

**Stochastic modeling of  
photoswitchable fluorophores for  
quantitative superresolution  
microscopy**

**Dissertation  
for the award of the degree**

**“Doctor rerum naturalium”**

**of the Georg-August-Universität Göttingen**

**within the  
International Max Planck Research School  
Physics of Biological and Complex Systems**

**of the  
Göttingen Graduate School for Neurosciences,  
Biophysics, and Molecular Biosciences**

**submitted by**

**Lars Frahm**

**from Bremen**

**Göttingen 2016**

## **Thesis Committee**

### **Prof. Dr. Stefan W. Hell (Referee)**

Department of NanoBiophotonics,  
Max Planck Institute for Biophysical Chemistry, Göttingen

### **Prof. Dr. Axel Munk (2<sup>nd</sup> Referee)**

Felix-Bernstein-Chair for Mathematical Statistics,  
Institute for Mathematical Stochastics, Georg-August-Universität Göttingen

### **Prof. Dr. Helmut Grubmüller**

Department of Theoretical and Computational Biophysics,  
Max Planck Institute for Biophysical Chemistry, Göttingen

## **Further members of the examination board**

### **Prof. Dr. Sarah Köster**

Research Group for Nanoscale Imaging of Cellular Dynamics,  
Institute for X-Ray Physics, Georg-August-Universität Göttingen

### **Dr. Florian Rehfeldt**

Research Group for Cell and Matrix Mechanics,  
3<sup>rd</sup> Institute of Physics, Georg-August-Universität Göttingen

### **Dr. Katrin Willig**

Research Group for Nanoscale Microscopy and Molecular Physiology of the Brain,  
Max Planck Institute of Experimental Medicine, Göttingen

**Date of oral examination: 23.11.2016**

# **Stochastic modeling of photoswitchable fluorophores for quantitative superresolution microscopy**

Lars Frahm

The diffraction limit of optical microscopy can be overcome by switching fluorophores between on- and off-states. In the on-state, the fluorophores emit fluorescence and can be detected. In the off-state, they remain dark. By modeling the stochastic switching of the fluorophores between the on- and the off-state, quantitative insights can be gained into the image formation process of superresolution microscopy.

Reversible saturable optical fluorescent transitions (RESOLFT) can be used to drive the fluorophores to a non-fluorescent off-state using light. In a RESOLFT microscope, the size of the region in which there is a high probability that the fluorophores remain in the on-state can be confined to the sub-diffraction scale. To count the number of fluorophores in a RESOLFT image, the photoswitching process of reversibly switchable fluorescent proteins was modeled. Based on this model, a method was developed to calibrate the brightness per fluorescent protein directly from the image data. The result of the analysis is an estimate of the number density of fluorophores in the image.

In stochastic optical reconstruction microscopy (STORM), a superresolved image can be reconstructed by switching fluorophores stochastically between the on- and off-state, and localizing the position of single molecule events. In activation-based multicolor STORM, several species of activator-reporter fluorophore pairs can be distinguished by activating them specifically at different points in time. In this method, the crosstalk between the color channels is typically high due to random activation of the reporter fluorophores, which can happen at the same time as the specific activation. We introduce a new approach to avoid this principal source of crosstalk in multicolor STORM, based on estimating the on-switching time of single molecule events with sub-frame precision. Our method enables the assignment of switching events to the correct molecular species with an improved error rate. This significantly reduces the crosstalk in activation-based multicolor STORM, and presents a solution for high resolution STORM imaging of multiple molecular species.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>List of abbreviations</b>	<b>vi</b>
<b>List of Figures</b>	<b>vii</b>
<b>1. Introduction</b>	<b>1</b>
1.1. Quantitative RESOLFT microscopy . . . . .	1
1.2. Activation-based multicolor STORM . . . . .	3
<b>2. Theoretical modeling of superresolution microscopy</b>	<b>5</b>
2.1. Photoswitchable fluorophores . . . . .	6
2.1.1. Properties of Markov chains . . . . .	7
2.1.2. Two state Markov Model . . . . .	8
2.1.3. Rate estimation for the exponential distribution . . . . .	10
2.2. RESOLFT microscopy . . . . .	14
2.2.1. Stochastic image formation . . . . .	14
2.2.2. Effective RESOLFT PSF for linear photoswitching . . . . .	17
2.2.3. Counting molecules in RESOLFT microscopy . . . . .	21
2.2.4. Deconvolution and measurement of the effective PSF . . . . .	22
2.2.5. Estimation of the brightness and the number density . . . . .	25
2.3. Single molecule localization microscopy . . . . .	28
2.3.1. Modeling of single molecule kinetics data and estimation of the on-switching rate . . . . .	29
2.3.2. Single molecule event timing and estimation of the off-switching rate . . . . .	32
2.3.3. Modeling of activation-based multicolor STORM . . . . .	37
<b>3. Quantitative RESOLFT microscopy</b>	<b>41</b>
3.1. Switching kinetics measurements . . . . .	42
3.1.1. Preparation of the initial state . . . . .	44
3.1.2. Switching kinetics for rsEGFP and rsEGFP2 . . . . .	44

## Contents

3.2. Counting molecules in simulated images . . . . .	47
3.3. Experimental results in RESOLFT images . . . . .	52
3.3.1. Estimation of the effective PSF and calibration of the effective on-switching probability . . . . .	52
3.3.2. Quantification of the excess variance and counting molecules in RESOLFT microscopy . . . . .	54
<b>4. Activation-based multicolor STORM</b>	<b>58</b>
4.1. Measuring Alexa 647 on-switching kinetics . . . . .	58
4.2. Single molecule event timing measurements . . . . .	61
4.2.1. Switching time estimation for fluorescent beads . . . . .	61
4.2.2. Switching time estimation for single molecules . . . . .	61
4.3. Alexa 647 off-switching kinetics . . . . .	64
4.4. Multicolor STORM using single molecule event timing . . . . .	64
4.4.1. Quantification of the sources of crosstalk . . . . .	69
<b>5. Discussion</b>	<b>74</b>
5.1. Two state model . . . . .	74
5.1.1. States on shorter timescales . . . . .	75
5.1.2. States on longer timescales . . . . .	75
5.2. Quantitative RESOLFT microscopy . . . . .	76
5.2.1. Blind deconvolution and regularization . . . . .	77
5.2.2. Counting error . . . . .	78
5.3. Single molecule event timing and multicolor STORM . . . . .	79
5.3.1. Single molecule event timing . . . . .	79
5.3.2. Crosstalk in multicolor SMLM . . . . .	80
<b>6. Outlook</b>	<b>81</b>
<b>A. Bibliography</b>	<b>83</b>
<b>B. Acknowledgements</b>	<b>93</b>
<b>C. Curriculum vitae</b>	<b>94</b>

## List of abbreviations

<b>FCS</b>	Fluorescence correlation spectroscopy
<b>FRET</b>	Fluorescence resonance energy transfer
<b>FWHM</b>	Full width at half maximum
<b>GFP</b>	Green fluorescent protein
<b>HMM</b>	Hidden Markov model
<b>PSF</b>	Point spread function
<b>RESOLFT</b>	Reversible saturable optical fluorescent transitions
<b>rsFP</b>	Reversibly switchable fluorescent protein
<b>rsEGFP</b>	Reversibly switchable enhanced green fluorescent protein
<b>SMLM</b>	Single molecule localization microscopy
<b>SOFI</b>	Superresolution optical fluctuation imaging
<b>STED</b>	Stimulated emission depletion
<b>STORM</b>	Stochastic optical reconstruction microscopy

## List of Figures

2.1. Two state Markov model. . . . .	9
2.2. RESOLFT image formation. . . . .	18
2.3. Deconvolution of the effective PSF. . . . .	24
2.4. Linear model for trace extraction . . . . .	30
2.5. Single molecule event timing . . . . .	33
2.6. Hidden Markov model (HMM) for identifying the transition frames. . . . .	36
2.7. Activation-based multicolor STORM . . . . .	38
3.1. Ensemble switching curves . . . . .	43
3.2. Switching kinetics data for rsEGFP2. . . . .	45
3.3. Switching kinetics for rsEGFP and rsEGFP2. . . . .	46
3.4. Simulated images of isolated clusters of rsEGFP2. . . . .	49
3.5. Simulated images of isolated lines of rsEGFP2. . . . .	50
3.6. Parameter estimation for $b$ in simulated data. . . . .	51
3.7. RESOLFT PSF estimation . . . . .	53
3.8. Log likelihood and posterior for $b$ . . . . .	56
3.9. Estimation of the number density of rsEGFP2 . . . . .	57
4.1. Estimation of the on-switching rate for Alexa 647. . . . .	60
4.2. Timing precision for fluorescent beads . . . . .	62
4.3. Single molecule measurements of the timing precision. . . . .	63
4.4. Off-switching rate measurement for Alexa647. . . . .	65
4.5. Multicolor STORM image of Tom20 and Tubulin. . . . .	67
4.6. On-switching transition time and color assignment. . . . .	68
4.7. Multicolor STORM image of gp210 and panFG. . . . .	71
4.8. Multicolor STORM image of panFG control sample . . . . .	72
4.9. Multicolor STORM image of gp210 control sample . . . . .	73

# 1. Introduction

With optical fluorescence microscopy, direct insights into biological processes can be gained. But due to diffraction, the resolution of fluorescence microscopy was limited to about half the wavelength of the fluorescence light [1]. The diffraction limit of far-field optical microscopy was first overcome by stimulated emission depletion (STED), by switching fluorescent dyes into a non-fluorescent off-state using stimulated emission [2, 3]. Since then, many different physical processes have been used to achieve superresolution by switching fluorophores between a fluorescent on- and a non-fluorescent off-state [4–6].

For light-driven processes, the incident light can change the probability that a fluorophore resides in the on-state. This can be used to confine the regions in which the fluorophores have a high probability to remain in the on-state to sub-diffraction sized spots. By illuminating the sample with patterns of light, this probability is driven close to zero everywhere else. These so called coordinate-targeted superresolution methods are based on reversible saturable optical (fluorescent) transitions (RESOLFT) [7].

By driving the probability that fluorophores reside in the on-state close to zero everywhere, these switching processes can also be exploited by localizing sparsely activated single molecules. The position of an individual bright emitter can be determined with a high precision [8, 9]. By repeatedly estimating the positions of random subsets of the population of fluorophores, a superresolved image can be reconstructed in single molecule localization microscopy (SMLM) [10–15].

## 1.1. Quantitative RESOLFT microscopy

RESOLFT microscopy based on reversibly switchable fluorescent proteins (rsFPs) is ideally suited for quantitative live-cell imaging [16–18]. Fluorescent proteins can be expressed natively in living organisms such that the structure of interest can be labeled quantitatively [19, 20]. Thus, not only can the structure be imaged with high resolution, but the number of fluorophores making up the image can be estimated from the data. At first glance, counting the number of fluorophores in a fluorescence



## 1. Introduction

image seems straightforward. As the image can be modeled as the superposition of the images of many single fluorophores, the number of fluorophores can be estimated by dividing the image intensity by the brightness per fluorophore [21]. In practice, however, the characteristic fluorophore brightness cannot be calibrated easily, as it depends on the illumination intensity distributions in the focal plane of the microscope, the response of the fluorophores and the detection efficiency of the optical system. Ideally, the brightness should be calibrated from the image data itself to take into account all these influences. To this end, several methods have been developed, both for diffraction-limited and superresolution fluorescence microscopy, which we compare in the following.

For single clusters of few fluorophores each, the number of molecules in each cluster can be determined by imaging them repeatedly until the fluorophores reach an irreversible photobleached state [22]. If the signal-to-noise ratio is large enough to determine steps in the decaying signal, the height of these bleaching steps can be used to calibrate the brightness per fluorophore.

Another approach is based on the fact, that a fluorophore in the excited state can only emit one photon at a time. Thus, by investigating the fluorescence photon statistics, the brightness per fluorophore can be calibrated [23]. This approach was recently also applied to map the number of fluorophores in STED images [24].

SMLM can also be used to quantify the number of fluorophores in an image. Instead of calibrating the brightness of the fluorophores in terms of measured photons, here, the number of switching cycles has to be estimated precisely [25,26]. However, due to photobleaching, the number of switching cycles for single fluorophores is typically geometrically distributed, as most fluorophores can switch to the on-state several times [27,28]. Also, some single molecules may not be detected because they emit an insufficient number of photons, leading to missed events [29]. Thus, the switching kinetics have to be calibrated carefully to avoid errors due to missed events and repeated activation of fluorophores [30]. The problem of geometrically distributed event numbers can be partly overcome by using binding and unbinding fluorescently labeled DNA strands [31]. The influx rate of DNA strands is held constant, such that the number of events follows a Poisson distribution. Nevertheless, the average binding rate still has to be calibrated from single binding targets.

Fluorescence correlation spectroscopy (FCS) can also be used to determine the brightness per single molecule [32–34]. Here, fluorophores diffuse through the detection volume, which can be diffraction-limited, or even below the diffraction limit using STED-FCS [35]. Given the same average signal, the fluctuations from many

## 1. Introduction

dim molecules are smaller than those from a few bright molecules diffusing through the focus. The brightness per molecule can be estimated from the mean and the variance of the observed temporal intensity fluctuations using a model for the number and brightness of the fluorophores [34].

For fluctuations due to photoswitching, the number and the brightness of the fluorophores can also be modeled using higher statistical moments of the distribution of measured photon counts. This approach is taken in balanced superresolution optical fluctuation imaging (SOFI) [36], where the brightness per fluorophore is estimated from the statistical cumulants of the signal in a wide-field fluorescence image.

Modeling the switching properties of single molecules due to different kinds of state transitions is the basis for these counting methods. In this work, we want to estimate the brightness of the fluorophores from the fluctuations due to the photoswitching process inherent in RESOLFT microscopy. We have to take into account both the shape of the intensity distribution in the focus, and the response of the rsFPs on this intensity. To this end, we measure the intensity-dependent switching rates for two rsFPs, the reversibly switchable enhanced green fluorescent proteins rsEGFP and rsEGFP2 [17, 18]. Using a simple two state model in which the switching rates are linearly proportional to the intensity, we calculate the mean and the variance of the fluorescence signal measured in RESOLFT microscopy. This model is applied to estimate both the brightness per fluorophore and the number density of the fluorophores in a single RESOLFT image in living tissue.

## 1.2. Activation-based multicolor STORM

In SMLM, each single molecule switches on individually, such that the position of each single molecule event can be localized with a high precision. Additional information can be gained from analyzing these single molecule signals. By measuring spectral information from single molecule data, multicolor images can be reconstructed from samples, in which different structures are labeled with different species of single molecules [37–41]. However, due to chromatic aberrations in the optical system, the images of single molecules of different colors may be distorted, such that the final image has to be corrected by transforming and aligning images for the different color channels [6, 42–44]. An alternative multicolor SMLM method images different color species by labeling and imaging the sample sequentially [45, 46]. But again, the images for the color channels have to be aligned. As the localization precision, and thus the resolution, reaches the single nanometer scale [47, 48], the error

## 1. Introduction

due to these image registration methods can be on the order of the localization precision, such that co-localization studies on the nanometer scale become increasingly difficult [6, 49].

An SMLM method that avoids the need for image registration is multicolor STORM (stochastic optical reconstruction microscopy) using activator-reporter dye pairs [11, 50]. Here, the sample is labeled with the same reporter fluorophore species for all color species in the sample. For color discrimination, each reporter fluorophore is paired with an activator dye. Different structures in the sample are labeled with activator dyes that react to light of different wavelengths. Several activator-reporter species can be discriminated by activating them periodically at different times. Localization events that are first detected in an image frame in which activation light is applied, are assigned to the corresponding color channel.

This method typically shows a high crosstalk between the color channels of 10 to 20% [42], while spectral separation of the fluorescence emission allows a crosstalk of lower than a few percent [39–41]. In sequential imaging methods, negligible crosstalk is reported [45, 46]. The high crosstalk in activation-based multicolor STORM is caused by two sources. First, the fluorescent dyes can be activated with activation light designated to activate another species. Second, random activation of the reporter dye can happen at the same time during which the activation light is applied. As events are typically assigned to each color channel by means of the frame in which they were first detected, single molecules that activate randomly during the same frame are assigned to the color channel as well.

To address the problem of the high crosstalk due to random activation, we develop a method to estimate the on- and off-switching transition time from the single molecule data with a sub-frame time precision. By estimating this time precisely, randomly activated single molecule events are recognized and discarded. The method for estimating the switching transition times is also applied to analyze the switching kinetics of Alexa 647, a single molecule fluorophore frequently used as reporter in activation-based multicolor STORM [51], based on a two state switching model.

## 2. Theoretical modeling of superresolution microscopy

In order to understand superresolution methods based on photoswitching fluorescent markers quantitatively, a mathematical model is needed to describe the switching behavior of the fluorophores. In this chapter, we will first discuss on- and off-switching fluorescence transitions which have been used for achieving superresolution in microscopy. Markov models can in general be used to describe the stochastic behavior of systems switching between different states. Methods for estimating the rate parameters for different kinds of data are introduced, including single molecule traces and ensemble switching measurements.

The switching kinetics estimated using these methods are then applied to quantitatively model two superresolution methods, RESOLFT microscopy based on reversibly switchable fluorescent proteins and STORM single molecule localization microscopy based on photoswitchable organic dyes.

In RESOLFT microscopy, the signal of an ensemble of many molecules is integrated onto the detector. This usually results in a large signal, but the individual behavior of each molecule cannot be identified any more. Additionally, the contribution of each molecule is weighted by the point spread function (PSF) of the microscope.

By taking into account prior knowledge about the imaged structure, we develop a method to estimate the effective PSF from an image based on a blind deconvolution procedure. Finally, we will discuss the statistical properties of the signal measured in RESOLFT microscopy. This stochastic image formation model is applied to estimate the average brightness of the fluorophores and to infer the number of fluorescent proteins in a RESOLFT image.

In SMLM, the stochastic behavior of single molecules is observed directly. However, this kind of measurement requires very bright fluorophores which can be detected one by one. The amount of single molecule data needed to estimate the switching rates with sufficient precision has to be very large, as the on- and off-switching rates span several orders of magnitude.

## 2. Theoretical modeling of superresolution microscopy

To measure the on-state dwell times from single molecule data, we develop a method to estimate the switching transition time for single molecule events with a sub-frame time precision.

The measured single molecule kinetics are then used to model the imaging process in SMLM. This model shows, that the crosstalk in activation-based multicolor STORM can be reduced by estimating the on-switching transition time, which in turn allows to assign single molecule events to each color channel with a lower error rate.

### 2.1. Photoswitchable fluorophores

Typically, two states are responsible for the resolution improvement in superresolution microscopy, a fluorescent on-state and a non-fluorescent off-state. The physical processes leading to on- and off-states useful for superresolution microscopy are manifold. Two different kinds of transitions can be distinguished: Light-driven and spontaneous switching transitions.

Light-driven transitions can be characterized by their effective switching cross-section [52,53], which is the effective area over which a single molecule absorbs light to undergo the transition. An important physical cross-section is the absorption cross-section  $\sigma$  for a absorbing a photon to reach the excited state. For single photon transitions,  $\sigma$  is typically on the order of  $10^{-14}$  to  $10^{-16}\text{cm}^2$ .

Once in the excited state, every fluorophore shows the spontaneous emission as one example of spontaneous switching transitions. Spontaneous switching transitions can also occur due to chemical reactions [54].

Spontaneous switching transitions from an excited state, can be described by an effective switching cross-section. The probability of the fluorophore to undergo the transition is proportional to the probability that it is excited. The effective switching cross-section can be directly determined by measuring the dependence of the switching rate  $k$  on the intensity  $I$ :

$$k = \sigma_{\text{eff}} \frac{\lambda}{hc} I. \quad (2.1)$$

Here,  $\lambda/hc$  is the inverse of the photon energy at wavelength  $\lambda$ .

The switching rates in light-driven switching transitions typically depend on the wavelength of the light used. For example, the reversibly switchable fluorescent proteins rsEGFP and rsEGFP2 can be switched on using ultraviolet light and switched off using green light [17, 18]. This means, that for ultraviolet illumination, the on-switching rate  $k_{\text{on}}$  from the dark off- to the fluorescent on-state is larger than the

## 2. Theoretical modeling of superresolution microscopy

competing spontaneous or light-driven off-switching rate  $k_{\text{off}}$  from the on- to the off-state. For green illumination light, the off-switching rate  $k_{\text{off}}$  is larger than the on-switching rate  $k_{\text{on}}$ , such that the proteins can be switched to the off-state. By applying light of different wavelengths, the dynamic equilibrium between the on- and the off-state can be shifted.

If at least one of the switching rates between the states depends on the intensity of light, the state of the fluorophore can be controlled spatially using patterns of light intensity. This effect is used in coordinate-targeted superresolution methods [3, 5, 7]. Here, the state of the fluorophores is prepared such that the probability that it remains in the on-state is very low everywhere except in a small region at the focus. This is achieved for example by stimulated emission or reversible switching of fluorescent proteins.

For superresolution methods based on stochastic switching, the switching rates do not necessarily depend on the intensity, as long as the overall probability of the fluorophores to be in the on-state is small. A superresolved image can be reconstructed from data of stochastically switching fluorophores which switch on one by one by localizing the position of each single emitter in the images [10–15].

