity, as there is more area that is excited with low power. For c close to 1, the distribution can be approximated as uniform. The average expected brightness λ is simply at the center of the distribution $\lambda_{\mathrm{max}} \cdot \frac{1+c}{2}$. The variance of a uniform distribution spanning the
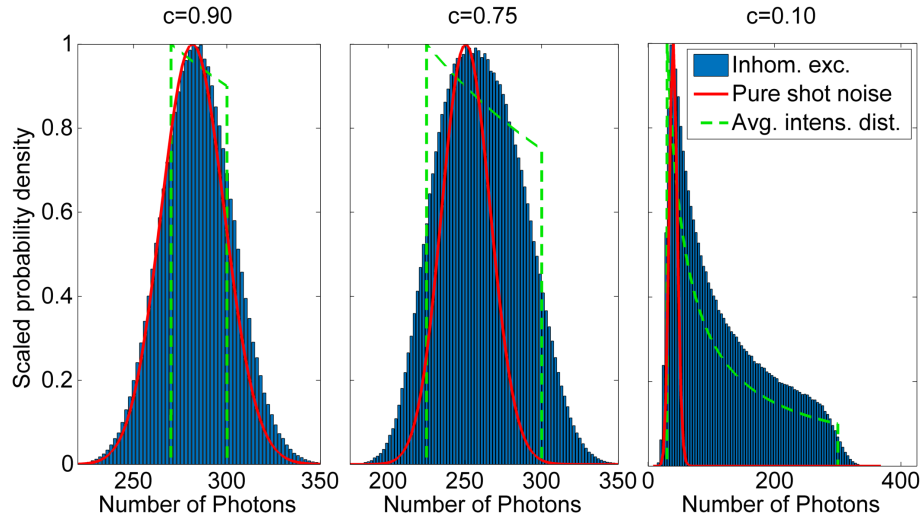
Figure 5.2: Influence of a Gaussian excitation profile on the measured intensity distribution. Molecules are only picked from a circular area around the focus, which's size is defined by the cutoff parameter c. At the border of the area, the intensity drops to c times the maximum.

interval Δx is $\frac{1}{12}\Delta x^2$, which yields

$$\text{Var}_{\text{Exc}}(\lambda, c) = \frac{1}{12}\left(\lambda_{\text{max}} \cdot (1 - c)\right)^2 = \frac{1}{12}\left(\frac{2\lambda}{1 + c} \cdot (1 - c)\right)^2. \tag{5.12}$$

Although the power of the excitation (and therefore emission) influences the strength of the other sources of noise (e.g. background, shot noise, photoblinking), for small cutoff radii the excitation is almost constant and we can approximate the processes as independent. In this case, the variances simply add up.

Photophysics

We model the photophysical behavior of fluorophores using a two-state model, in which it is either in the "on"-state (emitting photons) or in the "off"-state (dark state). The rates of transition between the states are k_+ (off to on) and k_- (on to off). This two-state system is a special case of a Continuous Time Markovian Chain (CTMC). The distribution of the occupation time $T_{\text{on}} \in [0, T_{\text{exp}}]$ of the on-state within exposure time T_{exp} is given by [113]

$$p_{\text{on}}(T_{\text{on}}; T_{\text{exp}}, k_+, k_-) = \tag{5.13}$$

$$e^{-k_- T_{\text{on}} - k_+(T_{\text{exp}} - T_{\text{on}})}\left\{ \frac{k_+}{k_+ + k_-}\delta(T_{\text{exp}} - T_{\text{on}}) + \frac{k_-}{k_+ + k_-}\delta(T_{\text{on}}) \right.$$

$$+ \left[\frac{k_+}{k_+ + k_-}\sqrt{\frac{k_- k_+ T_{\text{on}}}{T_{\text{exp}} - T_{\text{on}}}} + \frac{k_-}{k_+ + k_-}\sqrt{\frac{k_- k_+ (T_{\text{exp}} - T_{\text{on}})}{T_{\text{on}}}} \right] \cdot I_1\left(2\sqrt{k_- k_+ T_{\text{on}}(T_{\text{exp}} - T_{\text{on}})} \right)$$

$$\left. + \frac{2k_+ k_-}{k_+ + k_-} \cdot I_0\left(2\sqrt{k_- k_+ T_{\text{on}}(T_{\text{exp}} - T_{\text{on}})} \right) \right\},$$

with the modified Bessel function of order r

$$I_r(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k+r}}{k!(k+r)!}.$$ (5.14)

The expected accumulated on-time is

$$\langle T_{\text{on}} \rangle = \frac{k_+}{k_+ + k_-} \cdot T_{\text{exp}}.$$ (5.15)

The rate of emitted photons per time is constant in the on-state, giving a linear relationship between the time spent in the on-state and the expected number of emitted photon $\tilde{\lambda}$ during that time. If the average expected number of emitted photons within the exposure time is λ, the distribution for the expected number of photons can be expressed as

$$p_{\text{blink}}(\tilde{\lambda}; T_{\text{exp}}, k_+, k_-, \lambda) =$$
$$\begin{cases} p_{\text{on}}\left(\frac{\tilde{\lambda}}{\lambda} \cdot \frac{k_+}{k_+ + k_-} \cdot T_{\text{exp}}; T_{\text{exp}}, k_+, k_- \right) \cdot \frac{1}{\lambda} \cdot \frac{k_+}{k_+ + k_-} \cdot T_{\text{on}} & \frac{\tilde{\lambda}}{\lambda} \cdot \frac{k_+}{k_+ + k_-} \cdot T_{\text{exp}} \leq T_{\text{exp}} \\ 0 & \text{otherwise} \end{cases}$$ (5.16)

This can be easily combined with other effects, like shot noise, by integration. For example, the number of photons detected by an EMCCD (using the approximation above) is modeled by

$$p_{\text{blink+EMCCD}}(k) = \int_0^{\infty} p_{\text{blink}}(\tilde{\lambda}; T_{\text{exp}}, k_+, k_-, \lambda) \cdot p_{\text{EMCCD}}(k; \tilde{\lambda}) \, d\tilde{\lambda}$$ (5.17)

The average occupation times of the states $\tau_{\text{on}} = \frac{1}{k_-}, \tau_{\text{on}} = \frac{1}{k_+}$ are the inverse of the rates. The length of individual events follows a simple exponential distribution $p_{\text{on/off}}(\tau) = \frac{1}{\tau_{\text{on/off}}} \cdot \exp\left(-\frac{1}{\tau_{\text{on/off}}}\tau \right)$. A simulation of the two-state process was performed drawing random samples for the individual occupation times from their distributions and showed perfect agreement with the theory. Calculation of the analytical distribution using the Bessel functions built into MATLAB leads to numbers too small for processing even on 64-bit-systems if the ratio between exposure time and on-/off times becomes too high (approx. 1ms to 1μs). For this reason, simulations are used for illustration in Figure 5.3, which compares pure shot noise with a dye subject to photoblinking according to the model above.

It is easy to see that photoblinking visibly broadens the measured distribution, even for average occupation times more than 100-fold smaller than the exposure time. Note that distributions are much broader for higher values of the ratio $\tau_{\text{off}}/\tau_{\text{on}}$. However, this is easily explained: The mean number of photons is fixed to 300 in all cases. If the fluorophore spends less time in the off-state, it must have a higher instantaneous brightness in the on-state. Thus, there is a possibility to emit many more photons compared to a dimmer dye if the fluorophore, by chance, spends a long time in the on-state. As the analysis shows, photoblinking can have major effects on the brightness distribution. It is only negligible if the blinking time-scale is roughly 1000-fold below the exposure time and the average occupation times
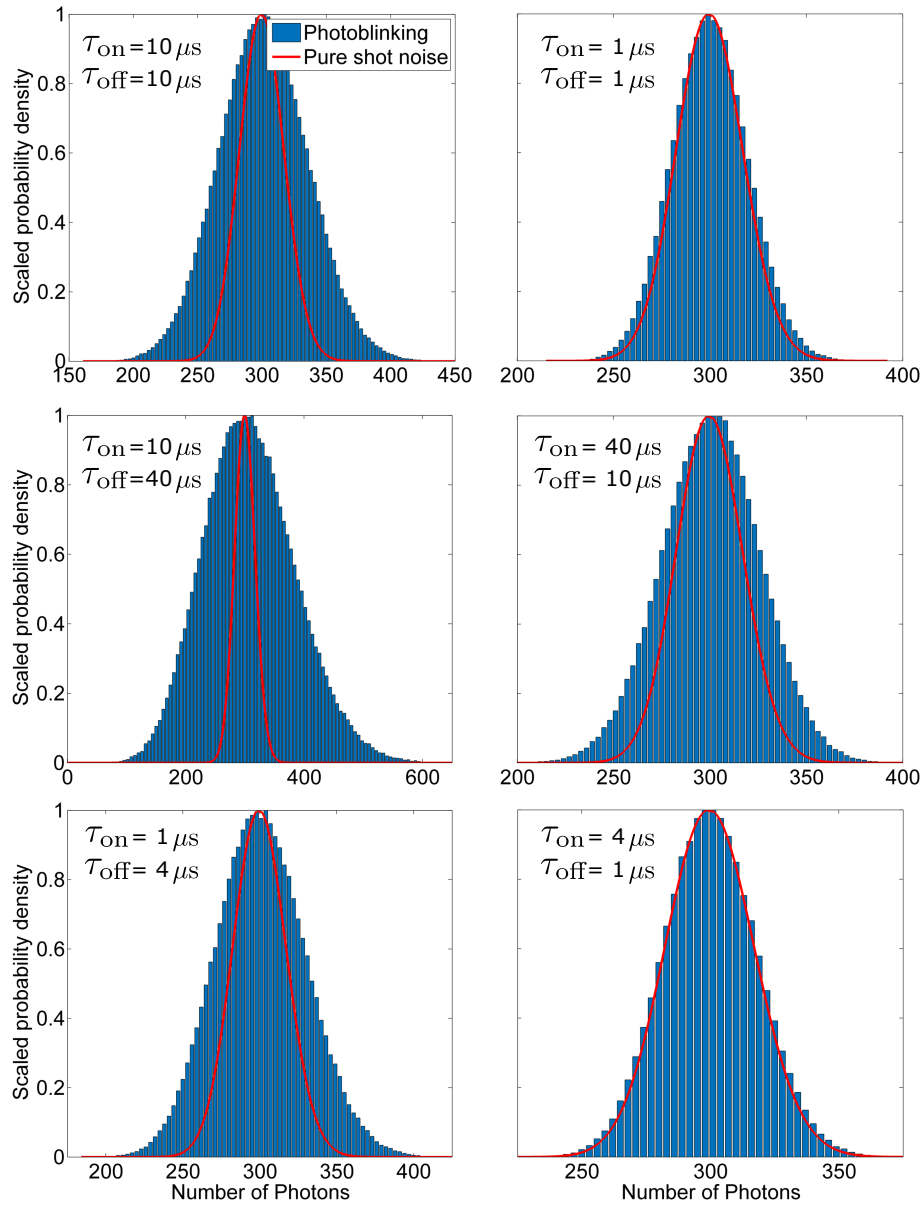
Figure 5.3: Influence of the average occupation times $\tau_{\mathrm{on}}, \tau_{\mathrm{off}}$ of the on-/off-state in a two-state blinking model on the measured intensity distribution. The red curve shows pure Poisson noise for comparison. The average photon count is fixed to $\lambda = 300$ photons. The occupation times of states are drawn from the corresponding exponential distributions and integrated within the exposure time $T_{\mathrm{exp}} = 1$ ms. Shot noise is added by drawing the number of measured photons from a Poisson distribution with an average set by the total occupation time of the on-state during exposure.

of the off- and on-states are comparable.

As the overall occupation time is the sum of many statistically independent on-states, its distribution is close to a Gaussian. To incorporate the blinking into our model, the variance of the distribution is derived as follows. Let $x(t), x \in \{0, 1\}$ describe the state of one molecule over time with $x(t) = 1$ representing the on-state and $x(t) = 0$ the off-state. Within one exposure time, the expected occupation time is

$$\langle T_{\mathrm{on}} \rangle = \left\langle \int_0^{T_{\mathrm{exp}}} x(t)\, \mathrm{d}t \right\rangle = \int_0^{T_{\mathrm{exp}}} \langle x(t) \rangle\, \mathrm{d}t = T_{\mathrm{exp}} \frac{k_+}{k_+ + k_-} \tag{5.18}$$

as $\frac{k_+}{k_+ + k_-}$ is the chance to observe the molecule in the on-state. The variance of this estimator is

$$\text{Var}\left(T_{\text{on}}\right) = \left\langle \left(\int_0^{T_{\text{exp}}} x(t)\,\text{d}t - \langle T_{\text{on}}\rangle\right)^2 \right\rangle = \int_0^{T_{\text{exp}}} \text{d}t \int_0^{T_{\text{exp}}} \langle x(t)x(t')\rangle \,\text{d}t' - \langle T_{\text{on}}\rangle^2. \tag{5.19}$$

This can be rewritten using the stationarity and time-inversion symmetry of the process

$$\text{Var}\left(T_{\text{on}}\right) = 2 \cdot \int_0^{T_{\text{exp}}} \text{d}t \int_0^{T_{\text{exp}}-t} \langle x(0)x(\tau)\rangle \,\text{d}\tau - \langle T_{\text{on}}\rangle^2. \tag{5.20}$$

The value $\langle x(t)x(t')\rangle$ is given by the chance to observe the molecule in the on-state at $t = 0$ times the conditional probability to still observe it after at time $t = \tau$, which is given by

$$\langle x(0)x(\tau)\rangle = \frac{k_+}{k_+ + k_-} \cdot \left(\frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-}\text{e}^{-(k_+ + k_-)\tau}\right) \tag{5.21}$$

$$= \frac{1}{(k_+ + k_-)^2} \cdot \left(k_+^2 + k_+ k_- \text{e}^{-(k_+ + k_-)\tau}\right). \tag{5.22}$$

Solving the double integral for the first term simply yields $\langle T_{\text{on}}\rangle^2$, leaving the variance

$$\text{Var}\left(T_{\text{on}}\right) = \frac{2k_+ k_-}{(k_+ + k_-)^2} \cdot \int_0^{T_{\text{exp}}} \text{d}t \int_0^{T_{\text{exp}}-t} \text{e}^{-(k_+ + k_-)\tau} \,\text{d}\tau \tag{5.23}$$

$$= \frac{2k_+ k_-}{(k_+ + k_-)^2} \cdot \left[\frac{T_{\text{exp}}(k_+ + k_-) + \text{e}^{-(k_+ + k_-)T_{\text{exp}}} - 1}{(k_+ + k_-)^2}\right]. \tag{5.24}$$

The number of expected photons has a linear relationship with the occupation time in the on-state. Thus, the variance of the expected number of photons under influence of photoblinking is

$$\text{Var}_{\text{blink}}\left(\lambda, k_+, k_-, T_{\text{exp}}\right) = \frac{\lambda^2}{\langle T_{\text{on}}\rangle^2} \cdot \text{Var}\left(T_{\text{on}}\right) = \lambda^2 \cdot \left(\frac{k_+ + k_-}{k_+ T_{\text{exp}}}\right)^2 \cdot \text{Var}\left(T_{\text{on}}\right), \tag{5.25}$$

where λ is the average expected number of emitted photons.

OVERALL VARIANCE OF THE BRIGHTNESS DISTRIBUTION

Compiling all influences on the brightness distribution, we model the following experimental scenario: A molecule is detected at a random position r and excited by the local laser intensity $I_{\text{exc}}(r)$. Within the exposure time it is subject to photoblinking and the average number of emitted photons depends on the excitation intensity, as well as on the time T_{on} it spends in the on-state. The number of finally emitted photons for this specific molecule is random, following a Poisson distribution (shot noise), and further randomness is added by the detection process via an EMCCD camera. At the same time, a certain number of background photons is collected, also depending on the local excitation intensity, shot noise and the EMCCD measurement. Assuming that the influence of excitation and photoblinking

on the distribution of the average number of expected photons is relatively small, the distributions of the EMCCD detection belonging to the different averages will be very similar. In this case, we can approximate the final number of photons as the sum of different independent noise sources and the final variance is just the sum of the individual variances:

$$
\mathrm{Var_{Int}}\left(\lambda, b, c, k_+, k_-, T_{\mathrm{exp}}\right) =
$$
$$
\mathrm{Var_{EMCCD}}\left(\lambda\right) + \mathrm{Var_{EMCCD}}\left(b\right) + \mathrm{Var_{Exc}}\left(\lambda, c\right) + \mathrm{Var_{Exc}}\left(b, c\right) + \mathrm{Var_{blink}}\left(\lambda, k_+, k_-, T_{\mathrm{exp}}\right) =
$$
$$
2(\lambda + b) + \frac{1}{12}\left(\frac{2(\lambda + b)}{1 + c} \cdot (1 - c)\right)^2 +
$$
$$
\lambda^2 \cdot \left(\frac{k_+ + k_-}{k_+ T_{\mathrm{exp}}}\right)^2 \cdot \frac{2k_+ k_-}{(k_+ + k_-)^2} \cdot \left[\frac{T_{\mathrm{exp}}(k_+ + k_-) + e^{-(k_+ + k_-)T_{\mathrm{exp}}} - 1}{(k_+ + k_-)^2}\right]
\tag{5.26}
$$

with λ and b the average number of photons detected from the fluorophore and background respectively, $c \in [0, 1]$ the ratio between the excitation intensity at the radius of a circular area around the maximum to the excitation intensity in the focus (only molecules within this area are considered for the analysis), k_+ and k_- the transition rates from the off- to the on state and vice versa, and the exposure time T_{exp}. After subtracting the background, the final distribution can be approximated by a Gaussian with the mean λ and the standard deviation $\sigma_I = \sqrt{\mathrm{Var_{Int}}\left(\lambda, b, c, k_+, k_-, T_{\mathrm{exp}}\right)}$.
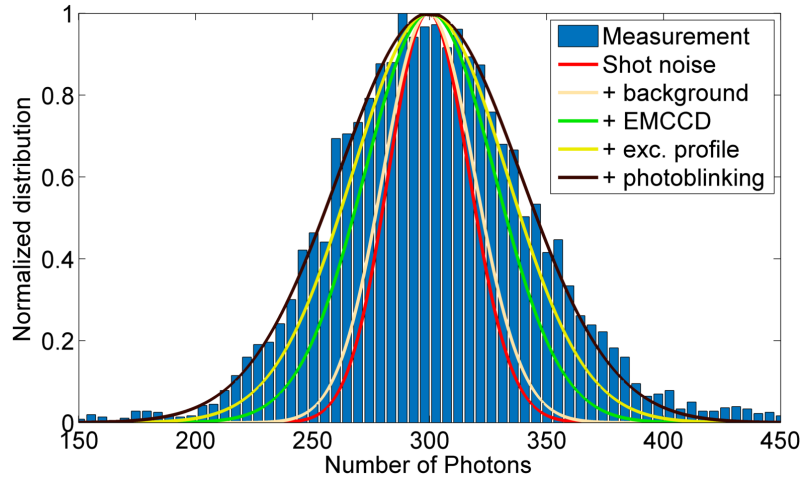
Figure 5.4: Nanochannel measurement of Cy5 molecules compared the theoretical model. The exposure time is $T_{\mathrm{exp}} = 1$ ms. Molecules were only picked up from a circular area defined by $c = 0.85$. Background subtraction was performed for every molecule individually and the average background $b \approx 83$ was used for the model. We estimated the state lifetimes to $\tau_{\mathrm{on}} \approx 12\,\mu s$, $\tau_{\mathrm{off}} \approx 6.7\,\mu s$ using fluorescence correlation spectroscopy. Measurements performed by *Hao Cheng*.

Figure 5.4 shows a measurement of Cy5 molecules inside the nanochannel. The measurement was performed by *Hao Cheng*. Molecules were only picked up from a circular area defined by $c = 0.85$. Background subtraction was performed for every molecule individually and the average background $b \approx 83$ was used for the model. We estimated the state lifetimes using fluorescence correlation spectroscopy (FCS). The obtained values were sensitive to the cutoff (fastest time-scale) that was used for the FCS evaluation, giving $2.3\,\mu s$ to $12.0\,\mu s$ for τ_{on} and $2.0\,\mu s$ to $6.7\,\mu s$ for τ_{off} for cutoffs between $0.1\,\mu s$ and $1.0\,\mu s$. This is caused by unknown dynamics that we observed on the fastest time-scale. Neverthe-

less, all combinations of values broaden the observed distribution and fit the curve reasonably well. We choose $\tau_{on} \approx 12\,\mu s$, $\tau_{off} \approx 6.7\,\mu s$ for the visualization in Fig.5.3, as the fast time-scale has the lowest influence for the largest cutoff[1]. Our model matches the data well and explains the observed broadening of the intensity distribution.

In practice, the photoblinking of fluorophores in the channel is hard to predict. When differently labeled species (different fluorophore or number of fluorophores) should be distinguished in a mixture, it has been proven useful to first measure the individual species under the same conditions, flushing the channel with buffer between each measurement. The individual measurement are fitted with a Gaussian model to extract the average brightness and the width of the histogram. This information is used when fitting the intensity distribution of the mixture, giving an accurate estimate of the fraction of each species in the sample. We are currently preparing a manuscript demonstrating this type of measurement. More information can be found in the PhD thesis of *Hao Cheng*[2].

Other phenomena that influence the shape of the intensity histogram include diffusion, bleaching, saturation by high excitation and multiple fluorescent states of the fluorophore. In our case, we reduced the influence of diffusion by choosing a sufficiently large integration area for the intensity measurement. This is a trade-off, as we also collect more background photons, but is much easier to model. Bleaching and saturation were avoided by an appropriately low laser intensity. Note that bleaching has a larger influence when trying to distinguish samples labeled with different numbers of fluorophores, as the chance that some of the fluorophores bleach during exposure increases with their number. Multiple fluorescent states can in principle be modeled in the same way that we used here for the two-state system. However, care has to be taken to avoid over-fitting the experimental data.

To sum up this section about modeling the brightness distribution of molecules in nanochannels:

- The measured brightness distribution is much broader than a shot-noise limited model.

- Major contributions to this broadening are the background signal, the EMCCD measurement process, the inhomogeneity of the excitation profile and photoblinking.

- Surprisingly, photoblinking is still relevant to the measurement, even if the lifetime of the on/off-states is multiple hundred times faster than the exposure time.

- The final model fits well to the experimental data.

---

[1]Larger cutoffs are not advisable, as the cutoff must be smaller than the time-scale of the observed blinking and we expect Cy5 to blink on the μs time-scale.

[2]Hao Cheng, "Probing Molecular Stoichiometry by Photon Antibunching and Nanouidics Assisted Imaging in Solution", Phd thesis, University of Göttingen. Publication expected for 2017.

5.2 SIMULTANEOUS MEASUREMENT OF THE THREE-DIMENSIONAL ORIENTATION OF EXCITATION AND EMISSION DIPOLES OF SINGLE MOLECULES

Fluorescent organic dyes are well described by an ideal electric dipole emitter, which is characterized by an oscillation frequency and the vectorial electric dipole moment, which captures the strength and orientation of the dipole. In more detailed models, each fluorescent molecule is modeled by a superposition of two dipoles, one describing its emission and one its excitation. The axes of these dipoles are in general not parallel, which is known, for example, from fluorescence anisotropy measurements[110]. As an improvement over previous techniques, we developed a method to measure the three-dimensional excitation and emission dipole orientations of single-emitters simultaneously. This enables us to measure distributions of the orientations in an ensemble of molecules. My contribution to the project focused on the evaluation of the single-molecule patterns.

The work presented in this section was published in the article [111]:

Karedla, N.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Chizhik, A.; Enderlein, J.; "Simultaneous Measurement of the Three-Dimensional Orientation of Excitation and Emission Dipoles", Physical Review Letters, 115, 173002, 2015.

The measurements are performed as follows: The linearly polarized light from a white light laser is passed through a liquid-crystal mode converter to generate a radially polarized beam. The beam is reflected into an 1.49 N.A. oil immersion objective by a non-polarizing beam-splitter, which focuses it onto the surface. In the first step of the measurement, images of the excitation dipole are generated by scanning the sample using a movable piezoelectric stage. Defocused images of the emission dipole are recorded in the second step, by sequentially parking the excitation focus on each registered molecule's position and guiding the emitted light onto an EMCCD camera using a replaceable mirror. Both measurements, radially polarized scanning as well as defocused imaging, generate images that are sensitive to the orientation of the recorded dipole [95, 114].

For the evaluation of the acquired images, we performed wave-optical calculations to compute a discrete set of theoretical excitation scan images for different azimuthal, φ and polar angles, θ, in $5°$ steps. A least-squares pattern matching was used find the closest matching template for each molecule's intensity pattern, yielding a rough estimate of its position and orientation [96]. The estimate is refined using a Poisson-noise-based maximum likelihood estimation with the log-likelihood function

$$l = -\sum_{\mathbf{r}} \{I(\mathbf{r}) \cdot \log[A\,P(\mathbf{r}|\mathbf{r}_P, \varphi, \theta) + B] - [A\,P(\mathbf{r}|\mathbf{r}_P, \varphi, \theta] + B]\} \tag{5.27}$$

where $I(\mathbf{r})$ is the measured image, $P(\mathbf{r}|\mathbf{r}_P, \varphi, \theta)$ a normalized theoretical pattern, A the integrated intensity, B the per-pixel background intensity and $\mathbf{r}_P$ the center position of the pattern. The optimization is executed using a nonlinear conjugate gradient algorithm [56]. In contrast to pattern matching, the produced parameter estimates are continuous, achieving higher precision without the computational
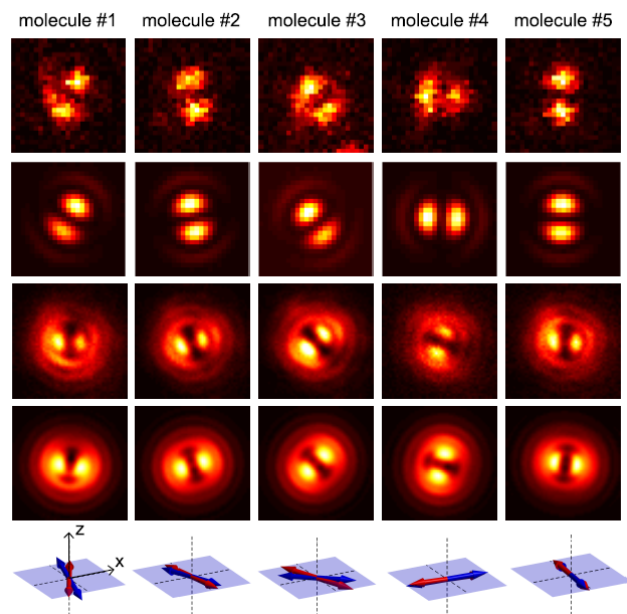
Figure 5.5: Excitation patterns (1st row) of five Atto655 molecules spincoated on a glass coverslip and their defocused images (3rd row, 0.9 μm defocusing) together with the corresponding fitted patterns (2nd and 4th row respectively). The last row is a visualization of the extracted dipole orientations. The excitation scan pixel size is 50 nm, whereas the camera pixel size is 60 nm. [This image has been published in the article [111].]

cost associated with a much larger template library. Pattern matching is still required to get rough initial guesses however, as the function to optimize is not convex, but describes a very complicated probability landscape. The presented evaluation was integrated into the *TrackNTrace* software (Chap. 4), to simplify similar experiments in the future.

Figure 5.5 shows recorded excitation and emission patterns of Atto655 molecules spincoated on a glass coverslip together with the fitted patterns. The total number of collected photons is on the order of a few thousand for the excitation pattern and a few hundred thousand for the emission pattern. While the overall quality of the fit is quite good, the measured patterns show a small asymmetry that can not be captured by our model. This is probably caused by a residual misalignment of the imaging system, as the influence of aberrations increases with the defocusing distance, making even small errors visible. The θ angle of all molecules is close to $90°$, indicating that they lie almost flat on the surface after spincoating. We applied a bootstrapping algorithm to estimate the random error of the fit: New synthetic images are created from the final fitting parameters and Poisson noise is added. Then the maximum likelihood estimator is used with random initial values to fit one thousand of the resampled images. The standard deviations of the resulting distributions are about $\pm 2°$ on average for the excitation patterns and $\pm 0.4°$ for the emission patterns for both angles.

Figure 5.6 shows the distribution of the angle β between both dipole axes for Atto655 on glass and Alexa488 in PVA. The Atto molecules show lower bootstrapping errors due to a higher photon output and defocusing value (0.9 μm compared to 0.6 μm for Alexa488). The determined θ values for Alexa488 vary between $0°$ and $90°$, indicating that there is no preferred orientation within the polymer, in contrast to the orientation parallel to the surface that was observed for the spincoated Atto molecules. We
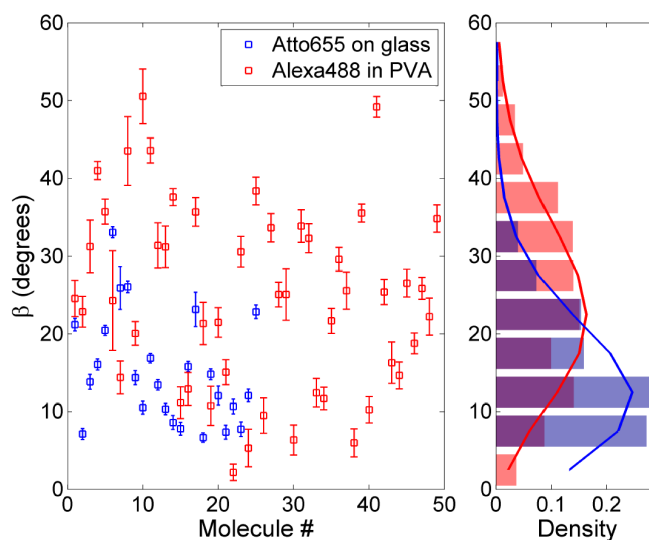
Figure 5.6: Measured distribution of the angle β between the emission and excitation dipole axes for 25 Atto655 molecules on a glass coverslip and 49 Alexa488 molecules embedded into a polymer film. The error bars were estimated by bootstrapping. [This image has been published in the article [111].]

extracted mean values of $\beta_{\text{Att}} = 14.6°$ for Atto655 and $\beta_{\text{Ale}} = 22.5°$ for the Alexa488 dye by fitting the data with a Poisson distributions. To validate the results, we performed conventional time-resolved anisotropy measurements on bulk solution, which yielded $14.9°$ and $19.9°$ for Atto655 and Alexa488 respectively, in good agreement with the imaging-based values. However, the new methods measures the full distributions instead of averages, which reveals a broader spectrum of values for the Alexa dye in PVA than for Atto on glass.

To sum up this section about the measurement of the emission and excitation dipole axes of single molecules:

- Images of the excitation dipole are recorded by scanning the sample with a radially polarized laser beam, while defocused imaging is used for the emission dipole.

- Patten matching with a discrete set of theoretical dipole images is used for a rough estimate of the orientation. This is used as an initial guess for iterative maximum likelihood estimation, yielding continuous parameter values.