Although the switching processes can have very different physical realizations, they can often be modeled using the Markov chain model, which describes the mathematical properties of a system which is able to switch stochastically between different states.

### 2.1.1. Properties of Markov chains

Markov chains are statistical models of time series which have the property that the future development of the system only depends on the current state of the system, and not on the sequence of states in the past [55]. This Markov (or “memoryless”) property is applicable in many different problems and makes it possible to efficiently calculate many useful results.

A discrete time Markov chain can be defined by a state space which describes the different states the system can be in and a transition matrix  $P$  with the matrix elements  $p_{ij}$  which describes the probability of state changes in every (time) step:

$$P(X(t_n) = i | X(t_{n-1}) = j) = p_{ij}. \quad (2.2)$$

## 2. Theoretical modeling of superresolution microscopy

In continuous time, the probability matrix  $P(t) = p_{ij}(t)$  depends on the elapsed time interval. It describes the probability of ending up in state  $j$  after the time interval  $t$ , if the system started in state  $i$  at  $t = 0$ . This matrix fulfills the master equation

$$P'(t) = P(t)Q, \quad (2.3)$$

with the rate matrix  $Q = q_{ij}$ . The positive off-diagonal elements of  $Q$  can be interpreted as the transition rates from a state to a specific other state. The diagonal elements  $q_{ii}$  are the negative sums of the off-diagonal elements in each row and give the total rate of leaving each state, the so called decay rate. The dwell times in each state follow an exponential distribution with the corresponding decay rate.

The solution to the system of differential equations is given by the matrix exponential of  $Q$ :

$$P(t) = \exp(Qt). \quad (2.4)$$

Together with an initial state, this equation describes the time evolution of the average probability of residing in each state.

### 2.1.2. Two state Markov Model

We want to consider the two state continuous time Markov chain with the bright on-state and the dark off-state, as shown in Figure 2.1a.

The time evolution of this model is determined by two rate parameters: The on-switching rate  $k_{\text{on}}$  and the off-switching rate  $k_{\text{off}}$ . The  $Q$  matrix is given by

$$Q = \begin{pmatrix} -k_{\text{off}} & k_{\text{off}} \\ k_{\text{on}} & -k_{\text{on}} \end{pmatrix}. \quad (2.5)$$

Solving the differential equation (2.3), starting in the state  $(\rho_0, 1 - \rho_0)$ , that is, in a random state with probability  $\rho_0$  to be in the on-state, we get the result for the time evolution of the probability  $\rho_{\text{on}}$  that the system is in the on-state:

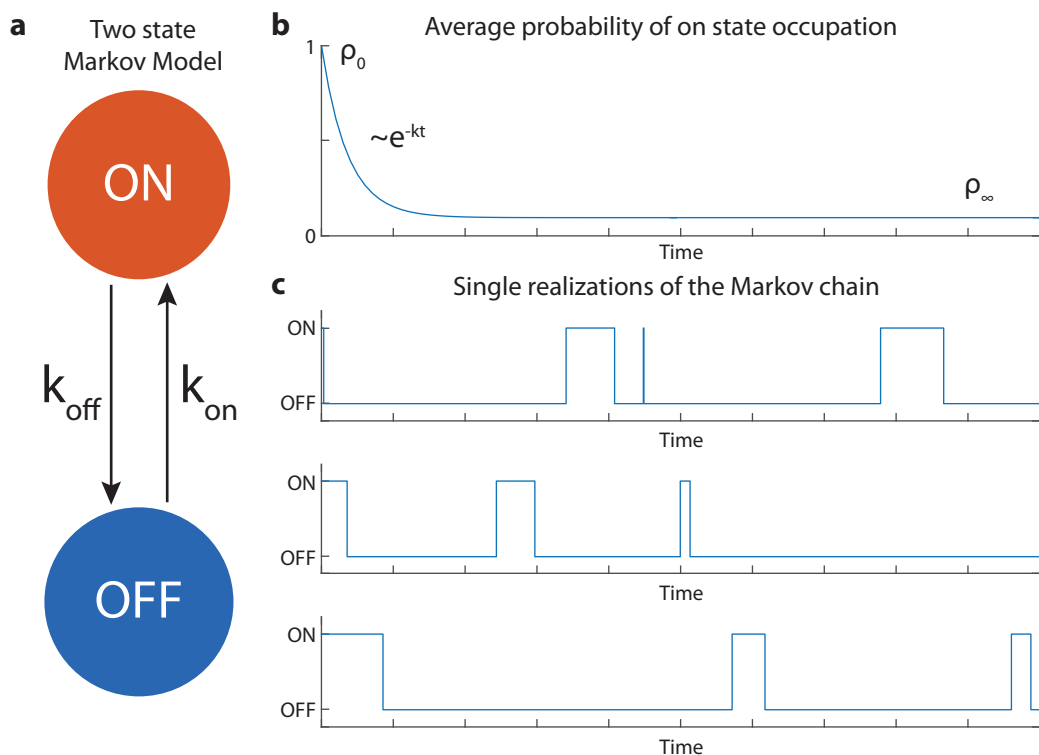
$$\rho_{\text{on}}(t) = \rho_{\infty} + (\rho_0 - \rho_{\infty}) \exp(-kt). \quad (2.6)$$

with the effective rate  $k = k_{\text{on}} + k_{\text{off}}$  and the equilibrium level or duty cycle

$$\rho_{\infty} = \frac{k_{\text{on}}}{k_{\text{on}} + k_{\text{off}}}. \quad (2.7)$$

For single realizations of the Markov chain, visualized in Figure 2.1c, the on- and off-state dwell times are distributed exponentially. To measure the parameters of this model for photoswitchable fluorophores, there are two principal ways: in ensemble measurements, the signals from many fluorophores are superimposed. In the simplest

## 2. Theoretical modeling of superresolution microscopy



**Figure 2.1. Two state Markov model and time development.** **a:** State diagram for the two state Markov model. The on- and off-switching rates  $k_{\text{on}}$  and  $k_{\text{off}}$  determine the time development of the system. **b:** On average, the probability of residing in the on-state decays exponentially from the initial value  $\rho_0$  with an effective rate  $k = k_{\text{on}} + k_{\text{off}}$  to the equilibrium level  $\rho_\infty$ . **c:** Single realizations of the Markov chain show the stochastic switching behavior.

case, if the parameters are the same for all fluorophores contributing to the signal, the signal is proportional to the on-state probability, as described by Equation (2.6). In the following, we will also derive an expression for the ensemble behavior if the switching rates are variable and follow a distribution. Another approach are single molecule measurements, which are feasible for bright fluorophores. Here, the switching behavior is observed for each molecule one by one. The dwell times of each molecule in the on- and off-state are extracted. From the distribution of on- and off-state dwell times the rate parameters are estimated. Due to a limited time resolution, short on-state dwell times may not be detected. As the data sets are of finite length, off-state dwell times which are longer than the measurement time are not detected. We thus develop tools to estimate the switching rate in these cases for single molecule data.



### 2.1.3. Rate estimation for the exponential distribution

The exponential distribution with the kinetic rate parameter  $k$  describes the dwell times  $t > 0$  in a state in a continuous time Markov Chain. It takes the following form:

$$P(t|k) = k \exp(-kt). \quad (2.8)$$

Here,  $k$  can describe for example the on- or the off-rate in the simple two-state Markov chain. The on-rate governs the off-state dwell times, and vice versa. The average dwell time in the state is given by the inverse of the rate parameter:

$$\langle t \rangle = \frac{1}{k}. \quad (2.9)$$

We want to infer the rate parameter from measured dwell times. One possible way to do this would be to calculate the inverse of the average dwell times. However, the measured data sets are of finite length. If the dwell time in one of the states is comparable to the duration of the measurement, this procedure would underestimate the switching rate for this state, as switching events longer than the measurement would not be detected. We have to take this into account to get an accurate estimation of the rate. Also, it may not be possible to measure very short dwell times due to the time resolution of our system. In both cases, we may use Bayesian probability theory [56–58]. If we can measure the dwell time  $t$  only in a finite interval given by  $T_1 < t < T_2$ , the resulting distribution of  $t$  is

$$P(t|k) = \begin{cases} 0 & , \quad t < T_1 \\ \frac{k \exp(-kt)}{\exp(-kT_1) - \exp(-kT_2)} & , \quad T_1 < t < T_2 \\ 0 & , \quad t > T_2 \end{cases}. \quad (2.10)$$

After measuring  $N$  dwell times in a given state  $\{t_i\} = (t_1, \dots, t_N)$ , we can write down the posterior probability distribution for  $k$  using Bayes' theorem:

$$P(k|\{t_i\}) = \frac{\prod_i P(t_i|k)P(k)}{P(\{t_i\})} \quad (2.11)$$

Here, the probability distribution from above as a function of the parameter  $k$  is the likelihood function,  $P(k)$  is the prior distribution of the parameter  $k$ , and the normalization factor  $P(\{t_i\}) = \int dk P(\{t_i\}|k)P(k)$  is also called evidence or marginal likelihood.

In practice, a useful choice for the prior distribution is the gamma distribution [56, 58]. This distribution is defined for positive values of  $k$  and has a mean of  $\alpha/\beta$  and a variance of  $\alpha/\beta^2$ :

$$P(k) = P_\Gamma(k; \alpha, \beta) = \frac{\beta^\alpha k^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta k). \quad (2.12)$$

## 2. Theoretical modeling of superresolution microscopy

The gamma distribution is the conjugate prior for the exponential distribution, taking the same functional form as the likelihood function. Therefore many calculations with the gamma distribution as the prior are tractable analytically. When the parameters are chosen such that the distribution is broad, for example by choosing small  $\alpha$  and  $\beta$ , the choice of the prior has little influence on the estimation result for large enough amounts of data.

### Rate estimation for exponential distribution with lower threshold

If the upper limit  $T_2$  is large in comparison to the dwell time, we can neglect the upper limit and evaluate the expression for the posterior mean analytically. Using the gamma distribution given in Equation (2.12) as the prior, we can now calculate the evidence term in Equation (2.11):

$$\begin{aligned}
 P(\{t_i\}) &= \int dk P(\{t_i\}|k)P(k) \\
 &= \int dk \frac{k^N \exp(-k \sum t_i)}{\exp(-kNT_1)} \frac{\beta^\alpha k^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta k) \\
 &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int dk k^{N+\alpha-1} \exp(-k(\sum t_i + \beta - NT_1)) \\
 &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(N + \alpha)}{(\sum t_i + \beta - NT_1)^{N+\alpha}}. \tag{2.13}
 \end{aligned}$$

The posterior distribution then also takes the form of a gamma distribution:

$$\begin{aligned}
 P(k|\{t_i\}) &= \frac{\prod_i P(t_i|k)P(k)}{P(\{t_i\})} \\
 &= \frac{(\sum t_i - NT_1 + \beta)^{N+\alpha} k^{N+\alpha-1}}{\Gamma(N + \alpha)} \exp(-k(\sum t_i - NT_1 + \beta)) \\
 &= P_\Gamma(k; N + \alpha, \sum t_i - NT_1 + \beta). \tag{2.14}
 \end{aligned}$$

With the parameters  $N + \alpha$  and  $\sum t_i - NT_1 + \beta$ . We can use the known relations for the mean and the variance of the gamma distribution to find the posterior mean and variance:

$$E[k|\{t_i\}] = \frac{N + \alpha}{\sum t_i - NT_1 + \beta} \tag{2.15}$$

$$\text{Var}(k|\{t_i\}) = \frac{N + \alpha}{(\sum t_i - NT_1 + \beta)^2}. \tag{2.16}$$

## 2. Theoretical modeling of superresolution microscopy

We note, that small  $\alpha$  and  $\beta$  do not have a strong influence on the resulting estimate, as they are dominated by the number of measured dwell times  $N$  and the sum of all the measured times above the threshold  $\sum t_i - NT_1 = \sum(t_i - T_1)$ . In the limit of  $\alpha, \beta \rightarrow 0$ , we can use

$$\hat{k} = \frac{N}{\sum t_i - NT_1} \quad (2.17)$$

as the estimator for estimating the switching rate from measured dwell times.

To analyze the precision of the estimate depending on the number of measurements, we consider the coefficient of variation, given by

$$\frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(k|\{t_i\})}}{E[k|\{t_i\}]} = \frac{1}{\sqrt{N}}. \quad (2.18)$$

Thus, for example to measure the rate parameter with a relative error of 10%, we would need to measure the dwell times of 100 events.

### Estimation of distribution of rates in ensemble measurements

Often it is not possible to measure the on- and off-state dwell times directly, as the signal from single molecules cannot be detected. Integrating the signal of many molecules on a detector, it is still possible to measure the average behavior which is determined by Equation (2.6). The signal measured is therefore expected to follow an exponential decay. In this kind of ensemble measurement, the decay law is often seen to be non-exponential, due to variations in the switching rates. The transition rates can vary for several reasons. If the switching process depends on the intensity, a non-uniform illumination will lead to a superposition of many different switching rates, as is the case for point scanning rate measurements with confocal intensity distributions. Also, there may be an intrinsic distribution of behavior for complex fluorophores like rsFPs [59].

In these cases, we can model the signal as the superposition of exponential decays with different rates, where the rates have a distribution with a given mean and variance. As rates can only be positive, the gamma distribution, introduced in Equation (2.12), is used to describe this behavior.

The fluorescence signal  $s(t)$  of an ensemble with rates distributed according to a gamma distribution with parameters  $\alpha$  and  $\beta$  is

$$\begin{aligned} s(t) &= \int dk \exp(-kt) P_{\Gamma}(k; \alpha, \beta) \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \int dk k^{\alpha-1} \exp(-k(\beta + t)) \\ &= \frac{\beta^{\alpha}}{(\beta + t)^{\alpha}}. \end{aligned} \quad (2.19)$$

## 2. Theoretical modeling of superresolution microscopy

This decay law can then be fitted to the measured signal to extract the parameters of the rate distribution. The time until the signal drops to  $1/e$  is longer than for an exponential decay with the average rate of the distribution  $\beta/\alpha$ , given by

$$\tau_{1/e} = \beta(\exp(1/\alpha) - 1). \quad (2.20)$$

For narrow distributions, characterized by a small coefficient of variation  $\frac{\sigma}{\mu} = \frac{1}{\sqrt{\alpha}}$ , and thus a large parameter  $\alpha$ , the mean lifetime is approximately the lifetime of the exponential decay  $\beta(\exp(1/\alpha) - 1) = \beta/\alpha + O(1/\alpha^2)$ .

## 2.2. RESOLFT microscopy

RESOLFT microscopy is based on the principle that fluorophores can be switched between a bright on- and a dark off-state using light. We first discuss the image formation in a point scanning RESOLFT microscope. Using focal intensity distributions with an intensity zero [60], the fluorophores are switched such that they are only able to fluoresce in a small volume, smaller than the diffraction limit, around the zero of the intensity. By scanning the pattern over the sample, the superresolved image can be retrieved. The interaction between the focal intensity distributions and the intensity-dependent switching kinetics of the fluorophore determine the shape of the effective PSF of the RESOLFT microscope. Thus, to model quantitatively the image formation process, we need to model both the intensity patterns as well as measure the switching rates for the fluorophore. After each switching step, a subset of fluorophores remains in the on-state, with a distribution which in the end forms the effective RESOLFT PSF. As these switching processes are stochastic, we examine the fluctuations in the signal due to the random switching of the fluorophores. These fluctuations can be used to extract the number and the brightness of the fluorophores in the image.

### 2.2.1. Stochastic image formation

In a point scanning RESOLFT microscope, at each scanning position, the fluorophores are first activated, and then switched off using a doughnut intensity pattern. Finally, the fluorescence of the remaining fluorophores in the center of the intensity zero is read out confocally [17]. In each of the switching steps, the fluorophores are switched stochastically, according to Equation (2.6).

As both the on- and the off-switching rate  $k_{\text{on}}$  and  $k_{\text{off}}$  depend on the intensity  $I$  and the wavelength  $\lambda$  of the incident light, so do the equilibrium level  $\rho_{\infty}$  and the effective switching rate  $k$ . The probability that a fluorophore residing at the position  $\vec{r}$  remains in the on-state after each switching step depends on the wavelength  $\lambda$  of the applied light, the intensity at the position of the fluorophore  $I(\vec{r})$ , the duration  $t$  for which intensity is applied and the probability  $\rho_0(\vec{r})$  that the fluorophore was in the on-state initially:

$$\rho_{\text{on}}(t, I(\vec{r}), \lambda, \rho_0(\vec{r})) = \rho_{\infty}(I(\vec{r}), \lambda) + (\rho_0(\vec{r}) - \rho_{\infty}(I(\vec{r}), \lambda)) \exp(-k(I(\vec{r}), \lambda)t). \quad (2.21)$$

For several switching steps, which may use different intensity distributions and wavelengths, this function is applied consecutively. The final state of the preceding

## 2. Theoretical modeling of superresolution microscopy

switching step takes the role of the initial probability  $\rho_0$  for the next switching step. We define the final distribution as the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$ .

Now, we want to calculate statistical properties of the measured signal. In a confocal microscope, the signal typically is modeled as following a Poisson distribution with an average given by the spatial distribution of the fluorophores convolved with the effective point spread function of the microscope. In the RESOLFT microscope, after the switching steps, the fluorophores in the vicinity of the focus are prepared in the on-state with a probability given by the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$ . This is modeled as a Bernoulli process, with a different on-switching probability for each fluorophore according to their position. This leads to fluctuations in the signal, due to the variance of this Bernoulli process.

In the readout step, a confocal excitation and detection is applied for a short time. The average measured signal from a fluorophore in the on-state is given by the product of its molecular brightness  $b$  per observation time and the confocal PSF  $h_{\text{conf}}(\vec{r})$ . We assume the confocal PSF to be scaled such that  $h_{\text{conf}}(0) = 1$ . Thus, the measured brightness of a fluorophore at the focus is  $b$ . If the readout pulse is short enough such that further switching can be neglected, the measured signal is Poisson distributed. The number of photons  $k$  measured from a fluorophore in the on-state at the focus is thus given by the Poisson distribution:

$$P_{\text{Pois}}(k|b) = \frac{b^k}{k!} \exp(-b). \quad (2.22)$$

To calculate the variance of the signal measured in the RESOLFT microscope, we consider the signal of a single fluorophore at position  $\vec{r}_1$  relative to the focus. After the switching steps, the probability that it is switched to the on-state is given by the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r}_1)$ . In the on-state, it emits fluorescence which can be measured as Poisson counts with an average of  $b$  times the confocal PSF  $h_{\text{conf}}(\vec{r}_1)$ . In the off-state, no fluorescence is measured. Thus, the probability distribution for the number  $k$  of measured photons from a single fluorophore is given by

$$P(k|\vec{r}_1) = P_{\text{Pois}}(k|b \cdot h_{\text{conf}}(\vec{r}_1))\rho_{\text{on}}^{\text{eff}}(\vec{r}_1) + \delta(k=0)(1 - \rho_{\text{on}}^{\text{eff}}(\vec{r}_1)). \quad (2.23)$$

## 2. Theoretical modeling of superresolution microscopy

To find the mean and the variance of this distribution, we calculate the first and second moment:

$$\begin{aligned}
\mathbf{E}[k] &= \sum_{k=0}^{\infty} k P(k|\vec{r}_1) \\
&= \sum_{k=0}^{\infty} k \frac{(bh_{\text{conf}}(\vec{r}_1))^k}{k!} \exp(-bh_{\text{conf}}(\vec{r}_1)) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) \\
&= bh_{\text{conf}}(\vec{r}_1) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1)
\end{aligned} \tag{2.24}$$

$$\begin{aligned}
\mathbf{E}[k^2] &= \sum_{k=0}^{\infty} k^2 P(k|\vec{r}_1) \\
&= \sum_{k=0}^{\infty} k^2 \frac{(bh_{\text{conf}}(\vec{r}_1))^k}{k!} \exp(-bh_{\text{conf}}(\vec{r}_1)) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) \\
&= \sum_{k'=0}^{\infty} (k'+1) \frac{(bh_{\text{conf}}(\vec{r}_1))^{k'}}{k'!} \exp(-bh_{\text{conf}}(\vec{r}_1)) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) \\
&= b^2 h_{\text{conf}}(\vec{r}_1)^2 \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) + bh_{\text{conf}}(\vec{r}_1) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1).
\end{aligned} \tag{2.25}$$

The mean and the variance are then given by

$$\mu_1 = b \cdot h_{\text{conf}}(\vec{r}_1) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) \tag{2.26}$$

$$\sigma_1^2 = b \cdot h_{\text{conf}}(\vec{r}_1) \rho_{\text{on}}^{\text{eff}}(\vec{r}_1) (1 - \rho_{\text{on}}^{\text{eff}}(\vec{r}_1)). \tag{2.27}$$

Assuming that different fluorophores switch and emit independently, the mean and the variance of the measured signal at the scan position  $\vec{s}$  during the readout step for  $N$  emitters at positions  $\vec{r}_i$  is the sum of the mean and variance of the signals from each individual emitter:

$$\mu = b \sum_i^N h_{\text{conf}}(\vec{r}_i - \vec{s}) \rho_{\text{on}}^{\text{eff}}(\vec{r}_i - \vec{s}) \tag{2.28}$$

$$\sigma^2 = b \sum_i^N h_{\text{conf}}(\vec{r}_i - \vec{s}) \rho_{\text{on}}^{\text{eff}}(\vec{r}_i - \vec{s}) (1 + b \cdot h_{\text{conf}}(\vec{r}_i - \vec{s}) (1 - \rho_{\text{on}}^{\text{eff}}(\vec{r}_i - \vec{s}))). \tag{2.29}$$

This expression depends on the positions of all the fluorophores in the image. These positions are not known exactly. To be able to apply this expression to real image data, we write it in terms of a number density  $n(\vec{r})$  on a discretized grid. Using the scanning positions as the grid for the number density, we can write the mean and the variance for each image pixel as a convolution of the number density with different convolution kernels for the mean and the variance of the signal.