- We measured the distribution of the angle between the axes for Atto655 molecules on a glass coverslip as well as Alexa488 in PVA. The estimated mean values are in good agreement with time-resolved fluorescence anisotropy measurements. We observed a broader spectrum of values for the Alexa dye.

6

Conclusion

I{.dropcap}N this thesis, I have introduced several improvements connected to the fields of single-molecule and super-resolution imaging, mainly concerning Super-resolution Optical Fluctuation Imaging (SOFI). As the study of its convergence properties shows, 2nd order SOFI requires only a small amount of frames to be recorded and works well even for densely labeled samples. Higher orders do, in principle, provide better resolution, but require a lot more frames to converge and are more sensitive to the emitter density, which limits their performance in practice. Interestingly, the image quality at high densities improves, even for higher orders, when the data is pre-processed to extract higher resolution information (Sec. 2.7.3). Potentially, this phenomenon could be used to widen the applicability of higher order SOFI. With *Fourier SOFI* we presented a new technique to generate super-resolved images on a finer pixel grid than the original camera recording. In contrast to established algorithms relying on spatial cross-cumulants, this method is practically free of artifacts and does not require any post-processing corrections. Next, the influence of noise on zero time-lag SOFI images was investigated and an algorithm was shown to correct for the noise-dependent bias. This extends the applicability of auto-cumulant SOFI – which correlates values only in time, not in space – to recordings where the time-scale of photoblinking is on the same order as the exposure time. However, cross-cumulant techniques might achieve a better SNR in some scenarios, indicating the need for a detailed comparison of these different methods. We also showed how the microscope PSF can be estimated from typical SOFI data. This works best with a TIRF microscope, where only emitters in a narrow axial range are visible, as variations in the axial distance cause errors. However, this issue is resolved if a 3D imaging system is employed for SOFI [28], which allows a three-dimensional reconstruction of the PSF. The second to last part of Chapter 2 deals with improving the resolution of SOFI images. We developed a

new method, termed *Fourier Preweighting*, which preprocesses the original data in the Fourier domain, resulting in an increased resolution after the SOFI calculation, outperforming current post-processing techniques as well as conventional iterative deconvolution schemes. As mentioned, this improves the convergence properties of higher order calculations at increased particle density. A related algorithm is presented which automatically matches the degree of resolution enhancement to the data quality, avoiding artifacts. While it performs well on both simulated and experimental data, the algorithm is partly built upon empirically derived measures and could be improved by a thorough theoretical examination, which was out of the scope of this thesis. In the final section, we show that SOFI can be successfully applied to cells labeled with carbon nanodots, a fairly new type of fluorophore. Carbon nanodots are bio-compatible and cheap to make, but difficult to use with most other super-resolution methods. Additionally, we performed an analysis of their photoblinking behavior, revealing a power law distribution for the occupation time of the dark state. This suggests electron transfer, rather than triplet-state dynamics, as the main cause of their fluorescence intermittency. Currently, the most exciting development in SOFI is the introduction of 3D imaging systems [28], which allows the simultaneous observation of three-dimensional samples with a wide-field microscope. This enables the study of dynamic processes in live cells and could be enhanced by extending the here presented improvements to 3D data, which is in many cases straightforward.

In the third chapter we demonstrated a drastically increased photostability and overall photon yield of Atto647N imaged at liquid nitrogen temperature of 89 K. The enhanced photon yield could be exploited by a newly developed evaluation scheme to localize single molecules with sub-nanometer precision. Furthermore, we proposed a setup based on polarized excitation and emission to resolve spectrally identical, non-blinking fluorophores spaced only few nanometers apart and validated our concept in simulations. We hope that our method allows easy and reliable distance measurements, in contrast to currently available methods that use less than 1% of the sample. Although the necessary setup was built and calibrated, the experimental verification is still pending due to a lack of suitable dual-label samples, which survive the preparation process necessary for cryogenic measurements.

We also developed *TrackNTrace*, an open-source MATLAB framework to support the development of fluorescence imaging applications. Its design is focused on easy extensibility through plugins, simplicity of coding, and rich visual feedback. The output of all stages of the processing pipeline can be visualized, facilitating parameter optimization and the recognition of errors. We demonstrated competitive performance and execution speed in SMLM applications compared to established software and include many state-of-the-art algorithms out-of-the-box. TrackNTrace was not only used for the development of many of the data evaluation algorithms presented in this thesis, but is already employed by different research groups in Europe.

Finally, Chapter 5 presents two techniques reliant on single-molecule imaging. We developed a model for the intensity distribution of fluorescent molecules flown through a nanochannel and imaged by a camera. As a surprising insight, we learned that photoblinking of the fluorophore has considerable influence on the distribution, even if its time-scale is 100 times faster than the exposure time of the camera. We are currently developing a method to accurately disentangle the contributions of differently labeled

samples from an intensity measurement of a mixture. In the second project we demonstrate a simultaneous measurement of the excitation and emission dipole axes of single molecules. This is achieved by scanning the sample with a polarized laser beam followed by recording defocused images of the previously detected molecules. We measured the ensemble distribution of the angle between the excitation and emission axis for two common fluorophores in different environments, and verified our results with time-resolved fluorescence anisotropy measurements.

The progress of fluorescence microscopy today depends first and foremost on clever new experimental techniques and the development of advanced data processing methods. In this thesis, I focused on the latter and introduced various improvements to Super-resolution Optical Fluctuation Imaging, cryo-fluorescence microscopy and single-molecule imaging. I hope that the presented work illustrates how smart new ways to handle the acquired data can improve upon techniques without the need to buy expensive new hardware or change long-established practices and protocols.

Bibliography

[1] Timothy C. Kriss and Vesna Martich Kriss. History of the operating microscope: From magnifying glass to microneurosurgery. Neurosurgery, 42(4):899–907, April 1998.

[2] George Sines and Yannis A. Sakellarakis. Lenses in antiquity. American Journal of Archaeology, 91(2):191, April 1987.

[3] Olaf Schmidt, Karl-Heinz Wilms, and Bernd Lingelbach. The visby lenses. Optometry and Vision Science, 76(9):624–630, September 1999.

[4] Al Van Helden. The galileo project - the telescope. http://galileo.rice.edu/sci/instruments/telescope.html, 1995. Last accessed 11th May 2017.

[5] Roshdi Rashed. A pioneer in anaclastics: Ibn sahl on burning mirrors and lenses. Isis, 81(3):464–491, September 1990.

[6] August Köhler. Ein neues beleuchtungsverfahren für mikrophotographische zwecke. Zeitschrift für wissenschaftliche Mikroskopie und für mikroskopische Technik, 10(4):433–440, 1893.

[7] Ernst Abbe. XV.-the relation of aperture and power in the microscope (continued)∗. Journal of the Royal Microscopical Society, 3(6):790–812, December 1883.

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

[9] William E. Moerner, Robert M. Dickson, Andrew B. Cubitt, and Roger Y. Tsien. On/off blinking and switching behaviour of single molecules of green fluorescent protein. Nature, 388(6640):355–358, July 1997.

[10] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. Science, 313(5793):1642–1645, September 2006.

[11] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein. Fast, background-free, 3d super-resolution optical fluctuation imaging (SOFI). Proceedings of the National Academy of Sciences, 106(52):22287–22292, December 2009.

[12] Hellen C. Ishikawa-Ankerhold, Richard Ankerhold, and Gregor P. C. Drummen. Advanced fluorescence microscopy techniques—FRAP, FLIP, FLAP, FRET and FLIM. Molecules, 17(12):4047–4132, April 2012.

[13] Hermann Helmholtz and H. Fripp. On the limits of the optical capacity of the microscope. The Monthly Microscopical Journal, 16(1):15–39, July 1876.

[14] Jörg Enderlein. Advanced fluorescence microscopy. In Anders Brahme, editor, Comprehensive Biomedical Physics, volume 4, chapter 9, pages 111–151. Elsevier, 2014.

[15] Fu Min Huang, Yifang Chen, F Javier Garcia de Abajo, and Nikolay I Zheludev. Optical super-resolution through super-oscillations. Journal of Optics A: Pure and Applied Optics, 9(9):S285–S288, August 2007.

[16] E. Betzig and R. J. Chichester. Single molecules observed by near-field scanning optical microscopy. Science, 262(5138):1422–1425, November 1993.

[17] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. Applied Optics, 46(10):1819–1829, March 2007.

[18] George Biddell Airy. On the diffraction of an object-glass with circular aperture. Transactions of the Cambridge Philosophical Society, 5:283–291, 1835.

[19] Claus B. Müller and Jörg Enderlein. Image scanning microscopy. Physical Review Letters, 104(19):198101, May 2010.

[20] M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy, 198(2):82–87, May 2000.

[21] Dominik Wildanger, Brian R. Patton, Heiko Schill, Luca Marseglia, J. P. Hadden, Sebastian Knauer, Andreas Schönle, John G. Rarity, Jeremy L. O'Brien, Stefan W. Hell, and Jason M. Smith. Solid immersion facilitates fluorescence microscopy with nanometer resolution and sub-ångström emitter localization. Advanced Optical Materials, 24(44):OP309–OP313, September 2012.

[22] Eva Wegel, Antonia Göhler, B. Christoffer Lagerholm, Alan Wainman, Stephan Uphoff, Rainer Kaufmann, and Ian M. Dobbie. Imaging cellular structures in super-resolution with SIM, STED and localisation microscopy: A practical comparison. Scientific Reports, 6(1), June 2016.

[23] Rainer Heintzmann, Thomas M. Jovin, and Christoph Cremer. Saturated patterned excitation microscopy—a concept for optical resolution improvement. Journal of the Optical Society of America A, 19(8):1599–1609, August 2002.

[24] M. G. L. Gustafsson. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. Proceedings of the National Academy of Sciences, 102(37):13081–13086, September 2005.

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

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

[27] Stefan Geissbühler, Claudio Dellagiacoma, and Theo Lasser. Comparison between SOFI and STORM. Biomedical Optics Express, 2(3):408–420, January 2011.

[28] Stefan Geissbühler, Azat Sharipov, Aurélien Godinat, Noelia L. Bocchio, Patrick A. Sandoz, Anja Huss, Nickels A. Jensen, Stefan Jakobs, Jörg Enderlein, F. Gisou van der Goot, Elena A. Dubikovskaya, Theo Lasser, and Marcel Leutenegger. Live-cell multiplane three-dimensional super-resolution optical fluctuation imaging. Nature Communications, 5:5830, December 2014.

[29] Thomas Dertinger, Jianmin Xu, Omeed Naini, Robert Vogel, and Shimon Weiss. SOFI-based 3d superresolution sectioning with a widefield microscope. Optical Nanoscopy, 1(1):2, 2012.

[30] Zhiping Zeng, Xuanze Chen, Hening Wang, Ning Huang, Chunyan Shan, Hao Zhang, Junlin Teng, and Peng Xi. Fast super-resolution imaging with ultra-high labeling density achieved by joint tagging super-resolution optical fluctuation imaging. Scientific Reports, 5(1), February 2015.

[31] Thomas Dertinger, Mike Heilemann, Robert Vogel, Markus Sauer, and Shimon Weiss. Super-resolution optical fluctuation imaging with organic dyes. Angewandte Chemie International Edition, 49(49):9441–9443, October 2010.

[32] P. Dedecker, G. C. H. Mo, T. Dertinger, and J. Zhang. Widely accessible method for superresolution fluorescence imaging of living systems. Proceedings of the National Academy of Sciences, 109(27):10909–10914, June 2012.

[33] Xi Zhang, Xuanze Chen, Zhiping Zeng, Mingshu Zhang, Yujie Sun, Peng Xi, Jianxin Peng, and Pingyong Xu. Development of a reversibly switchable fluorescent protein for super-resolution optical fluctuation imaging (SOFI). ACS Nano, 9(3):2659–2667, March 2015.

[34] Benjamien Moeyaert and Peter Dedecker. pcSOFI as a smart label-based superresolution microscopy technique. In Methods in Molecular Biology, pages 261–276. Springer New York, 2014.

[35] Gary C H Mo, Brian Ross, Fabian Hertel, Premashis Manna, Xinxing Yang, Eric Greenwald, Chris Booth, Ashlee M Plummer, Brian Tenner, Zan Chen, Yuxiao Wang, Eileen J Kennedy, Philip A Cole, Karen G Fleming, Amy Palmer, Ralph Jimenez, Jie Xiao, Peter Dedecker, and Jin Zhang. Genetically encoded biosensors for visualizing live-cell biochemical activity at super-resolution. Nature Methods, 14(4):427–434, March 2017.

[36] Seamus J Holden, Stephan Uphoff, and Achillefs N Kapanidis. DAOSTORM: an algorithm for high- density super-resolution microscopy. Nature Methods, 8(4):279–280, April 2011.

[37] Fang Huang, Samantha L. Schwartz, Jason M. Byars, and Keith A. Lidke. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. Biomedical Optics Express, 2(5):1377–1393, April 2011.

[38] Simon Christoph Stein, Anja Huss, Ingo Gregor, and Jörg Enderlein. Stochastic optical fluctuation imaging. In Alberto Diaspro and Marc AMJ van Zandvoort, editors, Super-Resolution Imaging in Biomedicine, chapter 9, pages 173–186. CRC Press, November 2016.

[39] Stefan Geissbühler, Noelia L Bocchio, Claudio Dellagiacoma, Corinne Berclaz, Marcel Leutenegger, and Theo Lasser. Mapping molecular statistics with balanced super-resolution optical fluctuation imaging (bSOFI). Optical Nanoscopy, 1(1):4, 2012.

[40] Thomas Dertinger, Ryan Colyer, Robert Vogel, Jörg Enderlein, and Shimon Weiss. Achieving increased resolution and more pixels with superresolution optical fluctuation imaging (SOFI). Optics Express, 18(18):18875–18885, August 2010.

[41] Wim Vandenberg, Sam Duwé, Marcel Leutenegger, Benjamien Moeyaert, Bartosz Krajnik, Theo Lasser, and Peter Dedecker. Model-free uncertainty estimation in stochastical optical fluctuation imaging (SOFI) leads to a doubled temporal resolution. Biomedical Optics Express, 7(2):467–480, January 2016.

[42] Xuehua Wang, Danni Chen, Bin Yu, and Hanben Niu. Statistical precision in super-resolution optical fluctuation imaging. Applied Optics, 55(28):7911–7916, September 2016.

[43] Joachim D. Müller. Cumulant analysis in fluorescence fluctuation spectroscopy. Biophysical Journal, 86(6):3981–3992, June 2004.

[44] Simon C. Stein, Anja Huss, Dirk Hähnel, Ingo Gregor, and Jörg Enderlein. Fourier interpolation stochastic optical fluctuation imaging. Optics Express, 23(12):16154–16163, June 2015.

[45] C.E. Shannon. Communication in the presence of noise. Proceedings of the IRE, 37(1):10–21, January 1949.

[46] Marcel Leutenegger. Balanced SOFI Toolbox. http://documents.epfl.ch/users/l/le/leuteneg/www/BalancedSOFI/index.html, 2017. Last accessed 8th March 2017.

[47] Dania Kambly, Christian Flindt, and Markus Büttiker. Factorial cumulants reveal interactions in counting statistics. Physical Review B, 83(7):075432, February 2011.

[48] Michael Hirsch, Richard J. Wareham, Marisa L. Martin-Fernandez, Michael P. Hobson, and Daniel J. Rolfe. A stochastic model for electron multiplication charge-coupled devices – from theory to practice. PLoS ONE, 8(1):e53671, January 2013.

[49] Robert L. Lucke. Fourier-space properties of photon-limited noise in focal plane array data, calculated with the discrete fourier transform. Journal of the Optical Society of America A, 18(4):777–790, April 2001.

[50] Gerd Weigelt. IV triple-correlation imaging in optical astronomy. In Progress in Optics, pages 293–319. Elsevier, 1991.