## 2. Theoretical modeling of superresolution microscopy

We define the convolution kernels as the effective RESOLFT PSF

$$h_{\text{eff}}(\vec{r}) = h_{\text{conf}}(\vec{r}) \cdot \rho_{\text{on}}^{\text{eff}}(\vec{r}) \quad (2.30)$$

and the excess variance PSF

$$h_{\text{exc}}(\vec{r}) = h_{\text{conf}}(\vec{r})^2 \rho_{\text{on}}^{\text{eff}}(\vec{r}) (1 - \rho_{\text{on}}^{\text{eff}}(\vec{r})). \quad (2.31)$$

The mean and variance of the signal measured in the RESOLFT image can thus be written as convolutions of the number density with the described convolution kernels:

$$\mu = b(n \otimes h_{\text{eff}}) + d \quad (2.32)$$

and

$$\sigma^2 = \mu + b^2(n \otimes h_{\text{exc}}). \quad (2.33)$$

Here, we included also a constant background  $d$ . Due to the large number of fluorophores from out of focus regions contributing to the background, and the large defocused PSF, it is reasonable to assume that it is homogenous over the field of view with a constant Poisson rate.

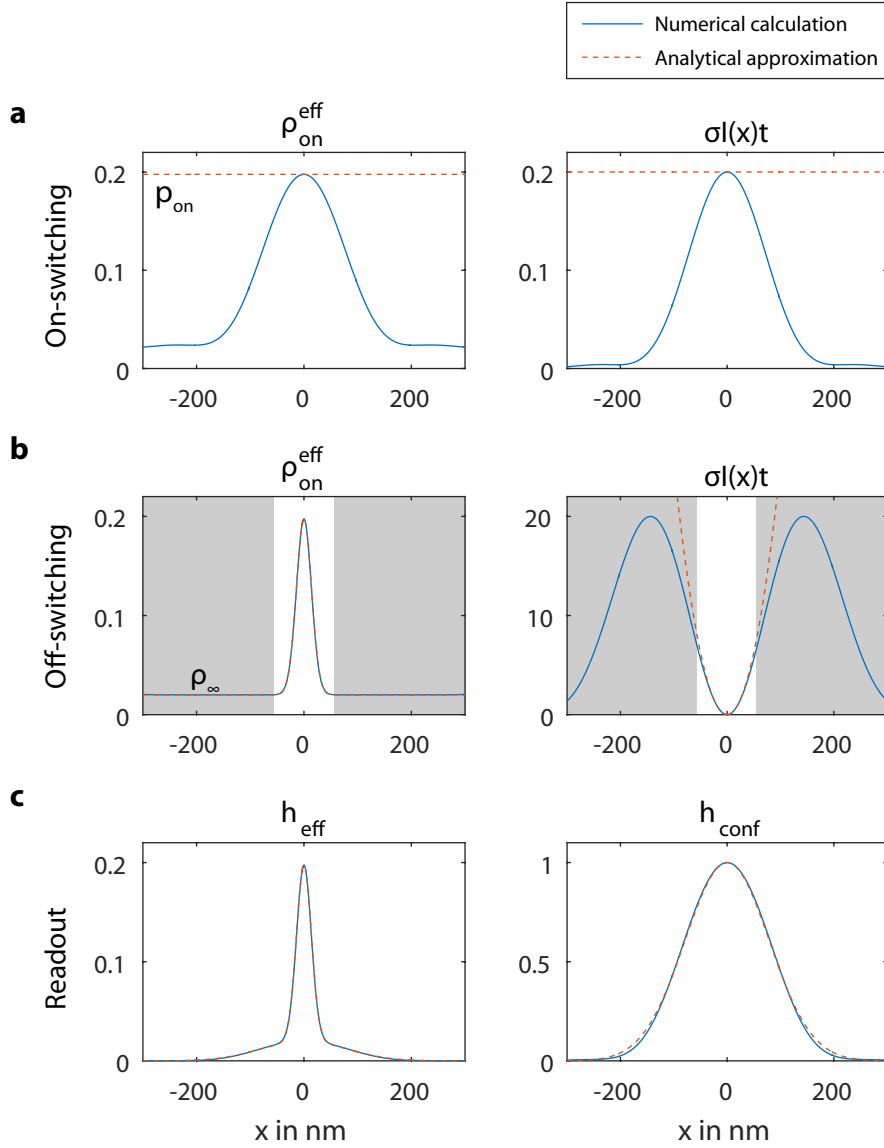
### 2.2.2. Effective RESOLFT PSF for linear photoswitching

Here, we want to calculate the effective PSF of a RESOLFT microscope under some simplifying considerations. We assume that after the on-switching step, all fluorophores are prepared in a random state with the same probability of  $p_{\text{on}}$  to be in the on-state. The second assumption is that during the off-switching step the behavior of the fluorophores can be described using a simple linear switching model, where both the on- and the off-switching rate depend linearly on the light intensity of the off-switching beam. The shape of the off-switching beam is approximated with a quadratic intensity profile. Finally, we approximate the confocal excitation and detection as a Gaussian PSF.

The on-switching step is typically performed with a diffraction limited PSF. However, the resolution enhancement is determined by the off-switching step. In the high intensity region around the focus, the fluorophores are switched into the equilibrium state, which is independent of the initial state. Thus, we can approximate the on-switching behavior in the first switching step with a constant value for the on-switching probability  $p_{\text{on}}$  for all fluorophores, as shown in Figure 2.2a.



2. Theoretical modeling of superresolution microscopy



**Figure 2.2. RESOLFT image formation.** Comparison of numerical calculation and analytical approximation of the PSF. **a:** Left: Probability of fluorophores residing in the on-state after the on-switching step. This is approximated by a constant on-switching probability  $p_{\text{on}}$ . Right: Applied effective switching dose. **b:** Left: In the off-switching step, a doughnut shaped intensity profile is used to switch off molecules in the periphery. In the regions of high intensity (shaded areas), the molecules switch to the equilibrium level  $\rho_{\infty}$ . Right: Applied effective switching dose. **c:** Left: Effective RESOLFT PSF  $h_{\text{eff}}$  after the confocal readout step. Right: Confocal PSF  $h_{\text{conf}}$ .

## 2. Theoretical modeling of superresolution microscopy

For the off-switching step, we consider linear photoswitching kinetics which can be characterized by the effective switching cross-sections  $\sigma_{\text{on}}$  and  $\sigma_{\text{off}}$ . At the wavelength  $\lambda_{\text{off}}$  of the off-switching laser, the fluorophores are driven linearly both from the on- to the off-state, and from the off- to the on-state, and there are no spontaneous switching rates between the states. The switching rates depending on the off-switching intensity  $I_{\text{off}}$  are given by

$$k_{\text{on}} = \sigma_{\text{on}} \frac{\lambda_{\text{off}}}{hc} I_{\text{off}} \quad (2.34)$$

and

$$k_{\text{off}} = \sigma_{\text{off}} \frac{\lambda_{\text{off}}}{hc} I_{\text{off}}. \quad (2.35)$$

In this simple model, the equilibrium level  $\rho_{\infty}$  is independent of the off-switching intensity  $I_{\text{off}}$ , and given by the ratio

$$\rho_{\infty} = \frac{\sigma_{\text{on}}}{\sigma_{\text{on}} + \sigma_{\text{off}}}. \quad (2.36)$$

The effective off-switching cross-section  $\sigma_{\text{off}}$  is assumed to be larger than the effective on-switching cross-section  $\sigma_{\text{on}}$ , such that  $\rho_{\infty}$  is small.

The effective switching rate  $k$  is proportional to the intensity, and we define the proportionality constant as  $\sigma$ :

$$\sigma = (\sigma_{\text{on}} + \sigma_{\text{off}}) \frac{\lambda_{\text{off}}}{hc}. \quad (2.37)$$

The effective switching rate is thus given by

$$k = \sigma I_{\text{off}}. \quad (2.38)$$

The off-switching step is usually performed until the fluorophores in the high intensity regions of the off-switching PSF are switched into equilibrium. Therefore, all fluorophores outside of the central region are assumed to be switched to the equilibrium state, and only fluorophores in the region which is well approximated by a parabola at the center remain in the on-state with a higher probability, as shown in Figure 2.2b. The off-switching intensity distribution can be approximated locally around the focus at  $\vec{r} = 0$  by a quadratic intensity profile, proportional to the illumination intensity  $I_0$ , with a steepness parameter  $\alpha$  which is determined by the exact shape of the applied intensity distribution [61]:

$$I_{\text{off}}(\vec{r}) \approx \alpha r^2 I_0. \quad (2.39)$$

The effective on-switching probability set up by the on- and the off-switching steps is found by inserting the initial probability  $p_{\text{on}}$ , the intensity dependent switching

## 2. Theoretical modeling of superresolution microscopy

rates given in Equation (2.38) and the quadratic intensity profile given in Equation (2.39) into Equation (2.21):

$$\rho_{\text{on}}^{\text{eff}} = \rho_{\infty} + (p_{\text{on}} - \rho_{\infty}) \exp(-\sigma \alpha r^2 I_0 t). \quad (2.40)$$

In the readout step, the fluorophores are read out confocally. We approximate the confocal PSF by a Gaussian PSF with the full width at half maximum (FWHM)  $w_{\text{conf}}$ :

$$h_{\text{conf}} = \exp(-4 \log(2) \frac{r^2}{w_{\text{conf}}^2}). \quad (2.41)$$

The effective RESOLFT PSF  $h_{\text{eff}}$ , shown in Figure 2.2c, is given by the product of the effective on switching distribution and the confocal readout PSF  $h_{\text{conf}}$ :

$$h_{\text{eff}} = \rho_{\infty} h_{\text{conf}} + (p_{\text{on}} - \rho_{\infty}) h_{\text{RESOLFT}}, \quad (2.42)$$

with

$$h_{\text{RESOLFT}} = \exp(-4 \log(2) \frac{r^2}{w_{\text{conf}}^2} - \sigma \alpha r^2 I_0 t) \quad (2.43)$$

The PSF thus consists of two terms, a term proportional to  $h_{\text{conf}}$  with confocal resolution, and a sharp central peak  $h_{\text{RESOLFT}}$ , with a sub-diffraction FWHM given by

$$w_{\text{RESOLFT}} = \frac{w_{\text{conf}}}{\sqrt{1 + w_{\text{conf}}^2 \alpha \sigma I_0 t}}. \quad (2.44)$$

The FWHM of the sharp central peak depends on the inverse square root of the radiation dose, the product of the illumination intensity  $I_0$  and the time  $t$  it is applied. The resolution can be improved increasing the radiation dose, either by using a higher intensity (within the range in which the linear model applies) or by employing longer switching times, which enables RESOLFT at low light levels.

The term with confocal resolution is determined by the switching kinetics of the photoswitch, namely by the equilibrium level or the switching background of the fluorophore. The contrast  $c$  can be defined as the ratio of the amplitudes of the two contributions to the PSF given in Equation (2.42):

$$c = \frac{p_{\text{on}} - \rho_{\infty}}{\rho_{\infty}}. \quad (2.45)$$

To model the effect of an imperfect off-switching intensity distribution, in which there is residual intensity at the focus, we can add a small constant  $\epsilon$  to the approximation given by Equation (2.39):

$$I_{\text{off}}(\vec{r}) = (\alpha r^2 + \epsilon) I_0. \quad (2.46)$$

## 2. Theoretical modeling of superresolution microscopy

In this case, although the FWHM of the central peak is the same as above, its amplitude decays exponentially:

$$h_{\text{eff}} = \rho_{\infty} h_{\text{conf}} + (p_{\text{on}} - \rho_{\infty}) h_{\text{RESOLFT}} \exp(-\sigma \epsilon I_0 t). \quad (2.47)$$

This reduces the signal from the central peak, but not from the confocal switching background, such that the contrast is decreased:

$$c = \frac{p_{\text{on}} - \rho_{\infty}}{\rho_{\infty}} \exp(-\sigma \epsilon I_0 t). \quad (2.48)$$

To maximize the contrast, it is therefore important to minimize imperfections in the optical system which can lead to a nonzero intensity at the center of the off-switching intensity distribution, and to maximize the initial on-switching probability  $p_{\text{on}}$ .

### 2.2.3. Counting molecules in RESOLFT microscopy

Our goal is to quantitatively estimate the number density  $n(\vec{r})$  in a RESOLFT image. Based on Equations (2.30) and (2.32), we need to calibrate

- the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}$ ,
- the shape of the confocal PSF  $h_{\text{conf}}$ ,
- the molecular brightness parameter  $b$ ,
- and the background  $d$ , which can typically be estimated from regions in the image devoid of signal.

#### Calibration of the effective on-switching probability and the confocal PSF

In the model described in Section 2.2.2, the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$  is given in Equation (2.40) by the constant equilibrium level  $\rho_{\infty}$  plus a Gaussian function with a sub-diffraction FWHM and the amplitude  $p_{\text{on}} - \rho_{\infty}$ . To calibrate the effective on-switching probability, in principle, the off-switching intensity distribution in the experiment and the switching kinetics of the fluorophore must be known. However, by estimating the effective PSF from the measured image data, the effective on-switching probability can be calibrated as well. In the analytical approximation, the effective PSF of the RESOLFT microscope is the product of the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$  and the confocal PSF  $h_{\text{conf}}(\vec{r})$ . The confocal PSF can be calculated using vectorial diffraction theory [62]. By measuring the effective PSF from the image, the parameters of the two Gaussian components, namely the FWHM of the two components and the ratio between the amplitudes can

## 2. Theoretical modeling of superresolution microscopy

be used to calibrate the parameters in the expression for the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$  given in Equation 2.40. In Section 2.2.4, we develop a method to estimate the effective PSF from an image, given that the structure is known to consist of sparse point- or line-like structures, which can be localized in the image.

### Estimation of the molecular brightness $b$

Equation (2.33) shows that the variance of the noise is larger than the variance of pure Poisson noise by an excess variance factor due to the stochastic switching of the fluorophores. The excess variance depends on the brightness  $b$  quadratically and has a spatial structure, as it is given by the number density  $n(\vec{r})$  convolved with a convolution kernel which depends on the confocal PSF  $h_{\text{conf}}(\vec{r})$  and the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$ . This can be used to calibrate the brightness of the fluorophores in an image as we show in Section 2.2.5.

### 2.2.4. Deconvolution and measurement of the effective PSF

The effective PSF of a RESOLFT microscope depends on the switching kinetics and the illumination light distribution and doses given in the experiment. It would be convenient to calibrate these in the experiment itself, for instance by estimating the PSF from the image data.

Although the structure underlying an image is usually not known, for sparse, well-separated point- or line-like objects, we can estimate it without knowing the PSF using a heuristic method. After smoothing the image with a Gaussian mask to reduce the influence of the noise, we find local maxima for sparse point-like structures, or locally maximum lines for sparse line-like structures. This heuristic estimate of the object can be used to estimate the PSF using the Richardson-Lucy deconvolution algorithm [63, 64]. This algorithm was first introduced to reconstruct the underlying structure in images blurred by a known PSF and degraded by Poisson noise, but as the convolution is commutative, we can as well apply it to estimate the PSF from an image given a known structure.

### Modeling the structure using local maxima

Assuming that the image is known to consist of well separated point objects, it is possible to localize these objects more precisely than the size of the PSF. Exact knowledge of the PSF is not needed, because the positions of the emitters can also be found by taking the position of the pixel with the maximum signal, or, for sub-pixel accuracy, calculating the center of mass of the intensity distribution.

## 2. Theoretical modeling of superresolution microscopy

A similar method can be applied for finding lines in an image. If the structure is known to consist of lines which are thinner than the size of the PSF, they can be localized without knowing the PSF by finding the locally maximal lines in the image. For each  $3 \times 3$  pixel array in the image, a filtering method which returns the locally maximal lines in an image can be defined by returning one in the case that the signal in the central pixel is among the three largest signals in the pixel array and zero otherwise.

For noisy image data, the noise has to be suppressed first, for example by smoothing the image with a Gaussian. To be able to extract lines from an image in this way, they have to be clearly separated, such that their images are still resolved in the smoothed image. This filtering method is layed out in Figure 2.3a-c. If the brightness of the lines varies strongly over the image, this can be taken into account by multiplying the result of the filter with the smoothed image.

### Richardson-Lucy deconvolution for known structure

The Richardson-Lucy deconvolution is based on the Poisson likelihood for an image. The image is modeled as an unknown object  $\rho(\vec{r})$  convolved with a known PSF  $h(\vec{r})$ , with an additional constant background  $d$ . We define the noise-free image model  $m_i$  in each pixel  $i$  of the image corresponding to a scan position  $\vec{s}_i$  as

$$m_i = (\rho \otimes h)_i + d \quad (2.49)$$

$$= \int d^n r \rho(\vec{r}) h(\vec{s}_i - \vec{r}) + d. \quad (2.50)$$

For the numerical calculation, we discretize the object  $\rho_j = \rho(\vec{r}_j)$  and the point spread function  $h_{ij} = h(\vec{s}_i - \vec{r}_j)$  in the image space such that

$$m_i = \sum_j \rho_j h_{ij} + d. \quad (2.51)$$

The measured counts  $\{k_i\}$  in each pixel  $i$  are degraded by Poisson noise, such that the logarithm of the likelihood function can be written as

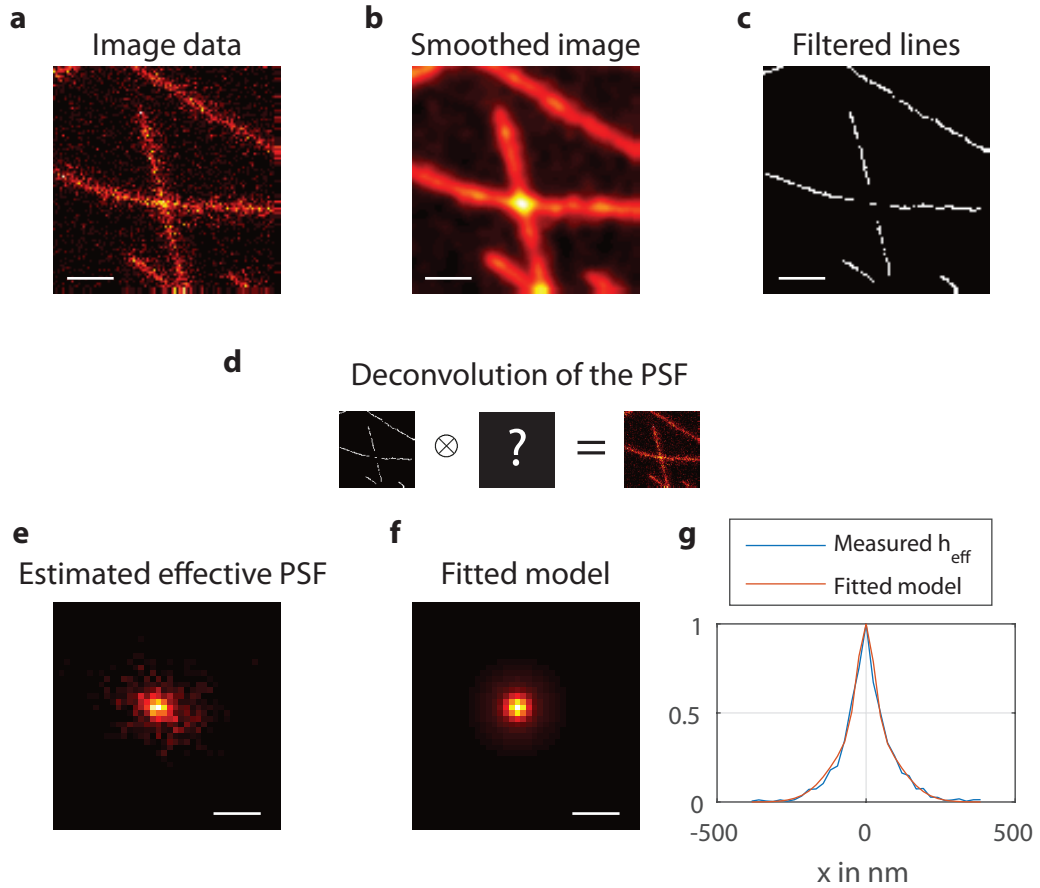
$$l(\{k_i\}|\rho, h, d) = \sum_i (k_i \log(m_i) - m_i) + \text{const}. \quad (2.52)$$

To find the maximum likelihood estimate for the structure at pixel  $l$ , we take the derivative with respect to  $\rho_l$ :

$$\frac{\partial L}{\partial \rho_l} = \sum_i \left( \frac{k_i}{m_i} - 1 \right) \frac{\partial m_i}{\partial \rho_l} \quad (2.53)$$

$$= \sum_i \left( \frac{k_i}{m_i} - 1 \right) h_{il}. \quad (2.54)$$

2. Theoretical modeling of superresolution microscopy



**Figure 2.3. Schematic of deconvolution of the effective PSF  $h_{\text{eff}}$ .** **a-c:** From the raw data **a**, lines are extracted by first smoothing the image using a Gaussian filter **b**, and then using a filter **c** which returns one if a pixel is among the three largest pixels in each  $3 \times 3$  pixel region. **d:** This model is used to extract the effective PSF using a Richardson-Lucy deconvolution. **e-g:** From this estimated PSF **e**, the parameters of the PSF model given by Equation (2.42) can be extracted by fitting the model to the data. **f-g.** Scale bars: Top row: 500 nm, Bottom row: 200 nm

## 2. Theoretical modeling of superresolution microscopy

Setting this derivative to zero, we get the necessary condition for the maximum likelihood solution

$$\sum_i \frac{k_i}{m_i} h_{il} = \sum_i h_{il}. \quad (2.55)$$

This can be written using the cross correlation  $\frac{k}{m} * h$  of the ratiometric deviation of the signal to the model with the PSF.

$$\left( \frac{k}{m} * h \right)_l = \sum_i h_{il}. \quad (2.56)$$

Thus we find that a multiplicative fixed-point iteration for the estimated object  $\hat{\rho}$  is given by

$$\hat{\rho}_i^{\text{iter}+1} = \hat{\rho}_i^{\text{iter}} \cdot \left( \frac{k}{\hat{\rho}^{\text{iter}} \otimes h + d} * h \right)_i / \sum_j h_j. \quad (2.57)$$

This procedure guarantees that the estimated structure will be positive, because if it is initialized with positive values for  $\hat{\rho}_i^0$ , the estimate will remain positive.

As the convolution operator commutes, we can directly write down the corresponding iteration procedure for the case of the unknown PSF with a known underlying structure  $\rho$ . This is given by

$$\hat{h}_i^{\text{iter}+1} = \hat{h}_i^{\text{iter}} \cdot \left( \frac{k}{\hat{h}^{\text{iter}} \otimes \rho + d} * \rho \right)_i / \sum_j \rho_j. \quad (2.58)$$

To extract the PSF from an image with sparse lines, we can now use the line extraction method described above to fix the structure  $\rho$ , and deconvolve the PSF by starting from a uniform distribution and iterating according to Equation (2.58), as shown in Figure 2.3d-g.

### 2.2.5. Estimation of the brightness and the number density

In Section 2.2.4, we showed how to estimate the shape of the effective PSF. Given that we calibrated the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$ , and that the shape of the confocal readout PSF  $h_{\text{conf}}(\vec{r})$  is known, the remaining unknown quantities in the stochastic image formation model described in Section 2.2.1 are the brightness parameter  $b$  and the number density  $n(\vec{r})$  underlying the image.

Here, we describe a method to estimate the brightness parameter  $b$  from a single image. The main idea is to make use of the excess variance of the signal, calculated in Equation (2.33), which is proportional to the square of the brightness parameter. If several independent images can be measured, the sample variance could also be



## 2. Theoretical modeling of superresolution microscopy

estimated by taking the pixel-wise mean square deviation from the average of the images.

One way to calculate the variance from only a single image, is to first estimate the mean signal by taking into account that the signal varies slowly due to the convolution with the known effective PSF. To estimate the variance, we then consider the deviations of the measured signal in the image pixels to the estimated mean. This can be achieved by using the Richardson-Lucy algorithm described in Section 2.2.4 to deconvolve the image with the calibrated effective RESOLFT PSF  $h_{\text{eff}}$ . The result of the deconvolution, the estimated object  $\hat{\rho}$ , is an estimate of the product of the number density  $n(\vec{r})$  and the brightness  $b$ :

$$\hat{\rho}(\vec{r}) = b \cdot n(\vec{r}) . \quad (2.59)$$

Convolving this result again with the effective RESOLFT PSF  $h_{\text{eff}}$  results in a smooth image with strongly reduced noise, which we use as the estimate for the mean signal in the RESOLFT image given in equation (2.32):

$$\hat{\mu} = \hat{\rho} \otimes h_{\text{eff}} + d . \quad (2.60)$$

In Equation (2.33), the variance of the signal measured in a RESOLFT image was calculated. It is given by the mean signal, as expected for Poisson noise, plus an excess variance term, which is due to the stochastic switching. The excess variance is given by the convolution of the number density  $n(\vec{r})$  with the excess variance PSF  $h_{\text{exc}}$  defined in Equation (2.31), times the square of the brightness  $b$ . The spatial distribution of the excess variance  $v$  can also be estimated using the estimated object  $\hat{\rho}$ :

$$\hat{v} = \hat{\rho} \otimes h_{\text{exc}} , \quad (2.61)$$

such that the variance of the image can be written as

$$\sigma^2 = \hat{\mu} + b\hat{v} . \quad (2.62)$$

Using the estimated mean signal  $\hat{\mu}$  and the estimated spatial distribution of the excess variance  $\hat{v}$ , we now approximate the likelihood function of the brightness parameter  $b$ , given the image data, with the likelihood function for a Gaussian distribution with the mean and the variance estimated in Equations (2.60) and (2.62). The log-likelihood function of the parameters for the mean  $\mu$  and the variance  $\sigma^2$  for the Gaussian distribution is given by

$$l(\mu, \sigma|x) = -\frac{(x - \mu)^2}{2\sigma^2} - \frac{1}{2} \log(\sigma^2) + \text{const} . \quad (2.63)$$

## 2. Theoretical modeling of superresolution microscopy

Using the estimated mean  $\hat{\mu}$  and variance  $\hat{\sigma}$ , we can write down the log-likelihood function for the brightness parameter  $b$  given the measured signals  $k_i$  in each pixel  $i$  in the Gaussian approximation:

$$l(b|k_i, \hat{\mu}_i, \hat{v}_i) = - \sum_i^N \left( \frac{(k_i - \hat{\mu}_i)^2}{2(\hat{\mu}_i + b\hat{v}_i)} + \frac{1}{2} \log(\hat{\mu}_i + b\hat{v}_i) \right) + \text{const}. \quad (2.64)$$

This likelihood function can be maximized with respect to  $b$  to estimate the brightness:

$$\hat{b} = \arg \max_b l(b|k_i, \hat{\mu}_i, \hat{v}_i). \quad (2.65)$$

A quantitative estimate of the number density  $\hat{n}(\vec{r})$  is then given by the estimated object, given in equation (2.59), divided by the estimated brightness:

$$\hat{n}(\vec{r}) = \frac{\hat{\rho}(\vec{r})}{\hat{b}}. \quad (2.66)$$

### 2.3. **Single molecule localization microscopy**

While two emitters which are closer than the diffraction limit emitting at the same time cannot be resolved, if they can be switched to an off-state sequentially, they can be localized at different points in time. Several physical processes can lead to a stochastic switching behavior that fulfills the condition that the fluorophores reside in the off-state for most of the time [10–15]. In this case, a superresolved image can be reconstructed by localizing single emitters, which emit light at different times, one by one.

The switching behavior of single emitters which are suitable for SMLM can be modeled in the simplest case as a two state Markov model, as described in Section 2.1.1, where the on-switching rate is much smaller than the off-switching rate, typically by more than three orders of magnitude. The on-state dwell times are typically on the order of the duration of the camera exposure time of 1-100 milliseconds, and the off-state dwell times are on the order of the duration of the whole experiment, which can take minutes and up to hours in some cases.

To measure the on-switching rate, single molecule experiments have to be analyzed such that the finite length of the dataset does not bias the estimate, as described in Section 2.1.3. To measure the on-state dwell times of the single molecules, we developed a method to estimate the transition time with a sub-frame timing precision. This is achieved by analyzing the signal of single molecule events frame by frame. Modeling the measured signal as photons emitted with a constant rate during the on-state, the transition time can be estimated by taking the ratio of the signal in the transition frame to the signal in a frame in which the molecule was in the on-state during the whole exposure.

The two-state switching kinetics can then be applied to model the imaging process of activation-based multicolor STORM microscopy [42,50]. Here, two or more color channels can be discriminated by activating specifically activator-reporter dye pairs with different activator dyes for each channel. The crosstalk between the color channels in this method can be on the order of 10 to 20 % [42]. Sources of crosstalk include the cross-activation, where the reporter dye belonging to the wrong species is activated by the activation light, and random activation, where a random blinking event is assigned to a color channel because it happened to be activated at the wrong time. We show that crosstalk due to randomly switching fluorophores can be reduced by applying sub-frame single molecule event timing analysis.

### 2.3.1. Modeling of single molecule kinetics data and estimation of the on-switching rate

To estimate the on- and the off-switching rates from single molecule data, we want to extract the dwell times of the molecules in the off- and in the on-state. The density of single molecules has to be low enough such that single molecules can be uniquely identified. However, a lower density of molecules means that a lower number of events can be measured, which reduces the precision of the estimation of the switching rates. Here, we present a method of extracting the signal traces for single molecule data of a relatively high density, in which single molecules can be clearly separated in the single molecule localization image but their images overlap in the widefield data because they are closer together than the diffraction limit.

To this end, the dataset is initially analyzed by a STORM analysis algorithm, localizing the positions of the blinking fluorophores in each image [11,65]. The resulting list of localizations is binned in a two-dimensional histogram. Single molecules are identified as clusters of localizations in this histogram. We extract the positions of these clusters by localizing again the position of the cluster in the superresolved image, as visualized in Figure 2.4.

The PSF of the widefield images of the single molecules can be well approximated by a two-dimensional Gaussian function with the FWHM  $w$  given by

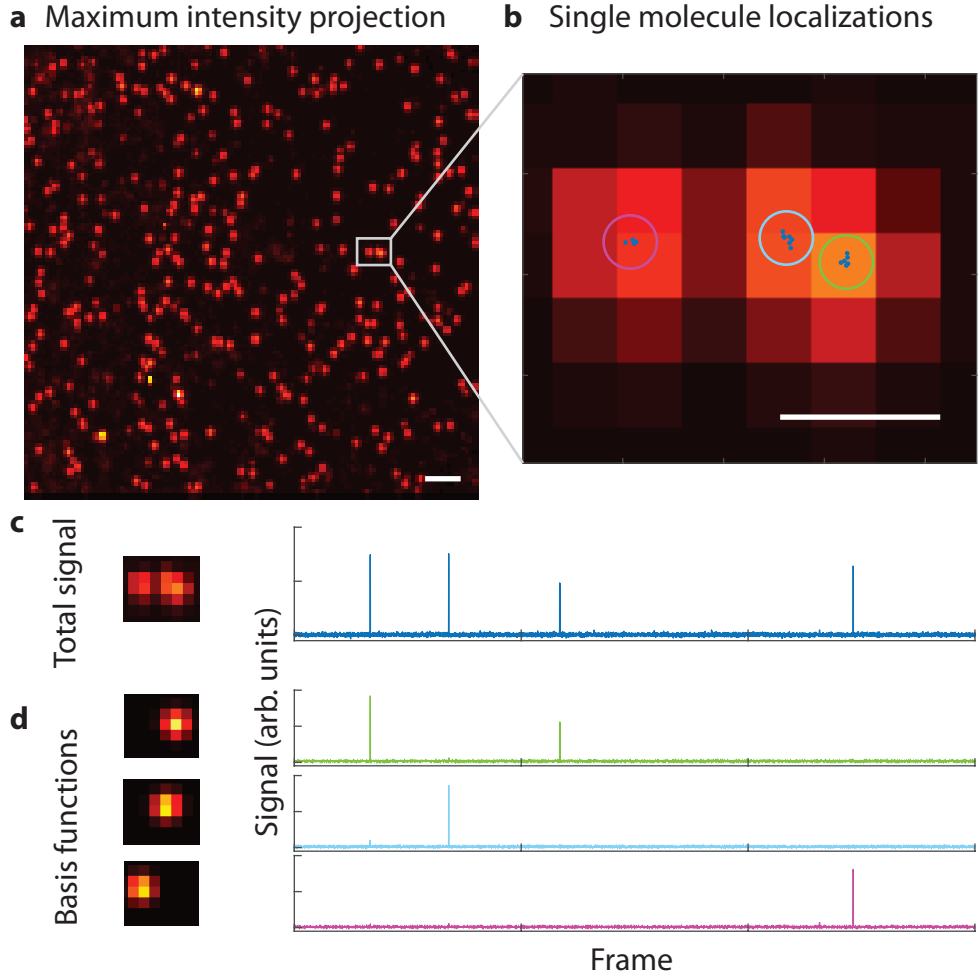
$$f(\vec{r}) = A \exp(-4 \log(2) \frac{r^2}{w^2}). \quad (2.67)$$

The single molecule dataset, given by the measured counts  $\{k_{i,t}\}$  in pixel  $i$  at position  $\vec{r}_i$  and frame  $t$  can be modeled as a constant background in each frame plus a superposition of Gaussian point spread functions at the positions  $\vec{r}_j$  of the single molecules in the image as basis functions:

$$m_{i,t} = d_t + \sum_{j=1}^N b_{jt} f(\vec{r}_j - \vec{r}_i) \quad (2.68)$$

The parameters of this linear model are the brightness of each single molecule in each frame  $b_{jt}$  and the constant background  $d$ . We can extract these parameters using regularized linear least squares optimization [66]. The estimation of the parameters can be written as the solution to a system of linear equations, which can be efficiently calculated. This assumes implicitly that the noise of the signal in each pixel is normally distributed.

If the distance between two molecules is below the diffraction limit, their point spread functions overlap. Due to the background noise, the least squares error



**Figure 2.4. Linear model for trace extraction.** **a:** The diffraction-limited images of single molecules in single molecule data may overlap, as seen in the maximum intensity projection of a movie of blinking single molecules. Scale bar,  $2 \mu\text{m}$ . **b:** The single molecule events are localized (blue dots) and clusters of events (circles) are identified. Scale bar,  $500 \text{ nm}$ . **c:** The signal from the region shown in **b** is the superposition of the signals from the single molecules in the region. **d:** To extract the signal from each single molecule, a set of basis functions is constructed. The basis function for each single molecule is given by a Gaussian PSF at the center of mass of each cluster. The single molecule signals can be extracted by estimating the parameters of the linear model using regularized least squares.

## 2. Theoretical modeling of superresolution microscopy

function may have a lower value if the brightness parameters of the two molecules are fitted as two very large numbers with a different sign. This overfitting effect can be reduced by regularizing the error function. By adding a term proportional to the sum of the squares of the parameters, which is also called weight decay regularization, the influence of overfitting can be strongly reduced. By defining the basis functions  $\phi_j(\vec{r}_i) = f(\vec{r}_j - \vec{r}_i)$  for  $j = 1..N$  and  $\phi_0(\vec{r}_i) \equiv 1$ , with  $b_{0t} = d_t$ , the regularized cost function is given by

$$L_t(\vec{b}_t) = \frac{1}{2} \sum_i \left( \sum_{j=0}^N b_{jt} \phi_j(\vec{r}_i) - k_{i,t} \right)^2 + \frac{1}{2} \lambda \sum_j b_{jt}^2, \quad (2.69)$$

where  $\lambda$  is the positive regularization parameter. The parameters minimizing this cost function are the solution of a linear system of equations for each frame  $t$ ,

$$(A + \lambda) \cdot \vec{b}_t = \vec{c}_t \quad (2.70)$$

where the elements of the matrix  $A$  are the overlap between the pairs of basis functions

$$A_{lj} = \sum_i \phi_l(\vec{r}_i) \phi_j(\vec{r}_i) \quad (2.71)$$

and the vector  $c$  is given by the weighted sum of the data

$$c_{jt} = \sum_i k_{i,t} \phi_j(\vec{r}_i). \quad (2.72)$$

The resulting single molecule traces  $b_{jt}$  for  $j \geq 1$  are the time-dependent brightness for each single molecule in the model. It is thus possible to extract the signal from single molecules which are not resolved in the diffraction-limited image, as shown in Figure 2.4. From the single molecule signal traces we want to extract the off- and on-state dwell times, which are used to estimate the on- and off-switching rates  $k_{\text{on}}$  and  $k_{\text{off}}$ .

### Estimation of the on-switching rate

To estimate the on-switching rate of the single molecules, we have to take into account the finite length of the dataset. Fluorescent dyes optimized for STORM microscopy exhibit off-state dwell times which are more than three orders of magnitude longer than the on-state dwell times [51]. Therefore, the duration of the measurement  $T$  is often of the same order of magnitude as the average off-state dwell times of the single molecules. This can bias the result of the estimation, as described in Section 2.1.3. Here, we approach this problem by only measuring the first off-time

## 2. Theoretical modeling of superresolution microscopy

for each single molecule, that is the time  $t_i$  until the first reappearance of the  $i$ th single molecule in the dataset, that is the first frame with a brightness higher than a threshold.

The on-switching rate  $k_{\text{on}}$  can then be estimated using Bayes' theorem for the posterior distribution given in Equation (2.11) with the likelihood function defined in Equation (2.10), where the upper boundary of the interval is given by the length of the dataset  $T$ , and the lower boundary is set to 0. The posterior probability distribution  $P(k_{\text{on}}|\{t_i\})$  in this case is proportional to

$$P(k_{\text{on}}|\{t_i\}) \propto \left( \prod_i^N \frac{k_{\text{on}} \exp(-k_{\text{on}} t_i)}{1 - \exp(-k_{\text{on}} T)} \right) P(k) \quad (2.73)$$

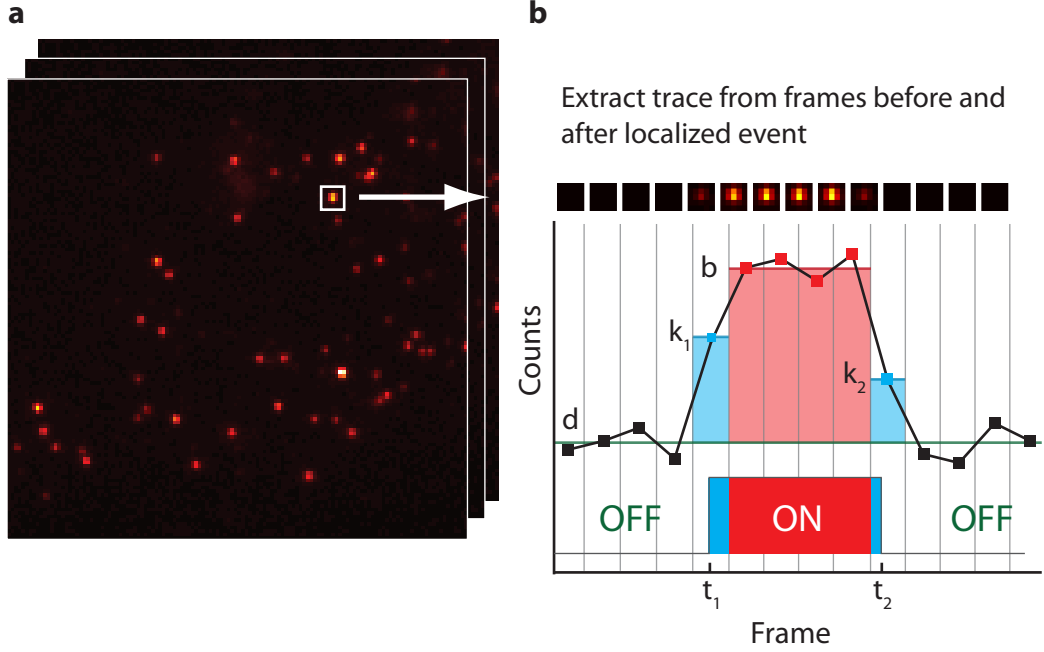
Using a uniform prior  $P(k)$  on a given interval, we calculate numerically the mean and the variance of the posterior distribution to estimate the on-switching rate and the error bars.

### 2.3.2. Single molecule event timing and estimation of the off-switching rate

To estimate the off-switching rate as described in Section 2.1.3, we first need to extract the on-state dwell times from the single molecule data. The time resolution of the camera is limited by a minimum exposure time, which is typically on the order of one millisecond. As the single molecule events happen randomly, the transition from the off- to the on-state can happen at any time during the exposure of the camera. After the transition, the single molecule resides in the on-state for a time which follows an exponential distribution governed by the off-switching rate and switches off again.

We want to measure the on-state dwell time accurately with a precision higher than the frame rate of the camera. The single molecule signal trace is extracted from the raw data, for example using the linear model described in Section 2.3.1. We assume, that the single molecule emits photons with a constant rate of  $b$  per frame while residing in the on-state. The average signal measured in each frame is then given by the fraction of time the single molecule spends in the on-state during each frame times  $b$ .