[51] E Pehlemann and O von der Lühe. Technical aspects of the speckle masking phase reconstruction algorithm. Astronomy and Astrophysics, 216:337–346, 1989.

[52] Solmaz Hajmohammadi, Saeid Nooshabadi, and Jeremy P. Bos. Massive parallel processing of image reconstruction from bispectrum through turbulence. Applied Optics, 54(32):9370–9378, November 2015.

[53] Julian Meng, E. Keith Hege, Jeffrey S. Morgan, and George J. M. Aitken. Triple-correlation sub-plane reconstruction of photon-address stellar images. Journal of the Optical Society of America A, 7(7):1243–1250, July 1990.

[54] William Hadley Richardson. Bayesian-based iterative method of image restoration*. Journal of the Optical Society of America, 62(1):55–59, January 1972.

[55] Leon B. Lucy. An iterative technique for the rectification of observed distributions. The Astronomical Journal, 79:745, June 1974.

[56] Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. The computer journal, 7(2):149–154, 1964.

[57] Leon B. Lucy. Statistical limits to super resolution. Astronomy and Astrophysics, 261(2):706–710, 1992.

[58] Leon B. Lucy. Resolution limits for deconvolved images. The Astronomical Journal, 104:1260–1265, September 1992.

[59] M. Shahram and P. Milanfar. Imaging below the diffraction limit: A statistical analysis. IEEE Transactions on Image Processing, 13(5):677–689, May 2004.

[60] M. Arigovindan, J. C. Fung, D. Elnatan, V. Mennella, Y.-H. M. Chan, M. Pollard, E. Branlund, J. W. Sedat, and D. A. Agard. High-resolution restoration of 3d structures from widefield images with extreme low signal-to-noise-ratio. Proceedings of the National Academy of Sciences, 110(43):17344–17349, October 2013.

[61] Sheila N. Baker and Gary A. Baker. Luminescent carbon nanodots: Emergent nanolights. Angewandte Chemie International Edition, 49(38):6726–6744, August 2010.

[62] Godefroy Leménager, Elisa De Luca, Ya-Ping Sun, and Pier Paolo Pompa. Super-resolution fluorescence imaging of biocompatible carbon dots. Nanoscale, 6(15):8617, July 2014.

[63] Syamantak Khan, Navneet Chandra Verma, Abhishek Gupta, and Chayan Kanti Nandi. Reversible photoswitching of carbon dots. Scientific Reports, 5(1), June 2015.

[64] Anna M. Chizhik, Simon Stein, Mariia O. Dekaliuk, Christopher Battle, Weixing Li, Anja Huss, Mitja Platen, Iwan A. T. Schaap, Ingo Gregor, Alexander P. Demchenko, Christoph F. Schmidt, Jörg Enderlein, and Alexey I. Chizhik. Super-resolution optical fluctuation bio-imaging with dual-color carbon nanodots. Nano Letters, 16(1):237–242, January 2016.

[65] Bo Shuang, David Cooper, J. Nick Taylor, Lydia Kisley, Jixin Chen, Wenxiao Wang, Chun Biu Li, Tamiki Komatsuzaki, and Christy F. Landes. Fast step transition and state identification (STaSI) for discrete single-molecule data analysis. The Journal of Physical Chemistry Letters, 5(18):3157–3161, September 2014.

[66] Pavel Frantsuzov, Masaru Kuno, Boldizsár Jankó, and Rudolph A. Marcus. Universal emission intermittency in quantum dots, nanorods and nanowires. Nature Physics, 4(5):519–522, July 2008.

[67] K.R. Diller and E.G. Cravalho. A cryomicroscope for the study of freezing and thawing processes in biological cells. Cryobiology, 7(4-6):191–199, November 1971.

[68] CKE Bleck, A Merz, MG Gutierrez, P Walther, J Dubochet, Benoît Zuber, and G Griffiths. Comparison of different methods for thin section em analysis of mycobacterium smegmatis. Journal of Microscopy, 237(1):23–38, January 2010.

[69] Werner Kuehlbrandt. Cryo-EM enters a new era. eLife, 3:e03678, August 2014.

[70] William E. Moerner and Michel Orrit. Illuminating single molecules in condensed matter. Science, 283(5408):1670–1676, March 1999.

[71] Rainer Kaufmann, Christoph Hagen, and Kay Grünewald. Fluorescence cryo-microscopy: current challenges and prospects. Current Opinion in Chemical Biology, 20:86–91, June 2014.

[72] Cindi L Schwartz, Vasily I Sarbash, Fazoil I Ataullakhanov, J Richard Mcintosh, and Daniela Nicastro. Cryo-fluorescence microscopy facilitates correlations between light and cryo-electron microscopy and reduces the rate of photobleaching. Journal of Microscopy, 227(2):98–109, August 2007.

[73] Rob Zondervan, Florian Kulzer, Mikhail A. Kol'chenk, and Michel Orrit. Photobleaching of rhodamine 6g in poly(vinyl alcohol) at the ensemble and single-molecule levels. The Journal of Physical Chemistry A, 108(10):1657–1665, March 2004.

[74] Weixing Li. Single Molecule Cryo-Fluorescence Microscopy. PhD thesis, Georg-August University Göttingen, 2016. http://hdl.handle.net/11858/00-1735-0000-002B-7C92-A.

[75] Weixing Li, Simon C. Stein, Ingo Gregor, and Jörg Enderlein. Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy. Optics Express, 23(3):3770–3783, February 2015.

[76] Siegfried Weisenburger, Bo Jing, Alois Renn, and Vahid Sandoghdar. Cryogenic localization of single molecules with angstrom precision. In Prabhat Verma and Alexander Egner, editors, Nanoimaging and Nanospectroscopy, page 88150D. International Society for Optics and Photonics, SPIE, September 2013.

[77] Steve Wolter, Anna Löschberger, Thorge Holm, Sarah Aufmkolk, Marie-Christine Dabauvalle, Sebastian van de Linde, and Markus Sauer. rapidSTORM: accurate, fast open-source software for localization microscopy. Nature Methods, 9(11):1040–1041, November 2012.

[78] Sarah R. Needham, Michael Hirsch, Daniel J. Rolfe, David T. Clarke, Laura C. Zanetti-Domingues, Richard Wareham, and Marisa L. Martin-Fernandez. Measuring EGFR separations on cells with ∼10 nm resolution via fluorophore localization imaging with photobleaching. PLoS ONE, 8(5):e62331, May 2013.

[79] Laura C. Zanetti-Domingues, Michael Hirsch, Christopher J. Tynan, Daniel J. Rolfe, Teodor V. Boyadzhiev, Kathrin M. Scherer, David T. Clarke, Marisa L. Martin-Fernandez, and Sarah R. Needham. Determining the geometry of oligomers of the human epidermal growth factor family on cells with 7 nm resolution. Progress in Biophysics and Molecular Biology, 118(3):139–152, September 2015.

[80] Siegfried Weisenburger, Bo Jing, Dominik Hänni, Luc Reymond, Benjamin Schuler, Alois Renn, and Vahid Sandoghdar. Cryogenic colocalization microscopy for nanometer-distance measurements. ChemPhysChem, 15(4):763–770, February 2014.

[81] Alexandros Pertsinidis, Yunxiang Zhang, and Steven Chu. Subnanometre single-molecule localization, registration and distance measurements. Nature, 466(7306):647–651, July 2010.

[82] E. A. K. Cohen and R. J. Ober. Analysis of point based image registration errors with applications in single molecule microscopy. IEEE Transactions on Signal Processing, 61(24):6291–6306, December 2013.

[83] Joerg Enderlein, Erdal Toprak, and Paul R. Selvin. Polarization effect on position accuracy of fluorophore localization. Optics Express, 14(18):8111–8120, 2006.

[84] Mikael P. Backlund, Matthew D. Lew, Adam S. Backer, Steffen J. Sahl, and W. E. Moerner. The role of molecular dipole orientation in single-molecule fluorescence microscopy and implications for super-resolution imaging. ChemPhysChem, 15(4):587–599, December 2013.

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

[86] Siegfried Weisenburger, Daniel Boening, Benjamin Schomburg, Karin Giller, Stefan Becker, Christian Griesinger, and Vahid Sandoghdar. Cryogenic optical localization provides 3d protein structure data with angstrom resolution. Nature Methods, 14(2):141–144, January 2017.

[87] Takashi Funatsu, Yoshie Harada, Makio Tokunaga, Kiwamu Saito, and Toshio Yanagida. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. Nature, 374(6522):555–559, April 1995.

[88] Achillefs N. Kapanidis and Terence Strick. Biology, one molecule at a time. Trends in Biochemical Sciences, 34(5):234–243, May 2009.

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

[90] Carlo Manzo and Maria F Garcia-Parajo. A review of progress in single particle tracking: from methods to biophysical insights. Reports on Progress in Physics, 78(12):124601, October 2015.

[91] Daniel Sage, Hagai Kirshner, Thomas Pengo, Nico Stuurman, Junhong Min, Suliana Manley, and Michael Unser. Quantitative evaluation of software packages for single-molecule localization microscopy. Nature Methods, 12(8):717–724, June 2015.

[92] Nicolas Chenouard, Ihor Smal, Fabrice de Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, Andrew R Cohen, William J Godinez, Karl Rohr, Yannis Kalaidzidis, Liang Liang, James Duncan, Hongying Shen, Yingke Xu, Klas E G Magnusson, Joakim Jaldén, Helen M Blau, Perrine Paul-Gilloteaux, Philippe Roudot, Charles Kervrann, François Waharte, Jean-Yves Tinevez, Spencer L Shorte, Joost Willemse, Katherine Celler, Gilles P van Wezel, Han-Wei Dan, Yuh-Show Tsai, Carlos Ortiz de Solórzano, Jean-Christophe Olivo-Marin, and Erik Meijering. Objective comparison of particle tracking methods. Nature Methods, 11(3):281–289, January 2014.

[93] Simon Christoph Stein and Jan Thiart. TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms. Scientific Reports, 6(1), November 2016.

[94] JP Lewis. Fast normalized cross-correlation. Vision Interface, 10(1):120–123, 1995.

[95] Martin Böhmer and Jörg Enderlein. Orientation imaging of single molecules by wide-field epifluorescence microscopy. Journal of the Optical Society of America B, 20(3):554–559, March 2003.

[96] Digambara Patra, Ingo Gregor, and Jörg Enderlein. Image analysis of defocused single-molecule images for three-dimensional molecule orientation studies. The Journal of Physical Chemistry A, 108(33):6836–6841, August 2004.

[97] I. Izeddin, J. Boulanger, V. Racine, C.G. Specht, A. Kechkar, D. Nair, A. Triller, D. Choquet, M. Dahan, and J.B. Sibarita. Wavelet analysis for single molecule localization microscopy. Optics Express, 20(3):2081–2095, January 2012.

[98] Stephen M. Anthony and Steve Granick. Image analysis with rapid and accurate two-dimensional gaussian fitting. Langmuir, 25(14):8152–8160, July 2009.

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

[100] Raghuveer Parthasarathy. Rapid, accurate particle tracking by calculation of radial symmetry centers. Nature Methods, 9(7):724–726, June 2012.

[101] Sameer Agarwal, Keir Mierle, and Others. Ceres solver. http://ceres-solver.org. Last accessed 1st May 2017.

[102] Jose Luis Blanco-Claraco. nanoflann: a c++ header-only library for nearest neighbor (nn) search wih kd-trees. https://github.com/jlblancoc/nanoflann, 2016. Last accessed 1st May 2017.

[103] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. Nature Methods, 5(8):695–702, July 2008.

[104] Martin Ovesnỳ, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M Hagen. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics, 30(16):2389–2390, May 2014.

[105] C. S. Smith, S. Stallinga, K. A. Lidke, B. Rieger, and D. Grunwald. Probability-based particle detection that enables threshold-free and robust in vivo single-molecule tracking. Molecular Biology of the Cell, 26(22):4057–4062, September 2015.

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

[107] Melanie Dannemeyer. A detailed study of axon initial segment maturation and structural organization by fluorescence microscopy. PhD thesis, Georg-August University Göttingen, 2016. http://hdl.handle.net/11858/00-1735-0000-0028-86B4-6.

[108] K. Xu, G. Zhong, and X. Zhuang. Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. Science, 339(6118):452–456, December 2012.

[109] Kerstin Weiß, Andreas Neef, Qui Van, Stefanie Kramer, Ingo Gregor, and Jörg Enderlein. Quantifying the diffusion of membrane proteins and peptides in black lipid membranes with 2-focus fluorescence correlation spectroscopy. Biophysical Journal, 105(2):455–462, July 2013.

[110] Joseph R. Lakowicz. Fluorescence anisotropy. In Principles of Fluorescence Spectroscopy, pages 353–382. Springer US, 2006.

[111] Narain Karedla, Simon C. Stein, Dirk Hähnel, Ingo Gregor, Anna Chizhik, and Jörg Enderlein. Simultaneous measurement of the three-dimensional orientation of excitation and emission dipoles. Physical Review Letters, 115(17):173002, October 2015.

[112] Nicholas P Machara, Peter M Goodwin, Jörg Enderlein, David J Semin, and Richard A Keller. Efficient detection of single molecules eluting off an optically trapped microsphere. Bioimaging, 6(1):33–42, March 1998.

[113] Shaun A. McKinlay. Markov Modulated Models for Derivatives Pricing. PhD thesis, University of Melbourne, 2009. http://www.ms.unimelb.edu.au/documents/thesis/Thesis-ShaunMcKinlay.pdf.

[114] Alexey I. Chizhik, Anna M. Chizhik, Anja Huss, Regina Jäger, and Alfred J. Meixner. Nanoscale probing of dielectric interfaces with single-molecule excitation patterns and radially polarized illumination. The Journal of Physical Chemistry Letters, 2(17):2152–2157, September 2011.
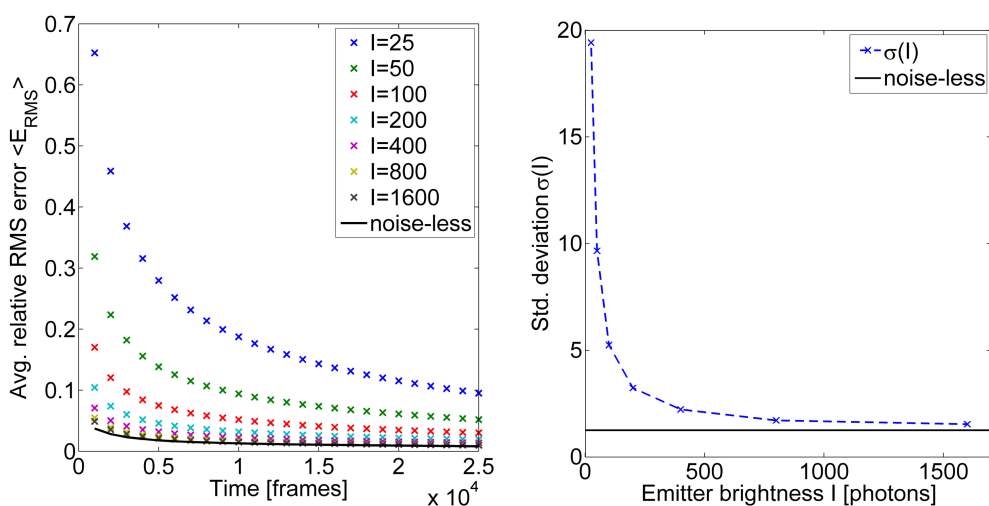
Appendices

A

Additional figures and tables

Figure A.1: Left: Pixel-averaged root mean squared error for the 2nd order cumulant for movies with different emitter intensity *I* subject to Poisson noise. The imaged structure is equivalent to the movies shown in Fig. 2.9. The blinking dynamics follow a two-state model with $\tilde{\tau}_{\mathrm{on}} = 1.67$ frames, $\tilde{\tau}_{\mathrm{off}} = 2.5$ frames. It is easy to see that convergence takes longer for dimmer emitters. Right: SOFI standard deviation against the emitter brightness. The standard error appears to be inversely proportinal to the emitter intensity. Note that the error values computed here do not compare well to the noise-free case, as the spatial correlation introduced by the PSF leads to a substantially better image than expected by the high variance of individual pixels (see end of Sec. 2.3).