For events which are detected in at least three consecutive frames, such that they are in the on-state for at least one full frame, the transition times can be measured with a precision higher than the frame rate of the camera. The on-switching time is measured by taking the ratio of the signal in the first frame to the average emission rate, and the off-switching time accordingly by taking the ratio of the signal in the



**Figure 2.5. Single molecule event timing.** **a:** From the raw data, the signal trace for each single molecule event is extracted. **b:** By taking the ratio of the signal in the transition frames  $k_1$  and  $k_2$  to the brightness  $b$  over the background  $d$ , measured in frames in which the molecule was on for the whole time, the switching times  $t_1$  and  $t_2$  can be estimated.

last frame to the average emission rate. The average emission rate  $b$  can be estimated from the frame(s) in which it is switched on for the whole frame. This procedure is sketched in Figure 2.5.

### Switching time estimator for Poissonian noise

We assume that in the on-state, the molecule emits photons with a constant known rate  $b$  per frame, and that there is a constant known rate of background photons  $d$  per frame. The measured counts  $k_1$  in the first frame, in which the molecule switches on at time  $t_1$  relative to the beginning of the frame are then Poisson distributed

$$P(k_1|t_1, b, d) = P_{\text{Pois}}((1 - t_1)b + d). \quad (2.74)$$

To estimate the switching time from the signal using Bayes' theorem, given by Equation (2.11), we now assume that the single molecule event is known to start during the particular camera frame under consideration. This is formalized by using a uniform prior distribution in the interval  $[0, 1]$  for  $t_1$ . The posterior distribution is



## 2. Theoretical modeling of superresolution microscopy

then proportional to

$$P(t_1|s_1, b, d) \propto ((1 - t_1)b + d)^{k_1} \exp(-((1 - t_1)b + d)) . \quad (2.75)$$

Thus, we find the maximum a posteriori estimator for  $t_1$  to be

$$\hat{t}_1 = 1 - \frac{k_1 - d}{b} . \quad (2.76)$$

The same method is applied to estimate the off-switching transition time  $t_2$  from the signal  $k_2$  in the last frame of the trace.

$$\hat{t}_2 = \frac{k_2 - d}{b} . \quad (2.77)$$

These estimators are equal to the maximum likelihood estimators, as the prior distribution was chosen to be constant over the given interval. We can calculate the Cramer-Rao lower bound for variance of the estimator  $t_1$ :

$$\text{Var}(\hat{t}_1) \geq \frac{b(1 - t_1) + d}{b^2} . \quad (2.78)$$

This shows an interesting property of the transition time estimation. The variance of the estimator depends on the switching time, and is lower for events which started later in the first frame. The reason for this dependence is, that as we take a ratio between two numbers, only the absolute width of the probability distribution of the signal determines the variance of the estimator. For larger  $t_1$ , the number of counts measured in the first frame is small. For the Poisson distribution, although the relative width of the distribution is inversely proportional to the square root of the mean, the absolute width is directly proportional to the square root of the mean and thus smaller for smaller signals.

In calculating the estimator, we assumed that the frame during which the transition takes place can be identified correctly and the brightness  $b$  and the background  $d$  were known precisely. In practice, both of these parameters and the transition frames also have to be estimated from the data.

### Hidden Markov model and Viterbi algorithm

To estimate the transition time of a single molecule event using the estimator described in Equations (2.76) and (2.77), we need to be able to identify the transition frames of the single molecule event. To this end, we employ a hidden Markov model (HMM) [67,68]. This model is based on the discrete time Markov chain discussed in Section 2.1.1. In addition to the state space and the transition matrix, in the hidden Markov model, we define emission probability distributions, which govern the data

## 2. Theoretical modeling of superresolution microscopy

$k_t$  measured during each time step  $t$  depending on the underlying “hidden” state of the model  $s_t$  in that time step.

We define an HMM with four hidden states, depicted in Figure 2.6, the off-state, a transition state from the off- to the on-state, the on-state, and a transition state from the on- to the off-state, with the transition matrix  $T_{ij} = P(s_{t+1} = j | s_t = i)$ :

$$T = \begin{pmatrix} 1 - p_{\text{on}} & p_{\text{on}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 - p_{\text{off}} & p_{\text{off}} \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (2.79)$$

The emission probability distributions are chosen as Gaussian distributions with equal standard deviations  $\sigma$ , centered around the minimum of the data for the off-state, the maximum of the data for the on-state and the center between both for the transition states.

Although these probability distributions could be chosen to model the process more realistically, for example using Poisson distributions, in practice the signal from single molecules is not stable over time, as for example triplet state blinking may happen on timescales shorter than the frame rate [69]. These effects, which also limit the precision of the timing measurements will be investigated in Section 4.2.2.

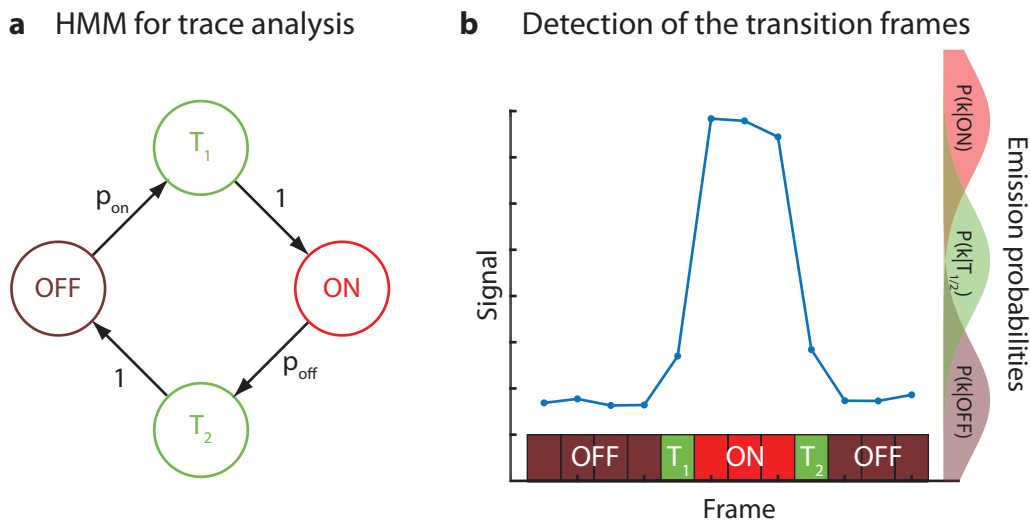
To infer the most likely sequence of states and thus to detect the transition frames, we use the Viterbi algorithm [70,71]. This algorithm is based on calculating the likelihood for only the most probable sequence of states by taking the Markov property into account. We first write down the probability distribution for the joint distribution of the sequence of states  $\{s_t\}$  and the emissions  $\{k_t\}$  at each time step  $t = 1 \dots N$ :

$$P(\{s_t\}, \{k_t\}) = P(s_1)P(k_1|s_1) \prod_{i=2}^N P(s_i|s_{i-1})P(k_i|s_i). \quad (2.80)$$

Here, in addition to the emission probability distributions  $P(k_t|s_t)$  and the transition probabilities  $P(s_i|s_{i-1})$  given in the matrix  $T$ , we need to define the initial state distribution given by  $P(s_1)$ . Now, to find the most likely sequence of states given the data, we want to maximize the conditional probability

$$P(\{s_t\}|\{k_t\}) = \frac{P(\{s_t\}, \{k_t\})}{P(\{k_t\})}. \quad (2.81)$$

As the term in the denominator is independent of  $\{s_t\}$ , this is the same as maximizing Equation (2.80) directly. By going through the data points sequentially, we calculate



**Figure 2.6. Hidden Markov model (HMM) for identifying the transition frames.** **a:** The model used to identify the transition states consists of four states: The on- and off-state, and transitions states in between. The probability of leaving the transitions states is 1. **b:** On the right-hand side, the emission probability distributions for the states are shown. The result of the Viterbi algorithm for estimating the most probable sequence of states is shown below.

first the probability of the most likely path which ends up in each state, and second, keep track of the state from which the last transition occurred. After doing this for each time step, the most likely sequence of states is found by going back through the list of states, starting from the final state with the highest likelihood.

### Estimation of the off-switching rate

Having identified the first and the last frame of each single molecule trace, we can now estimate the switching times using the ratiometric estimators given by Equation (2.76) for the on-switching time and Equation (2.77) for the off-switching time and calculate the on-state dwell time for each event by taking their difference. The switching times can only be estimated for events which are detected in at least three consecutive frames. For events which only emit signal in one or two frames, the Viterbi algorithm based on the four state model shown in Figure 2.6 is not able to identify the transition frames correctly. For these events however, the measured on-state dwell time is always less than two full frames. All events for which the measured on-state dwell time is longer than two full frames are detected in at least

## 2. Theoretical modeling of superresolution microscopy

three frames. The off-switching rate can thus be estimated from all the measured times which are longer than this lower threshold using the estimator (2.17).

### 2.3.3. Modeling of activation-based multicolor STORM

The image quality of SMLM critically depends on the switching kinetics of the dyes used for labeling the structure. The localization of single molecule events is only possible if the density of active fluorophores in each image is low enough. At the same time, the density of fluorophores which label the sample has to be high enough, otherwise the structure is not well represented by the labelling, and appears dotted in the reconstructed image [72]. The switching kinetics of the dye are modeled by the two state Markov chain. Several processes can contribute to the on- and off-rates, as shown in Figure 2.7b. The average density of fluorophores in the on-state at any given time is given by the total density of fluorophores  $n(\vec{r})$  times the probability  $\rho_{\text{on}}$  that each fluorophore is in the on-state, given by Equation (2.6). Typically, the equilibrium between the on- and the off-state is reached shortly after the beginning of the experiment, such that we can consider only the equilibrium level or duty cycle of the photoswitch  $\rho_{\infty}$ .

In addition to the molecules which are on at the beginning of each frame, the signal from single molecules which switch on during the exposure time is integrated onto the detector. The probability that a molecule switches from the off- to the on-state during the exposure time  $\Delta t$  of the camera is approximately given by the on-rate  $k_{\text{on}}$  times  $\Delta t$ . Thus, the average number density of fluorophores appearing in each camera frame is given by

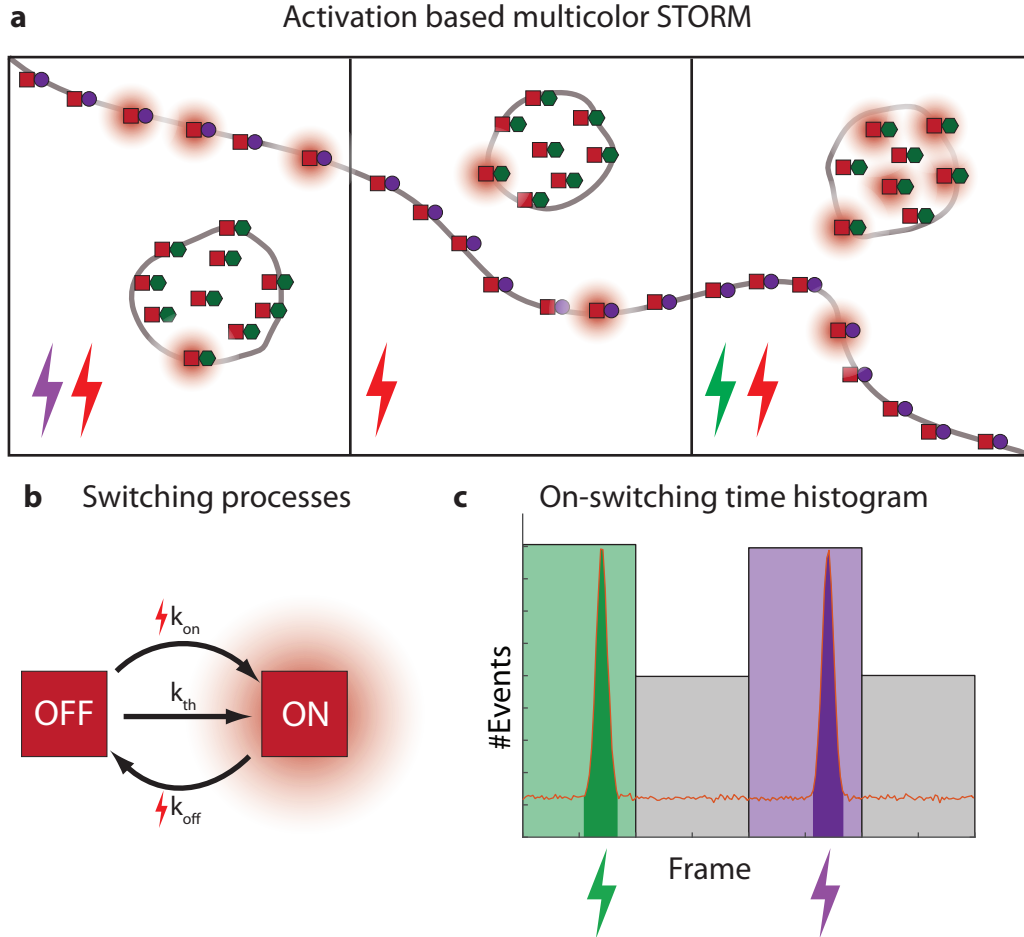
$$n_{\text{on}}(\vec{r}) = (\rho_{\infty} + k_{\text{on}}\Delta t) n(\vec{r}). \quad (2.82)$$

We denote the highest density for which the localization of single molecules is possible as  $n_{\text{max}}$ . It depends on the size and shape of the point spread function and the algorithm used for localization. For Gaussian point spread functions and single emitter fitting algorithms, this maximum density is usually around  $1 \mu\text{m}^{-2}$  [73]. The density of fluorophores in the on-state has to be lower than this maximum density such that single molecule localization is possible:

$$n_{\text{on}} < n_{\text{max}}. \quad (2.83)$$

One way to discriminate different color species in STORM microscopy is to use activator-reporter dye pairs [42, 50]. Here, the reporter molecule is the same for all color species. The reporter dye is paired in close proximity with an activator dye,

2. Theoretical modeling of superresolution microscopy



**Figure 2.7. Activation-based multicolor STORM.** **a:** The sample is labeled with activator-reporter dye pairs of two different species and excited with red laser light. Left: Using an ultraviolet pulse, the species with the UV activator dye is predominantly activated. Center: In frames without an activation, random activation can occur. Right: Using a green laser pulse, the species with the green activator dye is predominantly activated. **b:** State transition model for the reporter dye: A light driven off-switching rate  $k_{\text{off}}$  competes with a spontaneous on-switching rate  $k_{\text{th}}$  and a light driven on-switching rate  $k_{\text{on}}$ . **c:** Histogram of estimated on-switching transition times. Compared to the frame-based assignment of the events to the color channels, the number of randomly activated events in the channels can be reduced using sub-frame switching time estimation.

## 2. Theoretical modeling of superresolution microscopy

which can be excited using laser pulses of different colors, which in turn causes the reporter dye to switch to the on-state, as visualized in Figure 2.7a. Events which were first detected in the frame in which the activation pulse is given are assigned to the corresponding color species. However, all events which were activated randomly during that frame are also assigned to the same color species. This leads to crosstalk between the color channels, as some of the randomly activated molecules belong to the wrong species. Another source of crosstalk which can occur is cross-activation, where due to the spectral overlap of the excitation spectra of the activator dyes, a laser pulse also activates fluorophores of the wrong color with a low probability.

To estimate the amount of crosstalk due to random activation, we include an additional specific activation term in the model for the imaging process given in Equation (2.82). The power and duration of the activation laser pulse can be chosen such that the probability of activating a molecule in the off-state is  $p_{\text{act}}$ . This activation increases the number of molecules in the on-state in a frame with a specific activation pulse to

$$n_{\text{on}}(\vec{r}) = (\rho_{\infty} + k_{\text{on}}\Delta t + p_{\text{act}}) n(\vec{r}). \quad (2.84)$$

As this density still has to fulfill the condition (2.83), the activation probability is limited to

$$p_{\text{act}}(\vec{r}) < \frac{n_{\text{max}}}{n(\vec{r})} - (\rho_{\infty} + k_{\text{on}}\Delta t). \quad (2.85)$$

As the widefield activation illumination has the same intensity over the whole field of view, the activation probability is limited by the highest local density of fluorophores in the image,  $\max(n(\vec{r}))$ . Only events which are detected for the first time in the frame with the activation pulse are assigned to each color. The crosstalk is given by the number of molecules which randomly switch on during that frame. As part of the randomly activated events belong to the correct species, an upper limit for the crosstalk  $x$  due to random activation is the ratio of randomly activated events over the total number of activations in the frame:

$$x < \frac{k_{\text{on}}\Delta t}{k_{\text{on}}\Delta t + p_{\text{act}}}. \quad (2.86)$$

Choosing the activation probability of the pulse as high as possible according to Equation (2.85), the crosstalk can be reduced to

$$x < \frac{k_{\text{on}}\Delta t}{\frac{n_{\text{max}}}{\max(n(\vec{r}))} - \rho_{\infty}}. \quad (2.87)$$

This crosstalk is proportional to the exposure time of the camera, which is effectively the time during which randomly activated events are integrated into the specific color channel. Using the sub-frame switching time estimation method developed in Section 2.3.2, the crosstalk can be reduced.

### Histogram of switching times

The distribution of measured on-switching times in an activation-based multicolor STORM experiment is a sharp peak centered at the time of the activation pulse on a constant background of random activations, as shown in Figure 2.7c. Depending on the timing precision of the events in the single molecule experiment, a time interval  $\Delta t_{\text{timing}}$  shorter than the camera exposure time  $\Delta t$  can be chosen to assign the events to each channel. Thus only events which were activated by the activation pulse with a high probability are assigned to each color channel. This reduces the number of randomly activated events in each channel by the ratio of the timing interval to the exposure time of the camera.

The probability  $P(\text{act}|t_1)$  that an event with a measured switching time  $t_1$  was activated specifically by the activation laser pulse can again be calculated using Bayes' theorem.

$$P(\text{act}|t_1) = \frac{P(t_1|\text{act})P(\text{act})}{P(t_1|\text{act})P(\text{act}) + P(t_1|\text{rand})P(\text{rand})}. \quad (2.88)$$

Although randomly activated molecules can be assumed to activate at uniformly distributed times during a camera frame, the probability  $P(t_1|\text{rand})$  of the measured switching time given a random activation may not be uniform. The reason for this discrepancy is, that the variance of the estimated switching time depends on the switching time itself, as was shown in Equation (2.78). In practice, due to the estimation of the first frame of the trace using the hidden Markov model described in Section 2.3.2, additional artifacts can be induced if the first frame is not correctly identified. This may happen for switching times close to the boundary between two camera frames, or if the event is not long enough.

Due to the large number of events in a typical STORM dataset, both the distribution of the randomly activated events  $P(t_1|\text{rand})P(\text{rand})$  and the distribution of the specifically activated events  $P(t_1|\text{act})P(\text{act})$  can be estimated by binning the measured switching times as a histogram. These distributions are used to assign the probability given in Equation (2.88) to each event. Thus, events which were not specifically activated with a high probability can be identified and rejected, and the crosstalk in the final image due to random activation can be reduced.

### 3. Quantitative RESOLFT microscopy

In Section 2.2, we presented a stochastic imaging model for RESOLFT microscopy and developed a method to estimate the number of fluorophores in the image. The switching model used to describe the behavior of the fluorophores is based on a two state Markov chain model. The RESOLFT imaging process consists of three steps, the on-switching of the fluorophores, the off-switching using an intensity distribution with a zero used to confine the fluorophores which remain in the on-state to a sub-diffraction sized region, and the confocal readout. During the off-switching step, the switching rates between the on- and the off-state are both modeled to depend linearly on the intensity.

The fluorescent proteins rsEGFP and rsEGFP2 are widely used rsFPs in RESOLFT microscopy [17, 18, 20]. To find if the proposed switching model captures the switching behavior for rsEGFP and rsEGFP2, we measured the switching rates depending on the intensity. Using a setup with an extended illumination spot, it was made sure that the light intensity was constant over the observed volume.

The proposed method for counting molecules in a RESOLFT image depends on the accurate calibration of the effective PSF. In numerical simulations, we simulated images according to our forward model with known PSFs. As the ground truth is known, we can estimate how the error of the estimation of the brightness scales, independent of the additional errors which could be introduced by the calibration of the PSF.

Finally, the method was applied to RESOLFT images of microtubules in *Drosophila* embryos. To be applicable, the effective PSF was estimated from the image itself using the deconvolution procedure described in Section 2.2.4. From this PSF, the effective on-switching probability was calibrated, taking into account the measured switching kinetics.



### 3.1. Switching kinetics measurements

The model for the effective RESOLFT PSF described in Section 2.2.2 is based on a two state model in which both the on-switching rate  $k_{\text{on}}$  and the off-switching rate  $k_{\text{off}}$  of the stochastic switching of the fluorophore during the off-switching step depend linearly on the intensity. This dependence is characterized by the effective cross-section for the on-switching transition  $\sigma_{\text{on}}$  and the off-switching transition  $\sigma_{\text{off}}$ .

We measured the ensemble switching behavior of the reversibly switchable fluorescent proteins rsEGFP and rsEGFP2 depending on the intensity of the off-switching light with  $\lambda_{\text{off}} = 488$  nm using the widefield switching procedure. For an ensemble of fluorophores which switch stochastically, the signal is expected to be proportional to the on-state probability, described in Equation (2.6):

$$s(t) \propto \rho_{\infty} + (\rho_0 - \rho_{\infty}) \exp(-kt). \quad (3.1)$$

The effective switching rate  $k$  and the equilibrium level  $\rho_{\infty}$  can be directly measured by fitting an exponential curve to the measured signal, given that the initial state  $\rho_0$  is known. The on- and off-switching rates can be calculated from the measured parameters  $k$  and  $\rho_{\infty}$  as

$$k_{\text{on}} = \rho_{\infty} k \quad (3.2)$$

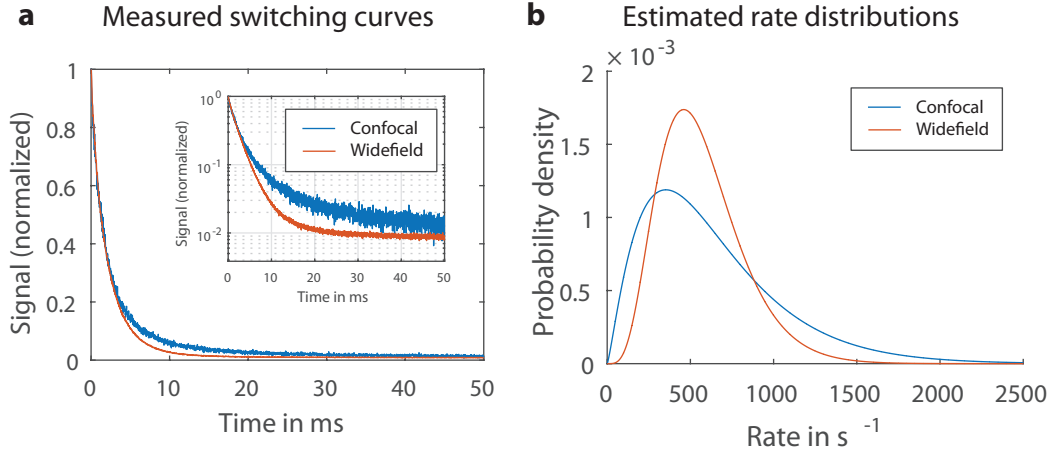
and

$$k_{\text{off}} = k - k_{\text{on}}. \quad (3.3)$$

However, in a confocal microscope, fluorophores contributing to the detected signal are subject to different illumination intensities, depending on the position of the fluorophore relative to the focus. As the switching rates depend on the intensity, the average signal measured when switching fluorophores using a confocal spot is a superposition of a distribution of rates. The shape of this distribution is determined by the intensity profile of the confocal spot and the density of the fluorophores, which may also vary spatially.

This is crucial for measuring the equilibrium level  $\rho_{\infty}$ , as there is a large volume in which the intensity of the confocal illumination is low, such that the relaxation to the equilibrium state takes a long time. In Figure 3.1, switching curves measured with a confocal illumination spot are compared to switching curves measured with homogeneous illumination. By fitting the model given by Equation (2.19) to the switching curves, we can extract the parameters of the distribution of rates. For the confocal switching signal, the distribution is broadened by the inhomogeneous intensity pattern. Therefore, to measure intensity dependent switching rates, this

### 3. Quantitative RESOLFT microscopy



**Figure 3.1. Switching curve for widefield and confocal illumination profiles.** **a:** switching curves are shown for confocal and widefield illumination on a linear scale. Inset: the same data is shown on a logarithmic scale. **b:** Estimated distribution of rates from fitting the curves with the model given by Equation (2.19).

effect has to be either modeled precisely or avoided experimentally. By illuminating the sample with a widened intensity distribution, which is homogeneous over the detection volume of the microscope, all detected fluorophores are exposed to the same intensity, such that the switching kinetics can be measured directly.

After avoiding artifacts due to the focal intensity distribution, the signal would be expected to follow an exponential decay, given in Equation (3.1). However, the decay seems to fall off more slowly after an initial steep decline. Thus, the decay model given by equation (2.19) with an amplitude  $a$  and a constant background level  $b$  was used to model the data:

$$s(t) = a \frac{\beta^\alpha}{(\beta + t)^\alpha} + b. \quad (3.4)$$

By fitting the parameters  $a$ ,  $b$ ,  $\alpha$  and  $\beta$  to the data using nonlinear least squares fitting, the average switching rate  $k$  can be calculated for each intensity

$$k = \frac{\alpha}{\beta} \quad (3.5)$$

and the equilibrium level is given by

$$\rho_\infty = \frac{b}{a + b}. \quad (3.6)$$

### 3. Quantitative RESOLFT microscopy

#### 3.1.1. Preparation of the initial state

We want to make sure that all proteins are in the on-state initially, that is  $\rho_0 = 1$ . To this end, samples of E. Coli cells expressing rsEGFP and rsEGFP2 were exposed to a broad intensity distribution with a FWHM of approximately  $17 \mu\text{m}$  of an ultraviolet laser beam ( $\lambda_{\text{on}} = 375 \text{ nm}$ ) for a time interval  $t_{\text{on}}$ . The fluorescence signal after the UV illumination was read out using an intensity distribution of the same size using blue light ( $\lambda_{\text{off}} = 488 \text{ nm}$ ), whereby the proteins switch to the equilibrium state  $\rho_{\infty}$ .

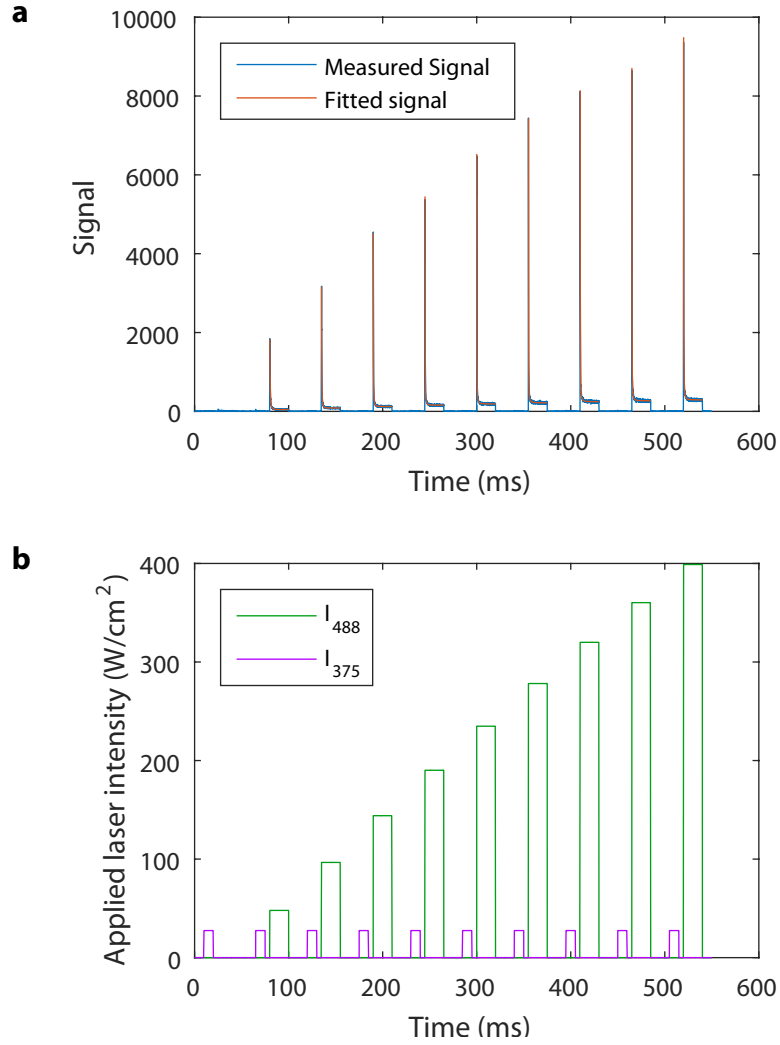
The time during which the UV light was applied was increased until the signal at the beginning of the switching curve reached a maximum. For rsEGFP,  $t_{\text{on}} = 25 \text{ ms}$  was chosen. For rsEGFP2, the used activation time was  $t_{\text{on}} = 10 \text{ ms}$ . In typical RESOLFT images, using a diffraction limited UV spot, the duration of the on-switching pulse given is on the order of  $10\text{-}30 \mu\text{s}$ , that is approximately 1000 times shorter [17, 18, 20]. The long on-switching time was chosen to maximize the probability that the fluorophores reside in the on-state at the beginning of each switching cycle, that is  $\rho_0$  is close to one.

#### 3.1.2. Switching kinetics for rsEGFP and rsEGFP2

To measure the on- and off-switching rates for the two proteins under the influence of blue light, switching cycles were conducted by first applying activation pulses to make sure that all proteins are prepared in the on-state ( $\rho_0 \cong 1$ ). This was achieved using long on-switching times  $t_{\text{on}}$ , as described in Section 3.1.1. Then blue light was applied for 20 ms, until the fluorophores in the sample were switched to the equilibrium state  $\rho_{\infty}$ . To make sure that the switching rates are measured under similar conditions, at each position in the sample, a linear ramp of 10 different illumination intensities was applied, as shown in Figure 3.2. This experiment was repeated using powers ranging over more than two orders of magnitude.

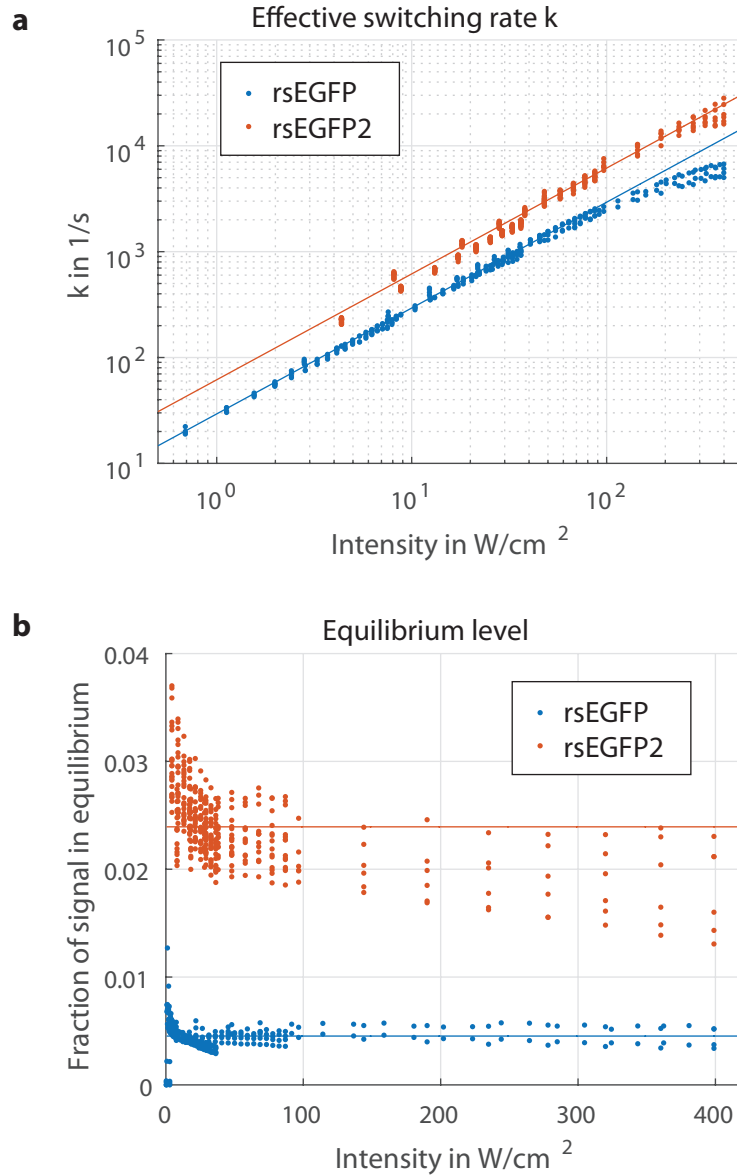
The results of the estimation of  $k$  and  $\rho_{\infty}$  are shown in Figure 3.3. The equilibrium level  $\rho_{\infty}$  does not seem to depend strongly on the intensity. The average equilibrium level  $\rho_{\infty}$  was measured to be 0.45% for rsEGFP and 2.4% for rsEGFP2. For both photoswitchable proteins, both the on- and the off-rate depend linearly on the intensity over a large range of intensities. For rsEGFP, the effective switching rate seems to saturate above approximately  $100 \text{ W/cm}^2$ .

### 3. Quantitative RESOLFT microscopy



**Figure 3.2. Switching kinetics data for rsEGFP2.** **a:** Switching signals measured with 488nm illumination light of different intensities. Each switching signal was fitted using the decay law for a distribution of rates. **b:** Applied intensity for the on-switching light ( $\lambda_{\text{on}} = 375\text{nm}$ ) and the off-switching light ( $\lambda_{\text{off}} = 488\text{nm}$ ). Data shown for the reversibly switchable fluorescent proteins rsEGFP2 expressed in E.Coli cells. Sample preparation and measurement was carried out by Philipp Alt.

### 3. Quantitative RESOLFT microscopy



**Figure 3.3. Switching kinetics for rsEGFP and rsEGFP2 expressed in E.Coli cells.** **a:** Effective switching rate  $k$  measured as a function of the intensity of the off-switching light ( $\lambda_{\text{off}} = 488\text{nm}$ ). A linear dependence of  $k$  on the intensity is fitted for intensities below  $100 W/cm^2$ . **b:** Equilibrium level  $\rho_{\infty}$  as a function of the intensity. The line shown is the average of the data for all intensities. Sample preparation and measurement was carried out by Philipp Alt.

### 3. Quantitative RESOLFT microscopy

By estimating the proportionality constant for this linear dependence, as described in Section 2.1 we can estimate the effective cross-sections for the on- and off-switching transition for rsEGFP:

$$\sigma_{\text{eff}}^{\text{on}} = 5.3 \times 10^{-20} \text{cm}^2 \quad (3.7)$$

and

$$\sigma_{\text{eff}}^{\text{off}} = 1.2 \times 10^{-17} \text{cm}^2. \quad (3.8)$$

For rsEGFP2, both effective cross-sections for on- and the off-switching are larger:

$$\sigma_{\text{eff}}^{\text{on}} = 5.7 \times 10^{-19} \text{cm}^2 \quad (3.9)$$

and

$$\sigma_{\text{eff}}^{\text{off}} = 2.5 \times 10^{-17} \text{cm}^2. \quad (3.10)$$

## 3.2. Counting molecules in simulated images

Using the forward model for the stochastic image formation process described in Section 2.2.1, RESOLFT images were simulated to test the workflow of counting molecules described in Section 2.2.3. To this end, point-like and line-like structures were simulated to find how precisely the brightness and the number can be measured, and which quantities determine the error of the estimate.

As we use the Richardson-Lucy deconvolution for the estimation of the object, the estimate of the brightness may be subject to bias and artifacts. The parameters in the simulation can be chosen freely, to determine for example a minimum brightness which is needed to estimate the number of fluorophores in the image precisely.

In the simulation, the structure was given by a number density in two dimensions on a  $1 \times 1 \mu\text{m}$  grid with 2 nm spacing. We simulated two kinds of structures: clusters of 50 molecules and single lines with 500 evenly distributed molecules. We assume the light intensity distributions in the focus to be known exactly, using numerically calculated intensity distributions of the diffraction limited on-switching, the doughnut-shaped off-switching and the confocal readout [62]. These light distributions are used to calculate the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$ , according to Equation 2.21. We define the peak dose  $D = \sigma I_{\text{max}} t$  to describe the switching parameters in the simulation, such that the final state after switching is given by

$$\rho_{\text{on}}(\vec{r}) = \rho_{\infty} + (\rho_0 - \rho_{\infty}) \exp(-Dh(\vec{r})), \quad (3.11)$$

where the applied intensity distribution  $h(\vec{r})$  is assumed to be normalized such that the maximum value is one. Two different cases were simulated, the saturated on-switching case, in which applied peak dose was  $D_{\text{on}} = 2$  and the low on-switching

### 3. Quantitative RESOLFT microscopy

case with  $D_{\text{on}} = 0.1$ . The resulting on-switching probability at the center of the focus in the two cases was 86.4 % and 9.5 %, respectively. For the off-switching step, the peak dose was set to  $D_{\text{off}} = 20$  in both cases, and the equilibrium level was assumed to be at 2%, approximately the level measured for rsEGFP2. The FWHM of the sharp RESOLFT peak after the two switching steps was about 40 nm.

To simulate the measurement, at each scanning position, all molecules in the sample are prepared in a random state according to the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$  by drawing binomially distributed random numbers at each position.

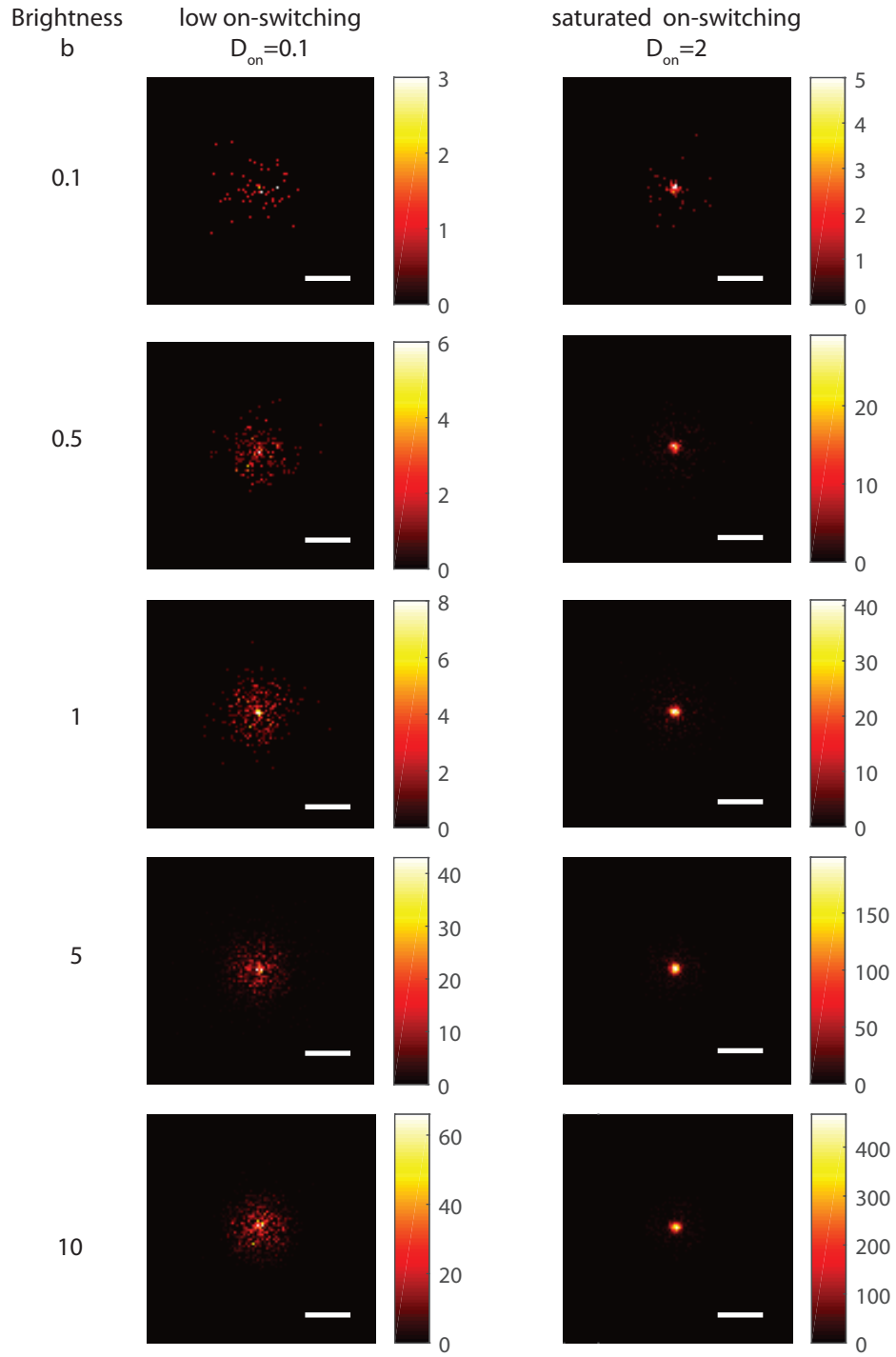
The measured signal at each scanning position is drawn from a Poisson distribution with the mean value calculated by integrating the distribution of molecules in the on-state weighted by the confocal readout PSF, and multiplied by the brightness for brightness parameters varying from 0.1 to 10. Simulated images for each of the parameter sets are shown in Figure 3.4 for clusters of molecules and in Figure 3.5 for line structures.

To estimate the number from the simulated images, we follow the procedure described in Section 2.2.5 to estimate the brightness from the image, given that the effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}(\vec{r})$  and the confocal PSF  $h_{\text{conf}}(\vec{r})$  are known.

In the first step, the image is deconvolved with the known effective PSF, which estimates the object  $\rho(\vec{r})$ , given by product of the number density  $n(\vec{r})$  and the brightness  $b$ . This estimated object is used to estimate the mean signal in the image and the spatial distribution of the excess variance. In the second step, the likelihood function for the parameter  $b$  is approximated using the likelihood for the Gaussian distribution, given by equation (2.64). This is a function of the parameter  $b$ , and the maximum of the likelihood function is calculated numerically to find a point estimate for the brightness  $\hat{b}$ . Finally, to estimate the number of fluorophores in the image, the estimated object is divided by the estimated brightness parameter  $\hat{b}$ . This is an estimate of the number density of fluorophores in the image. The total number of molecules in the image is then obtained by integrating this density over the whole image.