	$I = 25$	$I = 50$	$I = 100$	$I = 200$	$I = 400$	$I = 800$	Noise-less
σ, 2nd order	19.42	9.66	5.25	3.24	2.23	1.72	1.28

	$I = 25$	$I = 50$	$I = 100$	$I = 200$	$I = 400$	$I = 800$	Noise-less
$T_{10\%}$, 2nd order	37700	9333	2755	1048	497	294	164

Table A.1: Pixel-averaged SOFI standard deviation σ and number of necessary frames $T_{10\%}$ to converge to the 10% error margin $E_{\mathrm{RMS}}(T_{10\%}) = 0.1$ when simulating movies as shown in Fig. 2.9 subject to Poisson noise. SOFI images were zero-lag corrected (see Sec. 2.5) before calculating the error. The corresponding convergence curves are shown in Fig. A.1. Note that the error values computed here don't compare well to the noise-free case, as the spatial correlation introduced by the PSF leads to a substantially better image than expected by the high variance of individual pixels (see end of Sec. 2.3). Thus, the true image quality at a given number of frames is much better than expected by the error value given here.

B

Computation of joint cumulants

The n-th order joint cumulant $K_n(X_1, X_2 \ldots X_n)$ of n random variables X_i is a function that it is nonzero only if all of them are pairwise correlated. The formulas for the joint cumulants can be computed using an easy algorithm which is explained here. Formulas for ordinary cumulants can be extracted from joint cumulants by setting all arguments equal $X = X_1 = \cdots = X_n$.

We quickly repeat the computation of the 2nd order joint cumulant from Section 2.2. Joint cumulants are derived using the joint moment generating function c, which depends on the sum of all random variables X_i with individual prefactors t_i. For two variables, the joint moment generating function is

$$c(t_1, t_2) = \langle \exp(t_1 X_1 + t_2 X_2) \rangle. \tag{B.1}$$

Its logarithm is the joint cumulant generating function:

$$
\begin{aligned}
g(t_1, t_2) &= \log c(t_1, t_2) \\
&= \log \langle \exp(t_1 X_1 + t_2 X_2) \rangle \\
&= \log \langle \exp(t_1 X_1) \cdot \exp(t_2 X_2) \rangle \\
&= \log \left(\sum_{n,k=0}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n X_2^k \rangle \right) \\
&= \log \left(1 + \underbrace{\sum_{n=1}^{\infty} \frac{t_1^n}{n!} \langle X_1^n \rangle + \sum_{k=1}^{\infty} \frac{t_2^k}{k!} \langle X_2^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n X_2^k \rangle}_{\hat{=}\, x} \right)
\end{aligned}
\tag{B.2}
$$

The joint cumulant is composed of all terms proportional to $t_1 \cdot t_2 \cdot \ldots t_n$ in the power series of this expression. The power series of the logarithm is $\log(1 + x) \approx x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots$. Note that in the two-variable-case, terms proportional to $t_1 \cdot t_2$ can only exist up to the second order in the expansion, which is given by:

$$
\log c(t_1, t_2) \approx \sum_{n=1}^{\infty} \frac{t_1^n}{n!} \langle X_1^n \rangle + \sum_{k=1}^{\infty} \frac{t_2^k}{k!} \langle X_2^k \rangle + \underbrace{\sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n X_2^k \rangle}
$$
$$
- \frac{1}{2} \left[\sum_{n,k=1}^{\infty} \frac{t_1^{n+k}}{n!k!} \langle X_1^n \rangle \langle X_1^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_2^{n+k}}{n!k!} \langle X_2^n \rangle \langle X_2^k \rangle + \sum_{n,k,l,i=1}^{\infty} \frac{t_1^{n+l} t_2^{k+i}}{n!k!l!i!} \langle X_1^n X_2^k \rangle \langle X_1^l X_2^i \rangle \right.
$$
$$
\left. + 2 \left(\sum_{n,k,l=1}^{\infty} \frac{t_1^{n+l} t_2^k}{n!k!l!} \langle X_1^n X_2^k \rangle \langle X_1^l \rangle + \sum_{n,k,l=1}^{\infty} \frac{t_1^n t_2^{k+l}}{n!k!l!} \langle X_1^n X_2^k \rangle \langle X_2^l \rangle + \underbrace{\sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n \rangle \langle X_2^k \rangle}_{} \right) \right]
$$
$$
+ \frac{1}{3} \ldots
\tag{B.3}
$$

Only the underbraced parts contain terms proportional to $t_1 t_2$. Collecting these, we get for the joint cumulant of both variables:

$$
K_2(X_1, X_2) = \langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle
\tag{B.4}
$$

Settting $X_1 = X_2 = X$ yields

$$
K_2(X, X) = \langle X^2 \rangle - \langle X \rangle^2 = \kappa_2(X),
\tag{B.5}
$$

which is the second order cumulant for one variable, as calculated before. Computing the joint cumulant in this way quickly gets messy, however. For example, for three variables we get:

$$
\log \langle e^{t_1 X_1 + t_2 X_2 + t_3 X_3} \rangle = \log \left(\sum_{n,k,l=0}^{\infty} \frac{t_1^n t_2^k t_3^l}{n!k!l!} \langle X_1^n X_2^k X_3^l \rangle \right)
$$
$$
= \log \left(1 + \sum_{n=1}^{\infty} \frac{t_1^n}{n!} \langle X_1^n \rangle + \sum_{n=1}^{\infty} \frac{t_2^n}{n!} \langle X_2^n \rangle + \sum_{n=1}^{\infty} \frac{t_3^n}{n!} \langle X_3^n \rangle \right.
$$
$$
+ \sum_{n,k=1}^{\infty} \frac{t_1^n t_2^k}{n!k!} \langle X_1^n X_2^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_1^n t_3^k}{n!k!} \langle X_1^n X_3^k \rangle + \sum_{n,k=1}^{\infty} \frac{t_2^n t_3^k}{n!k!} \langle X_2^n X_3^k \rangle
$$
$$
\left. + \sum_{n,k,l=1}^{\infty} \frac{t_1^n t_2^k t_3^l}{n!k!l!} \langle X_1^n X_2^k X_3^l \rangle \right)
\tag{B.6}
$$

For two variables only terms in the Taylor expansion up to the second order (x^2) could contribute to the cumulant. For three variables, we must do this calculation up to the third order, which contributes the part proportional to $\langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle$ coming from the multiplication of the first three sums in the

equation above. Doing this, we get for the joint cumulant of three variables:

$$
\begin{aligned}
K_3 & (X_1 X_2 X_3) \\
&= \langle X_1 X_2 X_3 \rangle - \frac{1}{2} \cdot 2 \Big(\langle X_1 X_2 \rangle \langle X_3 \rangle + \langle X_1 X_3 \rangle \langle X_2 \rangle + \langle X_2 X_3 \rangle \langle X_1 \rangle \Big) + \frac{1}{3} \cdot 6 \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle \\
&= \langle X_1 X_2 X_3 \rangle - \langle X_1 X_2 \rangle \langle X_3 \rangle - \langle X_1 X_3 \rangle \langle X_2 \rangle - \langle X_2 X_3 \rangle \langle X_1 \rangle + 2 \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle
\end{aligned}
$$

Looking closely we can spot the pattern for constructing the cumulant. We search for terms proportional to $t_1 \cdot t_2 \cdot \ldots t_n$, which implies that every X_i must occur exactly once in each of these terms. The step $\log \sum[\ldots] \to \log(1 + \sum[\ldots])$ in equation (B.6) yields sums containing averages of all possible combinations of the X_i. This is similar to the power set

$$
\mathcal{P}(\{X_1, X_2, X_3\}) = \{\{\}, \{X_1\}, \{X_2\}, \{X_3\}, \{X_1, X_2\}, \{X_1, X_3\}, \{X_2, X_3\}, \{X_1, X_2, X_3\}\}, \quad \text{(B.7)}
$$

where each subset corresponds to an average $\langle \bullet \rangle$. These averages are the building blocks of the cumulant and must be combined by multiplication to assure that every X_i occurs once. This multiplication happens when inserting the sum of all averages x in the Taylor expansion $\log(1+x) \approx x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots$. The exponent n of x^n tells us, how many elements of the power set we are allowed to combine. Example:

Combinations containing all X_i:

$$
x \to \text{One element} \to \langle X_1 X_2 X_3 \rangle
$$
$$
x^2 \to \text{Two elements} \to \langle X_1 X_2 \rangle \langle X_3 \rangle + \langle X_1 X_3 \rangle \langle X_2 \rangle + \langle X_2 X_3 \rangle \langle X_1 \rangle
$$
$$
x^3 \to \text{Three elements} \to \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle
$$

The prefactor, i.e. how many times a combination of n elements exist, is given by the multinomial coefficient

$$
\binom{n}{k_1, k_2, \ldots, k_m} = \frac{n!}{k_1! k_2! \ldots k_m!}, \tag{B.8}
$$

which gives the number of occurrences of $x_1^{k_1} \cdot x_2^{k_2} \cdots \cdot x_m^{k_m}$ if one expands the term $x = (x_1 + x_2 + \cdots + x_m)^n$. In our simple case higher powers of elements are not allowed, i.e. $k_i \in \{0, 1\} \, \forall i$ and the multinomial coefficient is simply reduced to $n!$. Following this we get:

$$
x \to \text{One element} \to 1! \cdot \langle X_1 X_2 X_3 \rangle =: 1! \cdot M_1(X_1 X_2 X_3)
$$
$$
x^2 \to \text{Two elements} \to 2! \cdot \big(\langle X_1 X_2 \rangle \langle X_3 \rangle + \langle X_1 X_3 \rangle \langle X_2 \rangle + \langle X_2 X_3 \rangle \langle X_1 \rangle \big) =: 2! \cdot M_2(X_1 X_2 X_3)
$$
$$
x^3 \to \text{Three elements} \to 3! \cdot \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle =: 3! \cdot M_3(X_1, X_2, X_3),
$$

where M_j is the sum of all possible combinations of j elements where each X_i occurs exactly once. An additional prefactor with alternating sign comes from the series expansion of the logarithm $\log(1+x) \approx$

$x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \ldots$ itself. Combining everything, the cumulant of n random variables is given by:

$$K_n(X_1, \ldots, X_n) = \sum_{i=1}^{n} (-1)^{i-1} \frac{1}{i} \cdot i! \cdot M_i(X_1, \ldots, X_n) \tag{B.9}$$

$$= \sum_{i=1}^{n} (-1)^{i-1} \cdot (i-1)! \cdot M_i(X_1, \ldots, X_n) \tag{B.10}$$

As an example, we can compute the cumulant of four signals a, b, c, d (for clarity we use these instead of indices). First we build the sum of all possible averages:

$$\langle a \rangle + \langle b \rangle + \langle c \rangle + \langle d \rangle \qquad \binom{4}{1} = 4 \text{ possible combinations}$$

$$+ \langle ab \rangle + \langle ac \rangle + \langle ad \rangle + \langle bc \rangle + \langle bd \rangle + \langle cd \rangle \qquad \binom{4}{2} = 6 \text{ possible combinations}$$

$$+ \langle abc \rangle + \langle abd \rangle + \langle acd \rangle + \langle bcd \rangle \qquad \binom{4}{3} = 4 \text{ possible combinations}$$

$$+ \langle abcd \rangle \qquad \binom{4}{4} = 1 \text{ possible combination}$$

Like before, we combine one/two/three/four elements under the constraint that each X_i must show up exactly once in the result:

$x \rightarrow$ One element $\rightarrow 1! \cdot \langle abcd \rangle$

$x^2 \rightarrow$ Two elements $\rightarrow 2! \cdot \Big(\langle a \rangle \langle bcd \rangle + \langle b \rangle \langle acd \rangle + \langle c \rangle \langle abd \rangle + \langle d \rangle \langle abc \rangle$

$$+ \langle ab \rangle \langle cd \rangle + \langle ac \rangle \langle bd \rangle + \langle ad \rangle \langle bc \rangle \Big)$$

$x^3 \rightarrow$ Three elements $\rightarrow 3! \cdot \Big(\langle a \rangle \langle b \rangle \langle cd \rangle + \langle a \rangle \langle c \rangle \langle bd \rangle + \langle a \rangle \langle d \rangle \langle cb \rangle$

$$+ \langle b \rangle \langle c \rangle \langle ad \rangle + \langle b \rangle \langle d \rangle \langle ac \rangle + \langle c \rangle \langle d \rangle \langle ab \rangle \Big)$$

$x^4 \rightarrow$ Four elements $\rightarrow 4! \cdot \langle a \rangle \langle b \rangle \langle c \rangle \langle d \rangle$,

Combining this with the prefactor from the series expansion and summing everything up, the joint cu-

mulant of four signals is given by:

$$K_4(a, b, c, d) = 1! \cdot \langle abcd \rangle$$
$$- \frac{1}{2} \cdot 2! \cdot \Big(\langle a \rangle \langle bcd \rangle + \langle b \rangle \langle acd \rangle + \langle c \rangle \langle abd \rangle + \langle d \rangle \langle abc \rangle$$
$$+ \langle ab \rangle \langle cd \rangle + \langle ac \rangle \langle bd \rangle + \langle ad \rangle \langle bc \rangle \Big)$$
$$+ \frac{1}{3} \cdot 3! \cdot \Big(\langle a \rangle \langle b \rangle \langle cd \rangle + \langle a \rangle \langle c \rangle \langle bd \rangle + \langle a \rangle \langle d \rangle \langle cb \rangle$$
$$+ \langle b \rangle \langle c \rangle \langle ad \rangle + \langle b \rangle \langle d \rangle \langle ac \rangle + \langle c \rangle \langle d \rangle \langle ab \rangle \Big)$$
$$- \frac{1}{4} \cdot 4! \cdot \langle a \rangle \langle b \rangle \langle c \rangle \langle d \rangle$$
(B.11)

Setting $a = b = c = d = X$, we get the 4th-order ordinary cumulant:

$$\kappa_4(X) = K_4(X, X, X, X) = \langle X^4 \rangle - 4 \langle X \rangle \langle X^3 \rangle - 3 \langle X^2 \rangle^2 + 12 \langle X^2 \rangle \langle X \rangle^2 - 6 \langle X \rangle^4 \qquad \text{(B.12)}$$

The rules presented here can be implemented into a simple computer program that computes the formulas for arbitrary orders if required.

Theoretical cumulants of a two-state model

S_{on}

k_- k_+

S_{off}
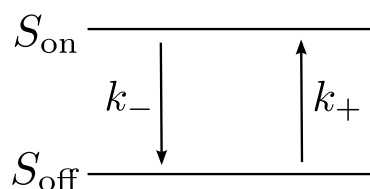
Figure C.1: Simple two state model of an excitable system with transition rates k_+ and k_-.