The results of this estimation procedure for the brightness parameter under the simulated conditions are shown in Figure 3.6.

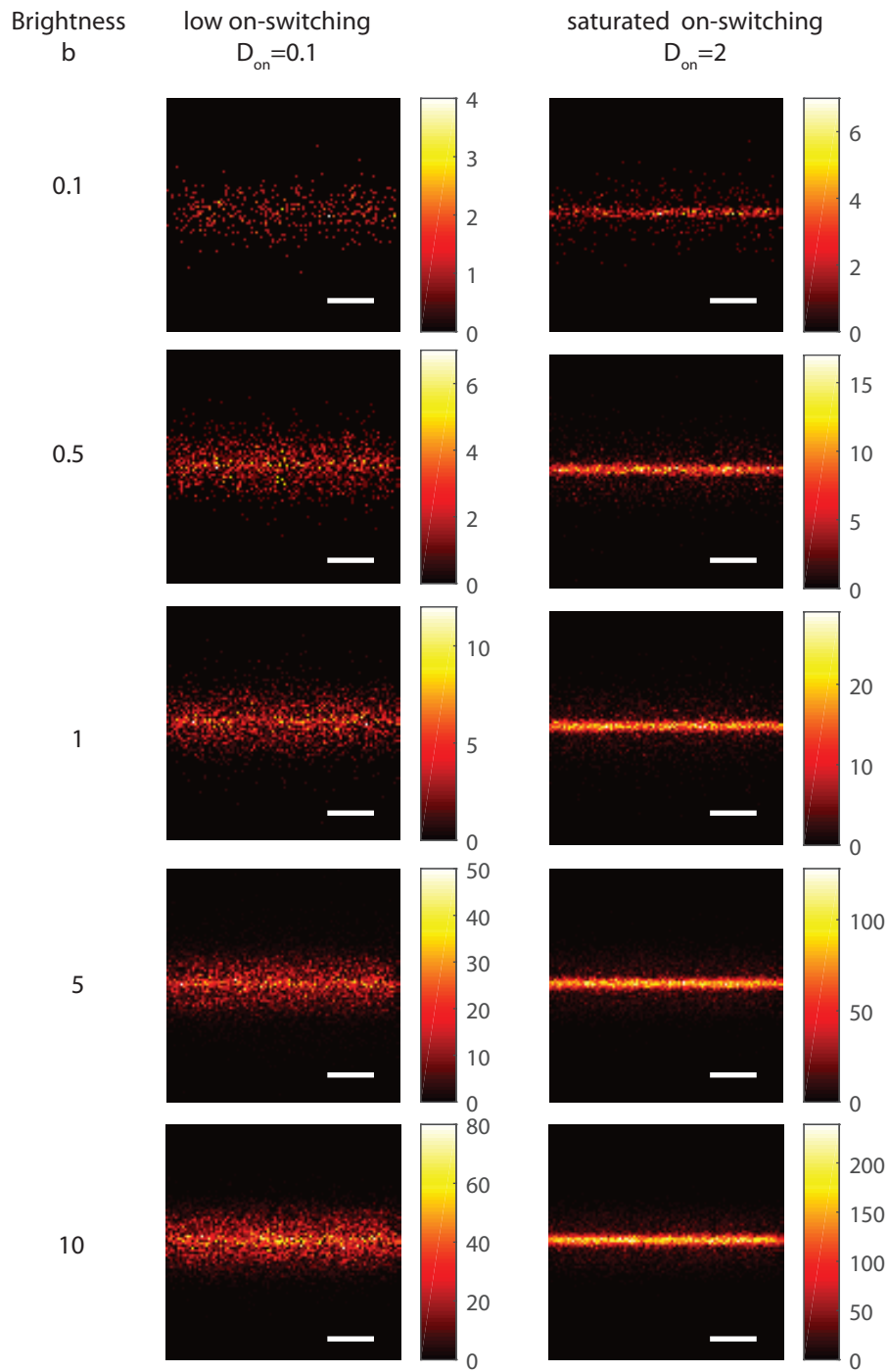
### 3. Quantitative RESOLFT microscopy



**Figure 3.4.** Simulated images of isolated clusters of rsEGFP2. The switching process was simulated according to the two-state Markov model for rsEGFP2 for different molecular brightness values  $b$ . Left: Low on-switching dose. Right: Saturated on-switching dose. Scale bars, 200nm

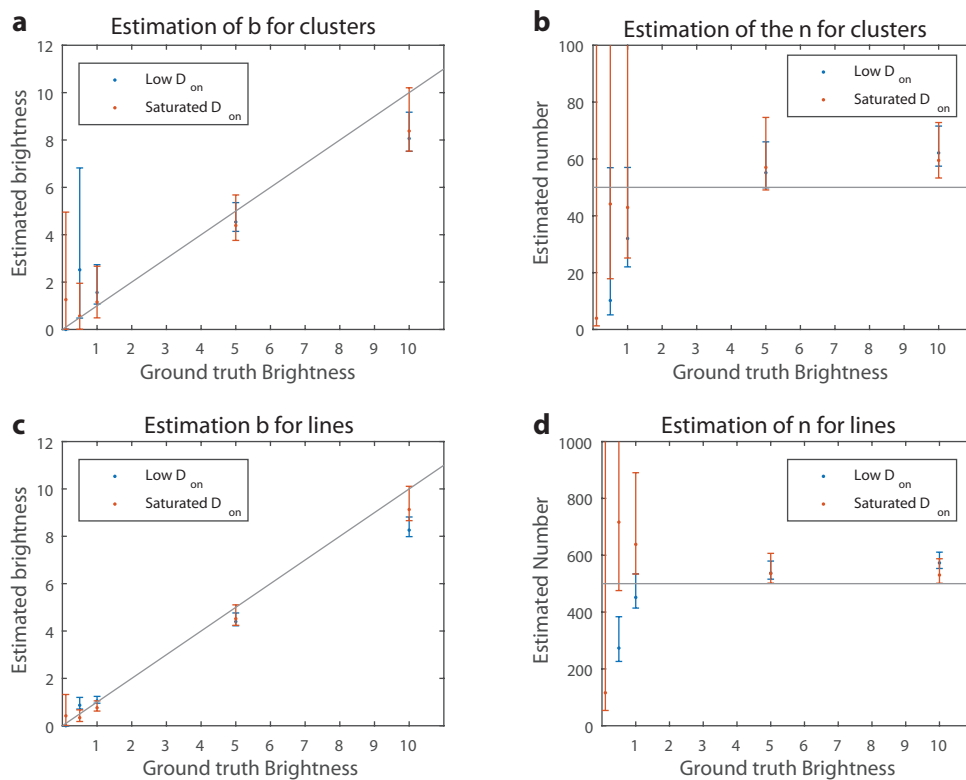


### 3. Quantitative RESOLFT microscopy



**Figure 3.5. Simulated images of isolated lines of rsEGFP2.** The switching process was simulated according to the two-state Markov model for rsEGFP2 for different molecular brightness values  $b$ . Left: Low on-switching dose. Right: Saturated on-switching dose. Scale bars, 200nm

### 3. Quantitative RESOLFT microscopy



**Figure 3.6. Parameter estimation for  $b$  in simulated data.** The ground truth of the simulation is shown as grey lines in all cases. **a:** Estimated brightness for clusters of 50 molecules. **b:** Estimated number for clusters of 50 molecules. **c:** Estimated brightness for lines of 500 molecules. **d:** Estimated number for lines of 500 molecules. Exemplary data for clusters shown in Figure 3.4 and for lines shown in Figure 3.5. Error bars are given as 5 and 95% quantile of the estimated brightness for 1000 simulated images each.

### 3.3. Experimental results in RESOLFT images

To validate the workflow described in Section 2.2.3, we use RESOLFT images of *Drosophila* embryo muscle tissue using the same microscope and samples as described in [20]. All cells in these flies express ubiquitously rsEGFP2 fused to  $\alpha$ -tubulin.

Tubulin is known to form helices with a diameter of approximately 25 nm. Each turn comprises of 13 dimers of  $\alpha$ - and  $\beta$ -tubulin and is spaced 8 nm apart [74, 75]. The total density of  $\alpha$ -tubulin along a single filament can thus be estimated to be approximately 1625 per  $\mu\text{m}$ .

Using Western blots of larval protein extracts, the ratio of endogenous  $\alpha$ -tubulin to rsEGFP2- $\alpha$ -tubulin in the *Drosophila* embryos was measured to be about 3:1 [20]. Thus, we expect the number density of rsEGFP2 proteins localized along a single microtubule fiber to be on the order of 400 per  $\mu\text{m}$ .

#### 3.3.1. Estimation of the effective PSF and calibration of the effective on-switching probability

By means of the line extraction filter, we build a model for the structure, which can be used to estimate the PSF using the Richardson-Lucy deconvolution, as described in section 2.2.4. The background value is estimated by averaging the signal in regions without any structure in it, which are identified by smoothing the image with a Gaussian function with a FWHM of 120 nm (5 pixels) and applying a threshold to the result. The average background counts per pixel are thus estimated to be

$$d \cong 0.35 \quad (3.12)$$

Figure 3.7a-d shows the result of the filtering procedure and the deconvolved effective PSF. Using nonlinear least-squares fitting, we fit the estimated PSF according to the analytical model, given in Equation (2.42), with a superposition of two Gaussian PSFs:

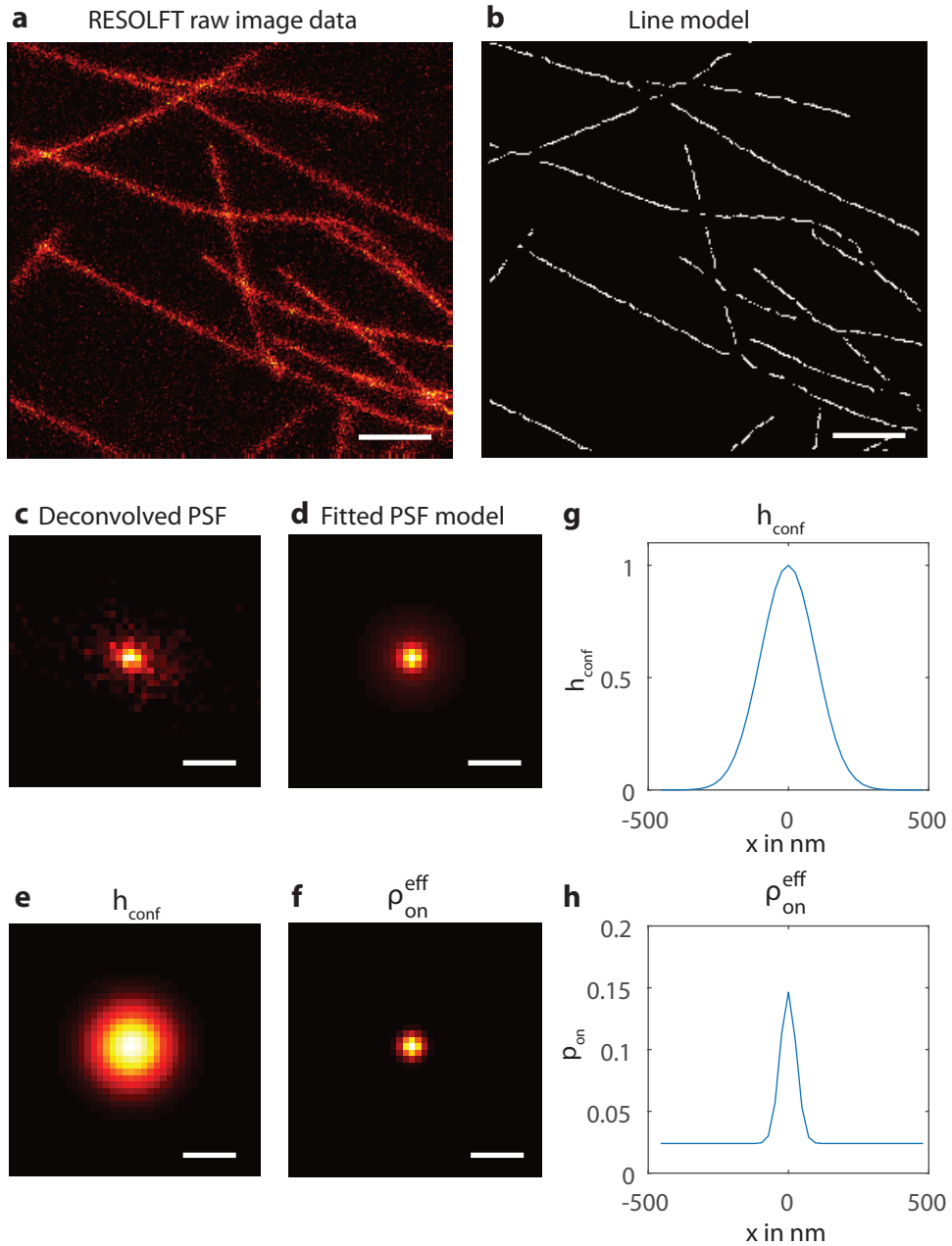
$$h_{\text{eff}} = a_1 h_{\text{conf}} + a_2 h_{\text{RESOLFT}} \quad (3.13)$$

$$= a_1 \exp\left(-4 \log(2) \frac{r^2}{w_{\text{conf}}^2}\right) + a_2 \exp\left(-4 \log(2) \frac{r^2}{w_{\text{RESOLFT}}^2}\right). \quad (3.14)$$

From this model, we find the values for the FWHM of the confocal background PSF to be

$$w_{\text{conf}} \cong 217 \text{ nm} . \quad (3.15)$$

### 3. Quantitative RESOLFT microscopy



**Figure 3.7. Estimation of the effective RESOLFT PSF** **a:** RESOLFT image of tubulin filaments in muscle tissue of a *Drosophila* embryo expressing rsEGFP2 fused to  $\alpha$ -tubulin. **b:** Localized tubulin filaments in the image. **c:** Result of the Richardson-Lucy deconvolution of the PSF. **d:** Fit of a two-dimensional Gaussian PSF with two components. **e:** Estimated confocal readout PSF  $h_{\text{conf}}$ . **f:** Estimated effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}$ . **g:** One-dimensional profile of the estimated confocal readout PSF  $h_{\text{conf}}$ . **h:** One-dimensional profile of the estimated effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}$ . Scale bars: **a,b:** 500 nm, **c-e:** 200 nm. Sample preparation and measurement was carried out by Sebastian Schnorrenberg.

### 3. Quantitative RESOLFT microscopy

The FWHM of the sharp RESOLFT PSF is measured to be

$$w_{\text{RESOLFT}} \cong 66 \text{ nm} . \quad (3.16)$$

As we have determined the equilibrium switching level  $\rho_\infty = 0.024$  for rsEGFP2 in Section 3.1, we can estimate the value of the on-switching probability  $p_{\text{on}}$  according to Equation (2.42) from the ratio of the measured amplitudes  $a_1$  and  $a_2$ , as

$$p_{\text{on}} = \rho_\infty \left(1 + \frac{a_2}{a_1}\right) . \quad (3.17)$$

The on-switching probability measured for this dataset is

$$p_{\text{on}} \cong 0.14 . \quad (3.18)$$

That means, at the focus, approximately 14% of the proteins remain in the on-state after the off-switching beam is applied.

Using these calibrated values, we can construct the two PSFs needed for the estimation of the brightness, the effective on-switching probability, given by Equation (2.40), and the confocal readout PSF, given by Equation (2.41).

The effective on-switching probability  $\rho_{\text{on}}^{\text{eff}}$  according to our model, is given by a constant  $\rho_\infty$  plus a Gaussian peak with a FWHM  $w_{\text{on}}$  which depends on the applied light dose, as shown in Figure 3.7f and h. The FWHM  $w_{\text{on}}$  of the effective on-switching probability can be calculated from the measured values of the confocal and the RESOLFT PSF.

$$w_{\text{on}} = \frac{w_{\text{RESOLFT}} w_{\text{conf}}}{\sqrt{w_{\text{conf}}^2 - w_{\text{RESOLFT}}^2}} \cong 63 \text{ nm} . \quad (3.19)$$

The effective on-switching probability, shown in Figure 3.7f, h is given by

$$\rho_{\text{on}}^{\text{eff}} = \rho_\infty + (p_{\text{on}} - \rho_\infty) \exp\left(-4 \log(2) \frac{r^2}{w_{\text{on}}^2}\right) . \quad (3.20)$$

The confocal readout PSF is given by a Gaussian PSF with the measured FWHM  $w_{\text{conf}}$  and is shown in Figure 3.7e, g.

#### 3.3.2. Quantification of the excess variance and counting molecules in RESOLFT microscopy

From the same RESOLFT image which was used in Section 3.3.1, we estimate the brightness  $b$  as described in Section 2.2.5. To this end, we first deconvolve the image using the Richardson-Lucy algorithm with the iteration procedure described in Equation (2.57) using the effective PSF and the value for the background which we

### 3. Quantitative RESOLFT microscopy

determined in Section 3.3.1. The initial estimate is chosen to be constant everywhere with the average value of the data set.

We stop the Richardson-Lucy algorithm after 50 iterations. The result of the deconvolution, the estimated object  $\hat{\rho}$  is an estimate of the product of the number density  $n(\vec{r})$  of the proteins in the sample times the brightness  $b$  per protein. We estimate the mean signal in the image  $\hat{\mu}$  by convolving  $\hat{\rho}$  with the effective PSF  $h_{\text{eff}}$  and adding again the calibrated constant background value  $d$  given in Equation (3.12):

$$\hat{\mu} = (\hat{\rho} \otimes h_{\text{eff}}) + d. \quad (3.21)$$

The estimated spatial distribution of the excess variance  $\hat{v}$  is calculated by convolving  $\hat{\rho}$  with the PSF of the excess variance given by Equation (2.31).

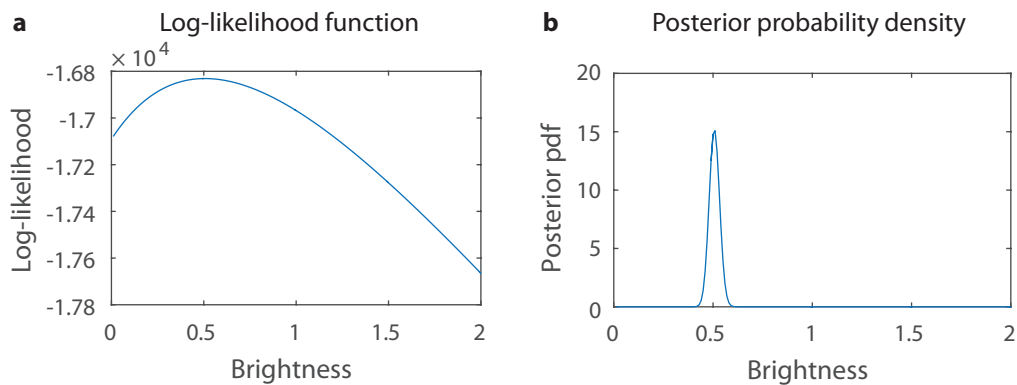
$$\hat{v} = \hat{\rho} \otimes h_{\text{exc}} \quad (3.22)$$

These results are used in the approximated log-likelihood function described in Equation (2.64). The values of the log-likelihood function are calculated numerically and shown in Figure 3.8. Assuming a uniform prior distribution, we can normalize this function to get the posterior probability density for the parameter  $b$ . The posterior mean estimate of the parameter  $\hat{b}$  from this procedure is given by

$$\hat{b} \cong 0.507 \quad (3.23)$$

with a 90% credible interval between 0.46 and 0.54. By dividing the result of the deconvolution  $\hat{\rho}$  by the estimated brightness  $\hat{b}$ , we get an estimate of the number density of fluorophores in the image, shown in Figure 3.9. Integrating the density over single microtubules identified in the image, we can estimate the linear density of rsEGFP2 along filaments in the image. For five microtubule segments in the image, the average linear density of rsEGFP2 was quantified, and the results are shown in Table 3.1.