A simple but in many cases sufficiently accurate model for a fluorophore assumes that the fluorophore can take one of two states: in the "on"-state the fluorophore emits photons, while in the "off"-state it stays completely dark. For every time interval Δt there is a fixed chance that the fluorophore switches its state, given by the transition rates k_+ and k_- (see Fig. C.1). To derive the cumulants, we will make use of the ergodicity of our system, i.e. that the time average $\langle F \rangle_t$ is equal to the ensemble average $\langle F \rangle_E$. Looking at an ensemble of molecules, let S_{on} and S_{off} be the ratio of molecules occupying the on- and off-state respectively. As molecules are neither generated nor destroyed, their number is conserved giving $S_{\mathrm{on}} + S_{\mathrm{off}} = 1$. The behavior of an ensemble of fluorescing two-state molecules can be encoded in a very simple differential equation. We will now solve that equation for the continuous time and the discrete time case. In the latter, the fluorophore state is frozen during the time interval given by the discretization and can only change state between intervals. This is useful, because the discrete dynamics are simpler to simulate using a computer.

C.1 Continuous time differential equation

The change of the occupation ratio of the excited state S_{on} is described by

$$\frac{dS_{on}}{dt} = -k_- S_{on} + k_+ S_{off} \tag{C.1}$$

Using the conservation of molecules $S_{on} + S_{off} = 1 \Rightarrow S_{off} = 1 - S_{on}$ this becomes:

$$\frac{dS_{on}}{dt} = k_+ - (k_+ + k_-)S_{on} \tag{C.2}$$

Solving this inhomogeneous differential equation combining the homogeneous solution with an inhomogeneous one:

homogeneous solution:

$$\frac{dS_{on}}{dt} + (k_+ + k_-)S_{on} = 0 \Rightarrow S_{on}(t) = A \cdot e^{-(k_+ + k_-)t} \tag{C.3}$$

inhomogeneous solution:

$$\frac{dS_{on}}{dt} = 0 = k_+ - (k_+ + k_-)S_{on} \Rightarrow S_{on} = \frac{k_+}{k_+ + k_-} \tag{C.4}$$

general solution:

$$S_{on}(t) = \frac{k_+}{k_+ + k_-} + A \cdot e^{-(k_+ + k_-)t} \tag{C.5}$$

Using the initial condition that every molecule is excited, i.e. $S_{on}(t = 0) = 1$, the occupation ratio of the on-state over time is given by:

$$S_{on}(t) = \frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} e^{-(k_+ + k_-)t} \tag{C.6}$$

C.2 Discrete time differential equation

The discrete time case starts similar to the continuous one, but can not be solved by integration. One iteration in time is modeled by:

$$S_{on}^{t+1} = S_{on}^t + k_+ S_{off}^t - k_- S_{on}^t = k_+ + S_{on}^t(1 - (k_+ + k_-)) \tag{C.7}$$

$$\equiv b + S_{on} \cdot a \quad \text{with} \quad b = k_+, \ a = [1 - (k_+ + k_-)] \tag{C.8}$$

An inhomogeneous discrete differential equation of the type $x_n = a \cdot x_{n-1} + b$ can be solved by recursion:

$$x_n = a \cdot x_{n-1} + b \tag{C.9}$$

$$= a^2 \cdot x_{n-2} + (a+1) \cdot b \tag{C.10}$$

$$= a^3 \cdot x_{n-3} + (a^2 + a + 1) \cdot b \tag{C.11}$$

$$= a^n \cdot x_0 + \underbrace{(a^{n-1} + a^{n-2} + \cdots + a + 1)} \cdot b \tag{C.12}$$

If $|a| < 1$ holds, the underbraced part is a geometric series with the known solution

$$\sum_{i=0}^{n-1} a^i = \begin{cases} n & , a = 1 \\ \frac{a^n - 1}{a - 1} & , |a| < 1 \end{cases} \tag{C.13}$$

giving us the solution to our differential equation

$$x_n = \begin{cases} x_0 + n \cdot b & , a = 1 \\ a^n \cdot x_0 + \frac{a^n - 1}{a - 1} \cdot b & , a \neq 1 \end{cases} \tag{C.14}$$

Replacing the general notation with our specific problem, we get:

$$S_{on}^t = [1 - (k_+ + k_-)]^t S_{on}^0 + \frac{[1 - (k_+ + k_-)]^t - 1}{[1 - (k_+ + k_-)] - 1} \cdot k_+ \tag{C.15}$$

$$= [1 - (k_+ + k_-)]^t S_{on}^0 - \frac{k_+ \cdot (1 - [k_+ + k_-])^t}{k_+ + k_-} + \frac{k_+}{k_+ + k_-} \tag{C.16}$$

Using the initial condition that every molecule is excited $S_{on}^0 = 1$ gives

$$S_{on}^t = [1 - (k_+ + k_-)]^t \cdot \left(1 - \frac{k_+}{k_+ + k_-} \right) + \frac{k_+}{k_+ + k_-} \tag{C.17}$$

$$= \frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} \cdot [1 - (k_+ + k_-)]^t \tag{C.18}$$

Comparison between the continuous and discrete time solutions reveals we can switch between them replacing

$$\textbf{continuous time} \ \exp(-(k_+ + k_-)t) \ \leftrightarrow \ [1 - (k_+ + k_-)]^{t_i} \ \textbf{discrete time} \tag{C.19}$$
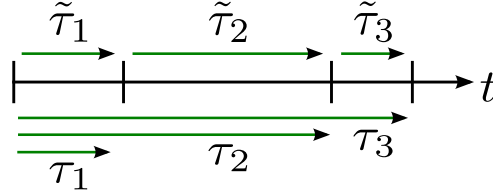
Figure C.2: Relative time-lags $\tilde{\tau}_i$ refer to points in time relative to each other, while the absolute time-lags τ_i are measured relative to the same starting point.

As an example we compute the theoretical joint cumulants of 2nd and 3rd order for a single molecule with brightness ε in the on-state and instantaneous brightness $F(t)$. Because the cumulant is additive, the cumulant of N molecules is simply N times the one-molecule cumulant. The calculation is done using the continuous time formula, but can be easily converted to discrete time as shown above. When calculating the probabilities involved in the problem, we make use of relative time-lags $\tilde{\tau}_i$ as shown in Fig. C.2 rather than absolute ones, as they are easier to handle in this context.

Assuming ergodicity, the solution for the behavior of the ensemble is (statistically) valid for the over-time behavior of a single emitter. Equation (C.5) reveals that probability to find a molecule in the excited state at an arbitrary point in time after equilibration is given by

$$P_{on}(t) = \frac{k_+}{k_+ + k_-} \Rightarrow \langle F(t) \rangle = \varepsilon \frac{k_+}{k_+ + k_-} \ , \tag{C.20}$$

which is the average ratio of molecules in the excited state in an ensemble and thus the average time spent in the excited state for a single molecule. Similarly, the conditional probability to find the molecule in the excited state at time $t + \tilde{\tau}$ given that it was excited at t is given by Eq. (C.6):

$$P_{on}(t + \tilde{\tau}|\text{on at } t) = \frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} e^{-(k_+ + k_-)\tilde{\tau}} \tag{C.21}$$

The joint cumulant of the time trace $F(t)$ and a time-shifted signal $F(t + \tilde{\tau})$ is given by

$$K_2(\tilde{\tau}) := K_2\left[F(t), F(t + \tilde{\tau})\right] = \langle F(t) F(t + \tilde{\tau}) \rangle - \langle F(t) \rangle^2 \tag{C.22}$$

Note that $F(t) F(t + \tilde{\tau})$ is only nonzero if the molecule is excited at both times and zero otherwise (assuming that the dark state is completely dark). The average is computed by multiplying the value of $F(t) F(t + \tilde{\tau})$ in this case (ε^2) with the probability that the situation occurs. We can denote this probability by $P_{on}(t) \cdot P_{on}(t + \tilde{\tau}|\text{on at } t)$. Fortunately, we already know their values and can directly

compute the cumulant by plugging them in:

$$K_2(\tilde{\tau}) = \langle F(t)\, F(t + \tilde{\tau})\rangle - \langle F(t)\rangle^2 \tag{C.23}$$

$$= \varepsilon^2 \frac{k_+}{k_+ + k_-} \cdot \left(\frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-}\, e^{-(k_+ + k_-)\tilde{\tau}}\right) - \left(\varepsilon \frac{k_+}{k_+ + k_-}\right)^2 \tag{C.24}$$

$$= \frac{\varepsilon^2}{(k_+ + k_-)^2} \cdot \left(k_+^2 + k_+ k_- e^{-(k_+ + k_-)\tilde{\tau}}\right) - \frac{\varepsilon^2}{(k_+ + k_-)^2} \cdot k_+^2 \tag{C.25}$$

$$\Rightarrow K_2(\tilde{\tau}) = K_2(\tau) = \frac{\varepsilon^2}{(k_+ + k_-)^2} \cdot k_+ k_- \cdot e^{-(k_+ + k_-)\tilde{\tau}} \tag{C.26}$$

Following the same scheme, we can compute the third order joint cumulant. The joint cumulant of three signals is given by

$$\langle \delta F \delta F_1 \delta F_2\rangle = \tag{C.27}$$

$$\langle FF_1F_2\rangle - \langle F\rangle\,\langle F_1F_2\rangle - \langle F_1\rangle\,\langle FF_2\rangle - \langle F_2\rangle\,\langle FF_1\rangle + 2\,\langle F\rangle\,\langle F_1\rangle\,\langle F_2\rangle \tag{C.28}$$

with $F = F(t)$, $F_1 = F(t + \tilde{\tau}_1)$, $F_2 = F(t + \tau_2) = F(t + \tilde{\tau}_1 + \tilde{\tau}_2)$ in our case. Note that the conditional probability if the molecule is still excited depends only on the difference between the two corresponding points in time. For clarity we divide the formula into three parts:

$$K_3(\tilde{\tau}_1, \tilde{\tau}_2) =$$

$$\textbf{(1)}\quad \langle F(t)\, F(t + \tilde{\tau}_1)\, F(t + \tilde{\tau}_1 + \tilde{\tau}_2)\rangle$$

$$-\textbf{(2)}\quad \langle F(t)\rangle \cdot \Big(\,\langle F(t)\, F(t + \tilde{\tau}_1)\rangle + \langle F(t)\, F(t + \tilde{\tau}_1 + \tilde{\tau}_2)\rangle + \langle F(t + \tilde{\tau}_1)\, F(t + \tilde{\tau}_1 + \tilde{\tau}_2)\rangle\,\Big)$$

$$+\textbf{(3)}\quad 2 \cdot \langle F(t)\rangle^3$$

For the first part we get

$$\textbf{(1)} = \varepsilon^3 \cdot \frac{k_+}{k_+ + k_-}\left(\frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} \cdot e^{-(k_+ + k_-)\tilde{\tau}_1}\right)\left(\frac{k_+}{k_+ + k_-} + \frac{k_-}{k_+ + k_-} \cdot e^{-(k_+ + k_-)\tilde{\tau}_2}\right)$$

$$= \frac{\varepsilon^3}{(k_+ + k_-)^3}\left(k_+^2 + k_+ k_- \cdot e^{-(k_+ + k_-)\tilde{\tau}_1}\right)\left(k_+ + k_- \cdot e^{-(k_+ + k_-)\tilde{\tau}_2}\right)$$

$$= \frac{\varepsilon^3}{(k_+ + k_-)^3}\left(k_+^3 + k_+^2 k_- \cdot e^{-(k_+ + k_-)\tilde{\tau}_1} + k_+^2 k_- \cdot e^{-(k_+ + k_-)\tilde{\tau}_2} + k_+ k_-^2 \cdot e^{-(k_+ + k_-)(\tilde{\tau}_1 + \tilde{\tau}_2)}\right)$$

For **(2)** we use the solution for terms of the type $\langle F(t)F(t + \tilde{\tau})\rangle$ known from the second order cumulant

$$\textbf{(2)} = \frac{\varepsilon^3}{(k_+ + k_-)^3}\left(3k_+^3 + k_+^2 k_- \cdot \left[e^{-(k_+ + k_-)\tilde{\tau}_1} + e^{-(k_+ + k_-)(\tilde{\tau}_1 + \tilde{\tau}_2)} + e^{-(k_+ + k_-)\tilde{\tau}_2}\right]\right)$$

$$\textbf{(3)} = 2\frac{\varepsilon^3}{(k_+ + k_-)^3} \cdot k_+^3$$

Assembling all parts, we get the overall solution

$$K_3(\tilde{\tau}_1, \tilde{\tau}_2) = (\mathbf{1}) - (\mathbf{2}) + (\mathbf{3}) = \frac{\varepsilon^3}{(k_+ + k_-)^3} \left(k_+ k_-{}^2 - k_+{}^2 k_-\right) e^{-(k_+ + k_-)(\tilde{\tau}_1 + \tilde{\tau}_2)} \qquad (\text{C.29})$$

Of course, this solution can also be expressed in terms of absolute time-lags using the variable transformation $\tilde{\tau}_1 = \tau_1, \tilde{\tau}_2 = \tau_2 - \tau_1, \tau_2 > \tau_1$, which depends only on the longest time-lag:

$$K_3(\tau_2) = (\mathbf{1}) - (\mathbf{2}) + (\mathbf{3}) = \frac{\varepsilon^3}{(k_+ + k_-)^3} \left(k_+ k_-{}^2 - k_+{}^2 k_-\right) e^{-(k_+ + k_-)\tau_2} \qquad (\text{C.30})$$

Appendix E lists the theoretical cumulants of the two-state model up to the 4th order.

D

Bias and error of the 2nd order ordinary cumulant

SOFI is based on the estimation of higher order statistical quantities called cumulants of time-correlated time series. Section 2.3 shows that 2nd order SOFI is most widely applicable as the estimation of the 2nd order cumulant (the *variance*) needs least amount of frames to converge and is most robust to high emitter density compared to higher orders. To understand this analytically, we derive the expected value of the variance estimator and its error for a time-correlated discrete stationary time series $\{x_i\}$ of T values. For ease of notation, we define $\delta x_i := x_i - \langle x_i \rangle$ and avoid indexing where it is not necessary due to stationarity, e.g. $\langle x_i \rangle = \langle x_j \rangle = \langle x \rangle$. As a reminder, κ_n denote ordinary cumulants of n-th order and K_n the joint cumulant of n signals.

D.1   Bias of the variance estimator for correlated time series

The usual estimator for the 2nd order cumulant / variance κ_2 of the distribution of x is

$$\langle \tilde{\kappa}_2 \rangle = \left\langle \frac{1}{T-1} \sum_{i=1}^{T} (x - \bar{x})^2 \right\rangle = \left\langle \frac{1}{T-1} \sum_{i=1}^{T} \left(x - \frac{1}{T} \sum_{j=1}^{T} x_j \right)^2 \right\rangle, \qquad \text{(D.1)}$$

which is unbiased for uncorrelated time series. To compute its value for correlated time series, a different expression is helpful:

$$\left\langle \frac{1}{T-1} \sum_{i=1}^{T} (x - \bar{x})^2 \right\rangle = \left\langle \frac{1}{(T-1)T} \sum_{i<j}^{T} (x_i - x_j)^2 \right\rangle \qquad \text{(D.2)}$$

159

$$\left\langle \frac{1}{(T-1)T} \sum_{i<j}^{T} (x_i - x_j)^2 \right\rangle = \left\langle \frac{1}{(T-1)T} \sum_{i<j}^{T} x_i^2 - 2x_i x_j + x_j^2 \right\rangle \tag{D.3}$$

$$= \frac{1}{(T-1)T} \sum_{i<j}^{T} \langle x_i^2 \rangle + \langle x_j^2 \rangle - 2\langle x_i x_j \rangle \tag{D.4}$$

$$= \langle x^2 \rangle - \frac{2}{(T-1)T} \sum_{i<j}^{T} \langle x_i x_j \rangle \tag{D.5}$$

For the last line we used that $\langle x_i^2 \rangle = \langle x_j^2 \rangle = \langle x^2 \rangle$ and that the number of elements of the double sum is $(T-1)T$. Adding $0 = -\langle x \rangle^2 + \langle x \rangle^2$ yields

$$= \langle x^2 \rangle - \langle x \rangle^2 - \frac{2}{(T-1)T} \left[\sum_{i<j}^{T} \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \right] \tag{D.6}$$

$$= \langle \delta x^2 \rangle - \frac{2}{(T-1)T} \sum_{i<j}^{T} \langle \delta x_i \delta x_j \rangle \tag{D.7}$$

$$= \langle \delta x^2 \rangle \cdot \left[1 - \frac{2}{(T-1)T} \sum_{i<j}^{T} \frac{\langle \delta x_i \delta x_j \rangle}{\langle \delta x^2 \rangle} \right] \tag{D.8}$$

Using stationarity, we can convert the double sum into:

$$= \langle \delta x^2 \rangle \cdot \left[1 - \frac{2}{(T-1)T} \sum_{i=0}^{T-1} \sum_{\tau=1}^{T-1-i} \frac{\langle \delta x_0 \delta x_\tau \rangle}{\langle \delta x^2 \rangle} \right] \tag{D.9}$$

$$\langle \tilde{\kappa}_2 \rangle = \kappa_2 \cdot \left[1 - \frac{2}{(T-1)T} \sum_{i=0}^{T-1} \sum_{\tau=1}^{T-1-i} G(\tau) \right] \tag{D.10}$$

where we also shifted the indices by 1, $\kappa_2 = \langle (x - \langle x \rangle)^2 \rangle$ is the 2nd order cumulant of the time series and $G(\tau)$ its autocorrelation function. The estimator is biased for finite T. It is noteworthy however, that an estimator using the true mean value $\langle x \rangle$ instead of the sample mean $\bar{x}$ to calculate the variance is unbiased even for correlated time series:

$$\left\langle \frac{1}{T} \sum_{i=1}^{T} (x_i - \langle x \rangle)^2 \right\rangle = \frac{1}{T} \sum_{i=1}^{T} \langle x^2 \rangle - 2\langle x \rangle^2 + \langle x \rangle^2 = \langle x^2 \rangle - \langle x \rangle^2 = \kappa_2 \tag{D.11}$$

Using the true mean for the estimator requires the prefactor T^{-1} instead of $(T-1)^{-1}$ for the estimator to be unbiased. As another remark, deriving the corresponding expression for a continuous signal $x(t)$ is analogous, giving

$$\left\langle \frac{1}{T} \int_0^T \left(x(t) - \frac{1}{T} \int_0^T x(t') \, dt' \right)^2 dt \right\rangle = \kappa_2 \cdot \left[1 - \frac{2}{T^2} \int_0^T dt \int_0^{T-t} G(\tau) \, d\tau \right] \tag{D.12}$$

The auto-correlation of a discrete two-state model is $G(\tau) = [1 - (k_+ + k_-)]^\tau$ (see appendix C, $K_2(\tau)/K_2(0)$) and as $|[1 - (k_+ + k_-)]| < 1$ holds, the double sum in Eq. (D.10) can be solved using the known solution for the geometric series. With $a := [1 - (k_+ + k_-)]$, the final estimator for the 2nd order cumulant of a two-state model is:

$$\langle \tilde{\kappa}_2 \rangle = \kappa_2 \cdot \left[1 - \frac{2}{(T-1)T} \frac{a(a^T - Ta + T - 1)}{(a - 1)^2} \right] \tag{D.13}$$

When processing movies using SOFI, it is important that the processed time interval T is at least a few times the typical correlation time to get an accurate value of the time average, which reduces the bias of the variance. For typical recording times and blinking dynamics, the bias can usually be neglected.

D.2 ERROR OF THE VARIANCE FOR CORRELATED TIME SERIES

We now derive the standard error for the 2nd order cumulant κ_2. We assume that the time series $\{x_i\}$ is long enough to measure the average brightness $\langle x \rangle$ with low error, i.e. that the recording time is longer than the typical blinking time-scale and the estimator for the cumulant is unbiased. This is equivalent to using the true-mean estimator

$$\langle \hat{\kappa}_2 \rangle = \left\langle \frac{1}{T} \sum_{i=1}^{T} (x_i - \langle x \rangle)^2 \right\rangle = \kappa_2 \tag{D.14}$$

for estimating the cumulant. Its variance is

$$\text{Var}\,(\hat{\kappa}_2) = \langle \hat{\kappa}_2^2 \rangle - \langle \hat{\kappa}_2 \rangle^2 \tag{D.15}$$

We already know the value for the second term $\langle \hat{\kappa}_2 \rangle^2 = \kappa_2^2$, leaving only the first term to be calculated.

$$\langle \hat{\kappa}_2^2 \rangle = \left\langle \left[\frac{1}{T} \sum_{i=0}^{T} (x_i - \langle x \rangle)^2 \right]^2 \right\rangle \tag{D.16}$$

$$= \frac{1}{T^2} \left\langle \left[\sum_{i=0}^{T} \delta x_i^2 \right]^2 \right\rangle \tag{D.17}$$

$$= \frac{1}{T^2} \left\langle \sum_{i=0}^{T} \delta x_i^4 + 2 \sum_{i<j}^{T} \delta x_i^2 \delta x_j^2 \right\rangle \tag{D.18}$$

$$= \frac{1}{T} \mu_4 + \frac{2}{T^2} \sum_{i<j}^{T} \langle \delta x_i^2 \delta x_j^2 \rangle \tag{D.19}$$

In the last line we introduced the definition of central moments $\mu_n = \langle (x_i - \langle x \rangle)^n \rangle$. We need to find an expression for $\langle \delta x_i^2 \delta x_j^2 \rangle$:

$$\langle \delta x_i^2 \delta x_j^2 \rangle = \langle (x_i^2 - 2x_i \langle x \rangle + \langle x \rangle^2)(x_j^2 - 2x_j \langle x \rangle + \langle x \rangle^2) \rangle \tag{D.20}$$

$$= \langle x_i^2 x_j^2 \rangle - 2 \langle x \rangle \left(\langle x_i^2 x_j \rangle + \langle x_i x_j^2 \rangle \right) + \langle x \rangle^2 \left(\langle x_i^2 \rangle + \langle x_j^2 \rangle + 4 \langle x_i x_j \rangle \right) - 3 \langle x \rangle^4 . \tag{D.21}$$

Using time inversion symmetry $\langle x_i x_j^2 \rangle = \langle x_j x_i^2 \rangle$ and stationarity $\langle x_i^2 \rangle = \langle x_j^2 \rangle = \langle x^2 \rangle$

$$\langle \delta x_i^2 \delta x_j^2 \rangle = \langle x_i^2 x_j^2 \rangle - 4 \langle x \rangle \langle x_i^2 x_j \rangle + \langle x \rangle^2 \left(2 \langle x^2 \rangle + 4 \langle x_i x_j \rangle \right) - 3 \langle x \rangle^4 . \tag{D.22}$$

The variance is always additive for mixtures of statistically independent variables. Thus, it can always be expressed by a combination of cumulants. We add $0 = \kappa_2^2 - \kappa_2^2$ with $\kappa_2^2 = (\langle x^2 \rangle - \langle x \rangle^2)^2 = \langle x^2 \rangle^2 - 2 \langle x^2 \rangle \langle x \rangle^2 + \langle x \rangle^4$ to get

$$= \kappa_2^2 + \langle x_i^2 x_j^2 \rangle - 4 \langle x \rangle \langle x_i^2 x_j \rangle + \langle x \rangle^2 \left(4 \langle x^2 \rangle + 4 \langle x_i x_j \rangle \right) - \langle x^2 \rangle^2 - 4 \langle x \rangle^4 . \tag{D.23}$$

The fourth order joint cumulant $K_4[x_i, x_i, x_j, x_j]$ (compare appendix E)

$$K_4[x_i, x_i, x_j, x_j] = \langle x_i^2 x_j^2 \rangle - \left(4 \langle x \rangle \langle x_i^2 x_j \rangle + \langle x^2 \rangle^2 + 2 \langle x_i x_j \rangle^2 \right)$$
$$+ 2 \left(2 \langle x \rangle^2 \langle x^2 \rangle + 4 \langle x \rangle^2 \langle x_i x_j \rangle \right) - 6 \langle x \rangle^4 \tag{D.24}$$
$$= \langle x_i^2 x_j^2 \rangle - 4 \langle x \rangle \langle x_i^2 x_j \rangle - \langle x^2 \rangle^2 - 2 \langle x_i x_j \rangle^2 + \langle x \rangle^2 \left(4 \langle x^2 \rangle + 8 \langle x_i x_j \rangle \right) - 6 \langle x \rangle^4 \tag{D.25}$$

can be identified within the expression, leaving

$$\langle \delta x_i^2 \delta x_j^2 \rangle = \kappa_2^2 + K_4[x_i, x_i, x_j, x_j] - 4 \langle x \rangle^2 \langle x_i x_j \rangle + 2 \langle x_i x_j \rangle^2 + 2 \langle x \rangle^4 \tag{D.26}$$
$$= \kappa_2^2 + K_4[x_i, x_i, x_j, x_j] + 2 \left(\langle x_i x_j \rangle - \langle x \rangle^2 \right)^2 \tag{D.27}$$
$$= \kappa_2^2 + K_4[x_i, x_i, x_j, x_j] + 2 K_2^2[x_i, x_j]. \tag{D.28}$$

With the second order joint cumulant $K_2[x_i, x_j] = \langle x_i x_j \rangle - \langle x \rangle^2$. Putting this into Eq. (D.19) and remembering that the double sum has $T(T-1)/2$ terms yields

$$\langle \hat{\kappa}_2^2 \rangle = \frac{1}{T} \mu_4 + \frac{2}{T^2} \sum_{i<j}^{T} \kappa_2^2 + K_4[x_i, x_i, x_j, x_j] + 2 K_2^2[x_i, x_j] \tag{D.29}$$

$$= \frac{1}{T} \mu_4 + \kappa_2^2 - \frac{1}{T} \kappa_2^2 + \frac{2}{T^2} \sum_{i<j}^{T} K_4[x_i, x_i, x_j, x_j] + 2 K_2^2[x_i, x_j] \tag{D.30}$$

Finally, assembling everything and inserting into Eq.D.15 gives the variance

$$\text{Var}\left(\hat{\kappa}_2\right) = \left\langle \hat{\kappa}_2^2 \right\rangle - \left\langle \hat{\kappa}_2 \right\rangle^2 \tag{D.31}$$

$$= -\kappa_2^2 + \frac{1}{T}\mu_4 + \kappa_2^2 - \frac{1}{T}\kappa_2^2 + \frac{2}{T^2}\sum_{i<j}^{T} K_4[x_i, x_i, x_j, x_j] + 2K_2^2[x_i, x_j] \tag{D.32}$$

$$= \frac{1}{T}\left(\mu_4 - \kappa_2^2\right) + \frac{2}{T^2}\sum_{i<j}^{T} K_4[x_i, x_i, x_j, x_j] + 2K_2^2[x_i, x_j] \tag{D.33}$$

The fourth order ordinary cumulant and the fourth moment have the relation $\kappa_4 = \mu_4 - 3\kappa_2^2$ (compare appendix E), which yields the final formula for the variance of a time-correlated, stationary time series with time reversal symmetry:

$$\text{Var}\left(\hat{\kappa}_2\right) = \frac{1}{T}\left(\kappa_4 + 2\kappa_2^2\right) + \frac{2}{T^2}\sum_{i<j}^{T} K_4[x_i, x_i, x_j, x_j] + 2K_2^2[x_i, x_j] \tag{D.34}$$

Like we did before for the estimator of the 2nd order cumulant, we can change the expression of the double sum using stationarity and shifting the index by 1, to get

$$\text{Var}\left(\hat{\kappa}_2\right) = \frac{1}{T}\left(\kappa_4 + 2\kappa_2^2\right) + \frac{2}{T^2}\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i} K_4[x_0, x_0, x_\tau, x_\tau] + 2K_2^2[x_0, x_\tau]. \tag{D.35}$$

Taking the cumulants of a two-state model, this is equivalent to

$$\text{Var}\left(\hat{\kappa}_2\right) = \frac{1}{T}\left(\kappa_4 + 2\kappa_2^2\right) + \frac{2}{T^2}\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i} K_4(\tau_1 = 0, \tau_2 = \tau, \tau_3 = \tau) + 2K_2^2(\tau) \tag{D.36}$$

With $\kappa_2^2 = K_2^2(0) = N_{\text{em}}^2 \varepsilon^4 \cdot k_+^2 k_-^2/(k_+ + k_-)^4$ and the autocorrelation function $G(\tau) = a^\tau = [1 - (k_+ + k_-)]^\tau$ of the discrete time model, the expressions for the joint cumulants are

$$K_2^2(\tau) = N_{\text{em}}^2 \varepsilon^4 \cdot \frac{k_+^2 k_-^2}{(k_+ + k_-)^4} a^{2\tau} = \kappa_2^2 \cdot a^{2\tau} \tag{D.37}$$

$$K_4(0, \tau, \tau) = N_{\text{em}} \varepsilon^4 \cdot \frac{k_+ k_-}{(k_+ + k_-)^2} \left[(k_-^2 - 2k_+ k_- + k_+^2) \cdot a^\tau - 2k_+ k_- \cdot a^{2\tau}\right] \tag{D.38}$$

$$= \frac{\kappa_2^2}{N_{\text{em}}} \left[\frac{(k_+ - k_-)^2}{k_+ k_-} \cdot a^\tau - 2a^{2\tau}\right] \tag{D.39}$$

Inserting this into the variance gives

$$\text{Var}\left(\hat{\kappa}_2\right) = \frac{1}{T}\left(\frac{\kappa_2^2}{N_{\text{em}}}\left[\frac{(k_+ - k_-)^2}{k_+ k_-} - 2\right] + 2\kappa_2^2\right)$$

$$+ \frac{2}{T^2}\sum_{i=0}^{T-1}\sum_{\tau=1}^{T-1-i} \frac{\kappa_2^2}{N_{\text{em}}}\left[\frac{(k_+ - k_-)^2}{k_+ k_-} \cdot a^\tau\right] + 2\kappa_2^2\left(1 - \frac{1}{N_{\text{em}}}\right) \cdot a^{2\tau} \tag{D.40}$$

As $|a| < 1$ holds, the double sums can be solved using the known solution for the partial sum of the geometric series. To investigate the convergence of the estimator to the true value, it is more interesting to look at its relative deviation instead of the absolute one. The relative variance of the 2nd order cumulant estimator is

$$
\frac{\mathrm{Var}\left(\hat{\kappa}_2\left(I(t)\right)\right)}{\kappa_2^2(I(t))} = \frac{1}{T}\left(2 + \frac{1}{N_{\mathrm{em}}k_+k_-}\left[k_+^2 - 4k_+k_- + k_-^2\right]\right)
$$
$$
+ \frac{2}{T^2}\left\{\frac{1}{N_{\mathrm{em}}}\frac{(k_+ - k_-)^2}{k_+k_-}\left[\frac{a(a^T - Ta + T - 1)}{(a-1)^2}\right]\right.
$$
$$
\left. +2\cdot\left(1 - \frac{1}{N_{\mathrm{em}}}\right)\left[\frac{a(a^{2T} - Ta^2 + T - 1)}{(a^2-1)^2}\right]\right\}.
$$

(D.41)

and its standard error the square root of this expression, as usual.

E

Collection of cumulant formulas

For reference, this part of the appendix gathers all formulas for cumulants used in this thesis. This is useful to implement SOFI calculations, correct images for the influence of noise, and for analytical modeling of two-state systems.

n	Ordinary cumulant $\kappa_n\left[F(\mathbf{r}, t)\right]$
1	$\langle F \rangle$
2	$\langle \delta F^2 \rangle = \langle F^2 \rangle - \langle F \rangle^2$
3	$\langle \delta F^3 \rangle = \langle F^3 \rangle - 3\langle F \rangle\langle F^2 \rangle + 2\langle F \rangle^3$
4	$\langle \delta F^4 \rangle - 3\langle \delta F^2 \rangle^2 = \langle F^4 \rangle - 4\langle F \rangle\langle F^3 \rangle - 3\langle F^2 \rangle^2 + 12\langle F \rangle^2\langle F^2 \rangle - 6\langle F \rangle^4$

Table E.1: List of ordinary cumulants up to the 4th order. $\delta F := F - \langle F \rangle$.

n	Joint cumulant $K_n\left[F(\mathbf{r}_1, t+\tilde{\tau}_1), \ldots, F(\mathbf{r_n}, t+\tilde{\tau}_n)\right]$
1	$\langle F_1 \rangle$
2	$\langle \delta F_1 \delta F_2 \rangle = \langle F_1 F_2 \rangle - \langle F_1 \rangle \langle F_2 \rangle$
3	$\langle \delta F_1 \delta F_2 \delta F_3 \rangle =$ $\langle F_1 F_2 F_3 \rangle - \langle F_1 \rangle \langle F_2 F_3 \rangle - \langle F_2 \rangle \langle F_1 F_3 \rangle - \langle F_3 \rangle \langle F_1 F_2 \rangle + 2\langle F_1 \rangle \langle F_2 \rangle \langle F_3 \rangle$
4	$\langle \delta F_1 \delta F_2 \delta F_3 \delta F_4 \rangle - \langle \delta F_1 \delta F_4 \rangle \langle \delta F_2 \delta F_3 \rangle - \langle \delta F_1 \delta F_3 \rangle \langle \delta F_2 \delta F_4 \rangle - \langle \delta F_1 \delta F_2 \rangle \langle \delta F_3 \delta F_4 \rangle$ $= \langle F_1 F_2 F_3 F_4 \rangle$ $-\Big(\langle F_1 \rangle \langle F_2 F_3 F_4 \rangle + \langle F_2 \rangle \langle F_1 F_3 F_4 \rangle + \langle F_3 \rangle \langle F_1 F_2 F_4 \rangle + \langle F_4 \rangle \langle F_1 F_2 F_3 \rangle$ $+ \langle F_1 F_2 \rangle \langle F_3 F_4 \rangle + \langle F_1 F_3 \rangle \langle F_2 F_4 \rangle + \langle F_1 F_4 \rangle \langle F_2 F_3 \rangle \Big)$ $+2\cdot\Big(\langle F_1 \rangle \langle F_2 \rangle \langle F_3 F_4 \rangle + \langle F_1 \rangle \langle F_3 \rangle \langle F_2 F_4 \rangle + \langle F_1 \rangle \langle F_4 \rangle \langle F_3 F_2 \rangle$ $+ \langle F_2 \rangle \langle F_3 \rangle \langle F_1 F_4 \rangle + \langle F_2 \rangle \langle F_4 \rangle \langle F_1 F_3 \rangle + \langle F_3 \rangle \langle F_4 \rangle \langle F_1 F_2 \rangle \Big)$ $-6\cdot\langle F_1 \rangle \langle F_2 \rangle \langle F_3 \rangle \langle F_4 \rangle$

Table E.2: List of joint cumulants up to the 4th order. $F_i := F(\mathbf{r}_i, t+\tilde{\tau}_i)$. $\delta F_i := F_i - \langle F_i \rangle$.

Two state model joint cumulant (absolute time-lags)	
$K_2(\tau_1)$	$\frac{N_{\text{em}}\varepsilon^2}{(k_++k_-)^2} \cdot k_+ k_- \cdot e^{-(k_++k_-)\tau_1}$
$K_3(\tau_1, \tau_2)$	$\frac{N_{\text{em}}\varepsilon^3}{(k_++k_-)^3} \cdot (k_+ k_-^2 - k_+^2 k_-) \cdot e^{-(k_++k_-)\tau_2}$
$K_4(\tau_1, \tau_2, \tau_3)$	$\frac{N_{\text{em}}\varepsilon^4}{(k_++k_-)^4} \cdot \big[(k_+ k_-^3 - 2k_+^2 k_-^2 + k_+^3 k_-) \cdot e^{-(k_++k_-)\tau_3} - 2k_+^2 k_-^2 \cdot e^{-(k_++k_-)(\tau_2+\tau_3-\tau_1)}\big]$

Table E.3: Two state model joint cumulant in terms of absolute time-lags (see Fig. C.2). All time-lags are positive and ordered $\tau_3 \geq \tau_2 \geq \tau_1 \geq 0$. N_{em} denotes the number of emitting molecules, ε their molecular brightness. The transition rate from the off- to the on-state is k_+, while k_- is its counterpart in the other direction.

Two state model joint cumulant (relative time-lags)	
$K_2(\tilde{\tau}_1)$	$\frac{N_{\text{em}}\varepsilon^2}{(k_++k_-)^2} \cdot k_+ k_- \cdot e^{-(k_++k_-)\tilde{\tau}_1}$
$K_3(\tilde{\tau}_1, \tilde{\tau}_2)$	$\frac{N_{\text{em}}\varepsilon^3}{(k_++k_-)^3} \cdot (k_+ k_-^2 - k_+^2 k_-) \cdot e^{-(k_++k_-)(\tilde{\tau}_1+\tilde{\tau}_2)}$
$K_4(\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3)$	$\frac{N_{\text{em}}\varepsilon^4}{(k_++k_-)^4} \cdot \big[(k_+ k_-^3 - 2k_+^2 k_-^2 + k_+^3 k_-) \cdot e^{-(k_++k_-)(\tilde{\tau}_1+\tilde{\tau}_2+\tilde{\tau}_3)} - 2k_+^2 k_-^2 \cdot e^{-(k_++k_-)(\tilde{\tau}_1+2\tilde{\tau}_2+\tilde{\tau}_3)}\big]$

Table E.4: Two state model joint cumulant in terms of relative time-lags (see Fig. C.2). All time-lags are positive $\tilde{\tau}_i \geq 0$. N_{em} denotes the number of emitting molecules, ε their molecular brightness. The transition rate from the off- to the on-state is k_+, while k_- is its counterpart in the other direction.

n	Factorial cumulants / Poisson noise corrected cumulants
1	κ_1^{Poiss}
2	$\kappa_2^{\text{Poiss}} - \kappa_1^{\text{Poiss}}$
3	$\kappa_3^{\text{Poiss}} - 3\kappa_2^{\text{Poiss}} + 2\kappa_1^{\text{Poiss}}$
4	$\kappa_4^{\text{Poiss}} - 6\kappa_3^{\text{Poiss}} + 11\kappa_2^{\text{Poiss}} - 6\kappa_1^{\text{Poiss}}$

Table E.5: List of factorial cumulants up to the 4th order. The factorial cumulants give the rules how to compute the n-th order ordinary cumulant of a noise-less random signal $x(t)$ from the linear combination of ordinary cumulants κ_n^{Poiss} of a measurement $y(t)$ of x subject to Poisson noise. This is useful for correcting (zero lag-time, same-pixel) SOFI images for the effects of the noise (see Section 2.5).

n	EMCCD noise corrected cumulants
1	$\frac{1}{a}\kappa_1^{\text{EMCCD}}$
2	$\frac{1}{a^2}\left(\kappa_2^{\text{EMCCD}} - 2a \cdot \kappa_1^{\text{EMCCD}}\right)$
3	$\frac{1}{a^3}\left(\kappa_3^{\text{EMCCD}} - 6a \cdot \kappa_2^{\text{EMCCD}} + 6a^2 \cdot \kappa_1^{\text{EMCCD}}\right)$
4	$\frac{1}{a^4}\left(\kappa_4^{\text{EMCCD}} - 12a \cdot \kappa_3^{\text{EMCCD}} + 36a^2 \cdot \kappa_2^{\text{EMCCD}} - 24a^3 \cdot \kappa_1^{\text{EMCCD}}\right)$

Table E.6: List of EMCCD noise corrected cumulants up to the 4th order. The photon conversion factor $a = g/s$ is the EM-gain g divided by the sensor sensitivity s and depends on the settings of the camera. The linear combinations show how to compute the n-th order ordinary cumulant of a noise-less random signal $x(t)$ from the linear combination of ordinary cumulants κ_n^{EMCCD} of a measurement $y(t)$ of x subject to noise from an EMCCD camera. This is useful for correcting (zero lag-time, same-pixel) SOFI images for the effects of the noise (see Section 2.5).

Wer nur einigermaßen zur Freiheit der Vernunft gekommen ist, kann
sich auf Erden nicht anders fühlen, denn als Wanderer, – wenn auch
nicht als Reisender nach einem letzten Ziele: denn dieses gibt es nicht.
Wohl aber will er zusehen und die Augen dafür offen haben, was Alles in
der Welt eigentlich vorgeht; deshalb darf er sein Herz nicht allzufest an
alles Einzelne anhängen; es muss in ihm selber etwas Wanderndes sein,
das seine Freude an dem Wechsel und der Vergänglichkeit habe. […]

Acknowledgements

*He who has come only in part to a freedom of reason cannot feel on earth
otherwise than as a wanderer – though not as a traveler towards a final goal,
for this does not exist. But he does want to observe, and keep his eyes open for
everything that actually occurs in the world; therefore he must not attach his
heart too firmly to any individual thing; there must be something wandering
within him, which takes its joy in change and transitoriness. […]*

Friedrich Nietzsche - *Menschliches, Allzumenschliches* (638, Auszug)

THIS THESIS WAS A JOURNEY that I fortunately did not have to travel alone. And while no scientific
conclusion is ever carved in stone[1], as our knowledge evolves, constantly changing our views of the
world, its the memories of this time of my life that stay, and that I will always remember fondly. I would
like to express my heartfelt thanks to the people that accompanied and supported me:

First, I want to thank my supervisor, JÖRG ENDERLEIN, not only for contributing so many great ideas,
countless fruitful discussions and giving me a lot of freedom in my research, but also for always being a
kind, sympathetic and supportive boss and person. This warded me from a lot of potential worries and
created a feel-good atmosphere at work, which allowed me to focus my mind on science – something I
really do not take for granted.

I am very grateful to my marvelous, exceptional, fantastic, tremendous, wonderful long-time friend,
colleague, co-author, officemate and gaming-companion JAN THIART, whom I honestly cannot thank
enough with almost all *Thesaurus.com* synonyms for "great". Without you this journey would have been
far less enjoyable and maybe I would not even be where I am today – or possibly anywhere else for that
matter. I am glad to have had such a good Leidensgenosse (companion-in-misfortune) during my PhD.

I thank all people I had the chance to collaborate with, especially ANJA HUSS, WEIXING LI, NARAIN
KAREDLA, SEBASTIAN ISBANER, HAO CHENG, and ANNA and ALEXEY CHIZHIK. I always enjoyed
working with you and am thankful for the chance to peak into topics outside my core projects.

Special thanks again to ANJA HUSS, WEIXING LI, NARAIN KAREDLA and HAO CHENG for supplying
me with lots of the experimental data you cook in those "labs" I keep hearing about. Without you, this
thesis would be mainly about meaningless ones and zeros.

---

[1]Although, as shown here, they are sometimes carved **by** a stone! – English-German pun intended.

This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Arno Pro, designed by Robert Slimbach in the style of book types from the Aldine Press in Venice, and issued by Adobe in 2007. A template, which can be used to format a PhD thesis with similar look and feel, has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/ or from the author at suchow@post.harvard.edu.

Curriculum Vitae

Simon Christoph Stein

Date of birth: 13.04.1988

Place of birth: Gehrden

Nationality: German

ACADEMICAL BACKGROUND

01.2014 - present	Doctoral Dissertation
	Thesis: *Advanced Data Processing for Super-resolution Microscopy*
	Georg-August-Universität, Göttingen
10.2011 - 10.2013	Master of Science *with distinction* in Physics
	Thesis: *Fast Bottom-Up Segmentation of 3D-Pointclouds using Cues in Local Geometry*
	Georg-August-Universität, Göttingen
08.2011-10.2011	Research Internship
	Fraunhofer Ernst-Mach Institut, Freiburg im Breisgau
10.2008 - 07.2011	Bachelor of Science *with distinction* in Physics
	Thesis: *Zweidimensionales Elektronengas in zufällig gestörten periodischen Magnetfeldern*
	Georg-August-Universität, Göttingen
07.1998 - 05.2007	Abitur (equivalent to A level)
	Engelsburg-Gymnasium, Kassel

AWARDS

2014	Dr. Berliner – Dr. Ungewitter – award for excellent Master's thesis

Conferences

Talks

2015	Focus on Microscopy (FOM), Göttingen
2014	PicoQuant 20th International Workshop on Single Molecule Spectroscopy and Ultra-sensitive Analysis in the Life Sciences, Berlin

Posters

2015	DPG Frühjahrstagung - *EPL Poster Award (2nd prize)*
2015	Quantitative Bio-Imaging (QBI)

Publications

2016   **Stein, S.C.**; Huss, A.; Gregor, I.; Enderlein, J.; "Stochastic Optical Fluctuation Imaging", In A. Diaspro and M. AMJ can Zandvoort, editors, Super-Resolution Imaging in Biomedicine, chapter 9, CRC Press, 2016

Stein, S.C.; Thiart, J.; "TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms.", Scientific Reports, 6, 37937, 2016.

Isbaner, S.; Karedla, N., Ruhlandt, D.; **Stein, S.C.**; Chizhik, A.; Gregor, I.; Enderlein, J.; "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging", Optics Express, 24, 9429-9445, 2016.

Chizhik, A.M.; **Stein, S.C.**; Dekaliuk, M.O.; Battle, C.; Li, W.; Huss, A.; Platen, M., Schaap, I.A.T.; Gregor, I.; Demchenko, A.P.; Schmidt, C.F., Enderlein, J.; Chizhik, A. I.; "Super-Resolution Optical Fluctuation Bio-Imaging with Dual-Color Carbon Nanodots", Nano Letters, 16, 237-242, 2016.

2015   Karedla, N.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Chizhik, A.; Enderlein, J.; "Simultaneous Measurement of the Three-Dimensional Orientation of Excitation and Emission Dipoles", Physical Review Letters, 115, 173002, 2015.

Stein, S.C.; Huss, A.; Hähnel, D.; Gregor, I.; Enderlein, J.; "Fourier interpolation stochastic optical fluctuation imaging", Optics Express, 23, 16154-16163, 2015.

Weixing, L.; **Stein, S.C.**; Hähnel, D.; Gregor, I.; Enderlein, J.; "Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy", Optics Express, 23, 3770-3783, 2015.

Tamosiunaite, M.; Sutterlütti, R.; **Stein, S. C**; Wörgötter, F.; "Perceptual influence of elementary three-dimensional geometry:(2) fundamental object parts", Frontiers in psychology, 6, 2015

2014   Sutterluetti, R.; **Stein, S. C.**; Tamosiunaite, M.; Woergoetter, F.; "Object names correspond to convex entities", Cognitive Processing,15 ,1 ,69-71 ,2014

Schoeler, M; **Stein, S. C.**; Papon, J.; Abramov, A.; Wörgötter, F.; "Fast self-supervised on-line training for object recognition specifically for robotic applications", Computer Vision Theory and Applications (VISAPP), 2, 94-103, 2014

Stein, S. C.; Wörgötter, F.; Schoeler, M.; Papon, J.; Kulvicius, T.; "Convexity based object partitioning for robot applications", Robotics and Automation (ICRA), 3213-3220, 2014

Stein, S. C.; Schoeler, M.; Papon, J.; Worgotter, F.; "Object partitioning using local convexity", Computer Vision and Pattern Recognition (CVPR), 304-311, 2014