### 3. Quantitative RESOLFT microscopy

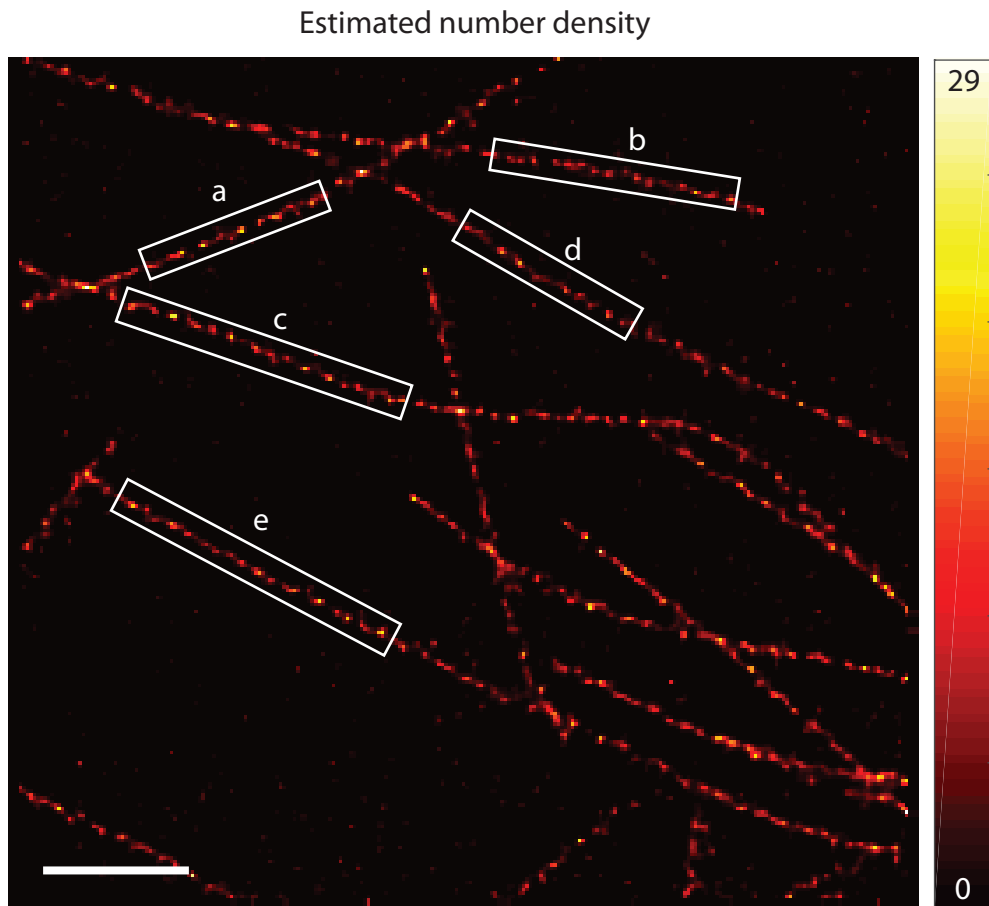


**Figure 3.8.** **Log likelihood and posterior distribution for  $b$ .** **a:** Log likelihood function for the brightness parameter  $b$ , given the image data. **b:** Numerically normalized posterior probability density for the brightness parameter  $b$ , given the image data.

**Table 3.1.** Estimated linear number density of rsEGFP2 for the regions in Figure 3.9.

Region	Number	Length in $\mu\text{m}$	Estimated density per $\mu\text{m}$
a	655	1.29	508
b	715	1.69	423
c	952	2.03	469
d	521	1.35	386
e	914	2.11	433
total	3757	8.47	444

### 3. Quantitative RESOLFT microscopy



**Figure 3.9.** Estimation of the number density of rsEGFP2. The number density was estimated for the RESOLFT image shown in Figure 3.7a. For each of the regions **a-e**, the linear density along the filament was quantified by integration the density over the area and dividing the result by the length of the filament. The results are shown in Table 3.1. Scale bar, 500 nm.



## 4. Activation-based multicolor STORM

The theoretical model of activation-based multicolor STORM presented in Section 2.3.3 is based on the two state Markov chain model. To validate this model, and to find the optimal imaging conditions given this model for the fluorophore, we measured the intensity-dependent kinetic on- and off-switching rates for the organic dye Alexa 647. To this end, single molecule datasets were taken with different illumination intensities of the excitation laser without additional activation of fluorophores, to measure the on-switching rate.

To measure the off-switching rate, we developed a method for single molecule event timing in Section 2.3.2. To evaluate the performance of the transition time estimation, we conducted experiments with fluorescent beads and single molecules. The method was then applied to measure the intensity-dependent off-switching rate of Alexa 647.

The model described in Section 2.3.3 predicts that the crosstalk between the color channels in activation-based multicolor STORM can be reduced by measuring the on-switching time precisely, to identify specifically activated single molecule events. We describe multicolor STORM experiments with precisely timed activation pulses and quantify the crosstalk due to the random activation of single molecules and due to cross-activation of the species.

### 4.1. Measuring Alexa 647 on-switching kinetics

To prepare a single molecule sample, single Alexa 647-labeled double stranded DNA molecules were bound to a biotin-streptavidin coated glass surface at 500 pM concentration. The sample was immersed in a buffer containing an enzymatic oxygen scavenging system (pyranose oxidase and catalase) and primary thiol (mercaptoethylamine). The samples were imaged using an inverted fluorescence microscope as described previously [42]. Single molecule experiments were performed at six intensities of the 642 nm excitation laser ranging from 0.4 kW/cm<sup>2</sup> to 6.5 kW/cm<sup>2</sup>. For each intensity value, a measurement of 300 seconds was taken with different frame-rates adjusted to the expected event duration time. The measurement parameters

#### 4. Activation-based multicolor STORM

are shown in Table 4.1. From these datasets, events were localized using a STORM image analysis algorithm [11, 65]. In the localization data, clusters of localizations were identified as single molecules. Fitting a linear model to the dataset, as described in Section 2.3.1, the brightness of each single molecule in each frame was estimated to obtain the signal trace.

For each single molecule signal trace, the first off-state dwell time was measured as the first time the signal was measured above a manually chosen threshold value for each dataset. The on-rate was estimated by modeling the distribution of measured off-state dwell times according to equation (2.73) with the upper limit of  $T = 300s$ . In Figure 4.1, we show histograms of the off-times with the corresponding model distribution given by equation (2.10) with the estimated parameters. The on-switching rate appears to be linearly dependent on the intensity  $I_{642}$  with an offset, the constant thermal on-switching rate:

$$k_{\text{on}}(I_{642}) = k_{\text{th}} + \sigma_{\text{on}} \frac{\lambda}{hc} I_{642}. \quad (4.1)$$

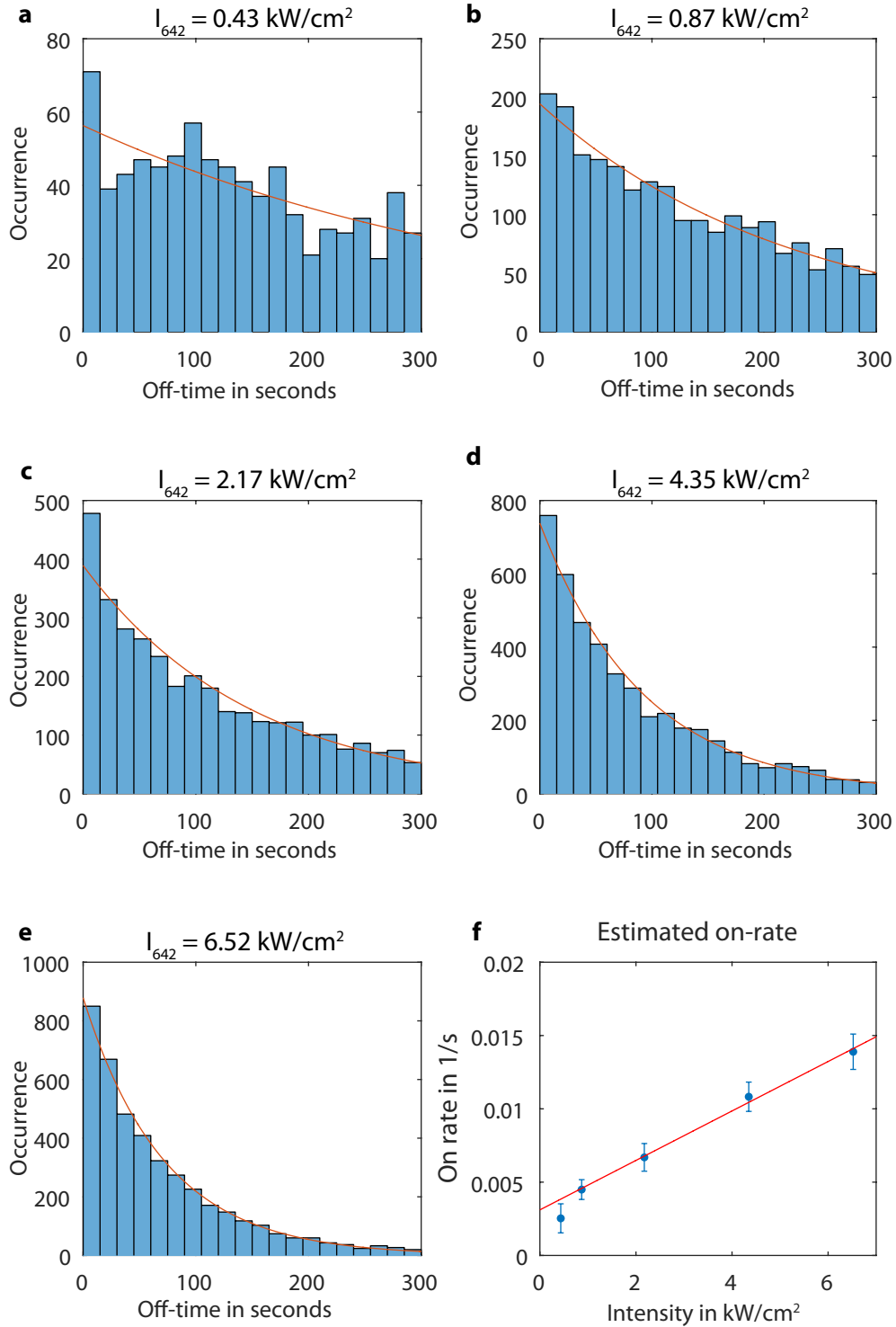
By fitting this model to the measured on-rates we can estimate the effective on-switching cross-section for Alexa 647:

$$\sigma_{\text{on}} = 5.3 \times 10^{-22} \text{cm}^2. \quad (4.2)$$

**Table 4.1.** Measurement parameters for the single molecule switching kinetics experiments.

Intensity in kW/cm <sup>2</sup>	Frame rate in Hz
0.43	20
0.87	20
0.87	50
2.17	50
2.17	100
4.35	100
4.35	200
6.52	100
6.52	200

#### 4. Activation-based multicolor STORM



**Figure 4.1. Estimation of the on-switching rate for Alexa 647. a-e:** Measured off-times for different intensity values. The plotted curve is the distribution expected for the on-switching rate parameter estimated from the data in each histogram. **f:** The intensity-dependence of the on-switching rate is shown. A linear fit of the dependence is shown in red. Sample preparation and measurement was carried out by Michael Weber and Mark Bates.

## 4.2. Single molecule event timing measurements

In Section 2.3.2, we introduced a method to estimate the transition times of single molecule events with a sub-framerate timing precision by calculating the ratio of photon counts in the first frame in which a single molecule event is detected and the average rate of photon counts measured during a frame in which the molecule was in the on-state for the whole time. We calculated the Cramer-Rao lower bound for the variance of the estimator in Equation (2.78), using a Poisson noise model with constant background counts.

### 4.2.1. Switching time estimation for fluorescent beads

To find how well the switching time can be estimated, we first conducted an experiment with fluorescent beads. The excitation laser was controlled such that the beads were exposed to light with constant intensity for a given time interval, starting at a fraction of 0.8 of a camera frame. This pulse scheme was repeated 150 times. The switching time was estimated by first identifying the transition frames using the hidden Markov model described in Section 2.3.2. Using the estimator given in Equation (2.76), the on-switching transition time was estimated for beads under different intensity conditions. We calculated the standard deviation of the switching time estimation and the number of photon counts for each bead under different illumination intensities. The result is shown in Figure 4.2 together with the square root of the Cramer-Rao lower bound calculated for the estimator without background noise:

$$\sigma_{\text{CRLB}} = \sqrt{\frac{1 - t_1}{b}}. \quad (4.3)$$

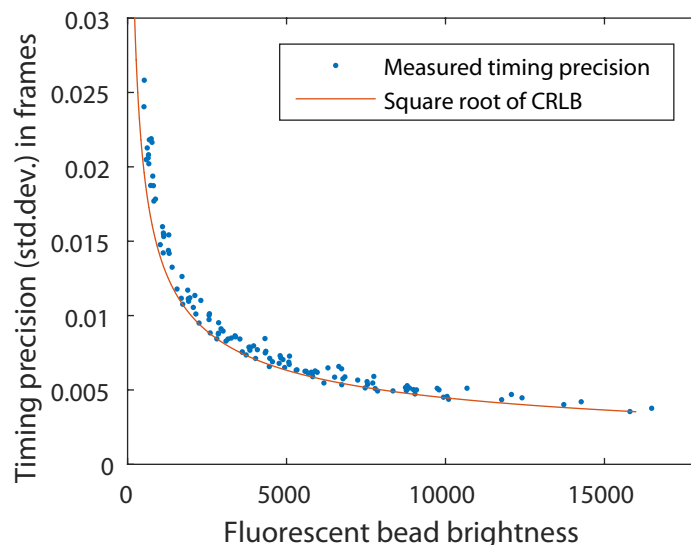
For fluorescent beads, the measured timing precision is close to the Cramer-Rao bound.

### 4.2.2. Switching time estimation for single molecules

Single molecule experiments were conducted using streptavidin bound to a glass surface coated with biotinylated bovine serum albumin. The streptavidin was labeled with Cy3 and Alexa 647, such that the Alexa 647 dyes can be activated via excitation of the Cy3 dye using a green 532nm laser.

The activation pulses were timed at specific delays during the exposure time of 50ms of the camera. For each delay, 3000 frames were recorded. The single molecule switching events were localized from the dataset. For each event, the signal trace was extracted from the data and the switching time was estimated.

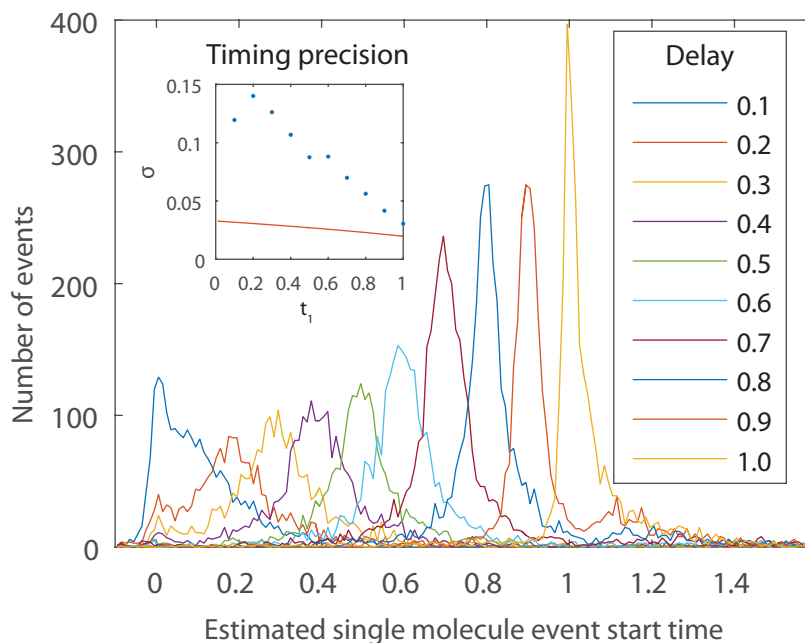
#### 4. Activation-based multicolor STORM



**Figure 4.2. Timing precision for fluorescent beads.** For each data point, the brightness and switching time was measured for 150 times. The red curve is the square root of the Cramer-Rao lower bound for the number of photon counts and the starting time during the frame 0.8. Sample preparation and measurement was carried out by Mark Bates.

In Figure 4.3, the results of the switching time estimation relative to the beginning of the frame with an activation pulse are shown. To estimate the timing precision for the different delays, we fit Gaussian distributions to the histogram. The Cramer-Rao bound is shown for the average brightness and the background measured from the single molecule events. The measured timing precision is larger than the estimated Cramer-Rao bound for Poisson noise by up to a factor of five. This discrepancy may be explained by a higher variance in the signal due to sub-framerate blinking of the single molecules.

#### 4. Activation-based multicolor STORM



**Figure 4.3. Single molecule measurements of the timing precision for Alexa 647.** The measured single molecule event transition times are shown as histograms relative to the beginning of the frame in which the activation pulse was given with different delays. Inset: The standard deviation for each delay (blue dots) is shown together with the square root of the Cramer-Rao lower bound (red line). Sample preparation and measurement was carried out by Mark Bates.

### 4.3. Alexa 647 off-switching kinetics

The single molecule event timing method can be applied to the single molecule kinetics data described in Section 4.1 to measure the duration of the events, that is the on-state dwell time, depending on the intensity. The on- and off-switching transition times are measured for each single molecule event detected in the dataset. Here, the events are not activated specifically, such that the on-switching time is not synchronized with the camera. To be able to measure the transition times precisely, the first and last frames of the single molecule event have to be identified using the hidden Markov model algorithm developed in Section 2.3.2. This is only possible if the signal can be detected in the transition frames and the molecule is in the on-state for at least one frame in between. Thus, the on-state dwell times can only be measured for events which are longer than two camera frames. To estimate the off-switching rate, we have to take this into account when using the estimator derived in Section 2.1.3 given by equation (2.17). The results of the estimation are shown in Figure 4.4. The off-switching rate is proportional to the applied intensity of the excitation laser. The number of photons per event, given by the estimated rate  $b$  times the duration of the event, does not depend on the intensity. By fitting a linear dependence to the measured off-rate, we can estimate the effective off-switching cross-section for Alexa 647:

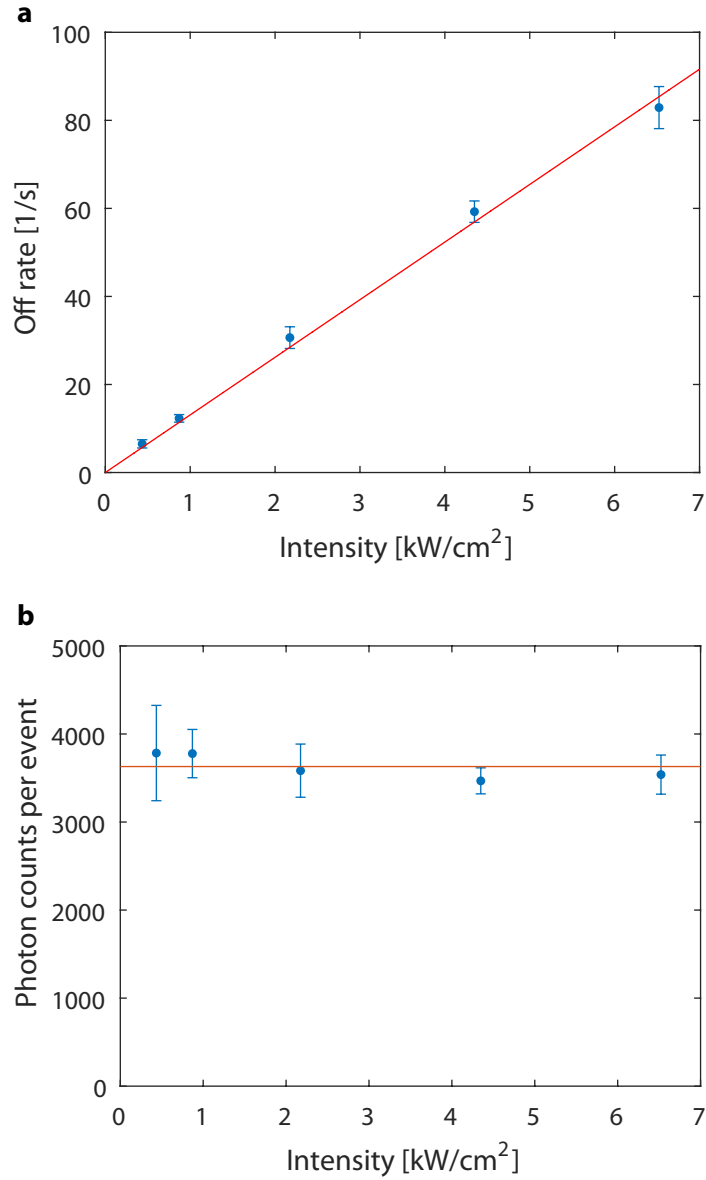
$$\sigma_{\text{off}} = 4.1 \times 10^{-18} \text{cm}^2. \quad (4.4)$$

### 4.4. Multicolor STORM using single molecule event timing

An application for the single molecule event timing method is the reduction of the crosstalk in activation-based multicolor STORM. In Section 4.2.2, we showed that the activation time can be measured with a sub-frame time precision for single molecules. As discussed in Section 2.3.3, the random activation of single molecules is a source of crosstalk, which can be reduced by using the switching time estimation to discriminate between randomly and specifically activated single molecule events.

We performed multicolor STORM experiments using activator-reporter dye pairs. For the first color channel, the mitochondrial protein Tom20 was labeled with activator-reporter dye pairs of Alexa 647 and Cy3. For the second color channel, Tubulin was labeled with dye pairs of Alexa 647 and Alexa 405. In both cases, the fluorophores were conjugated to antibodies which were bound to the protein of interest, as described previously [42,50]. The two species of dye pairs were activated

#### 4. Activation-based multicolor STORM



**Figure 4.4. Off-switching rate measurement for Alexa647. a:** Intensity-dependent off-switching rate. A linear dependence (red line) is fitted to the measured data (blue dots). **b:** Total number of photon counts per single molecule event for single molecules of Alexa 647 (blue dots). The mean of the number of photons is shown in red.



#### 4. Activation-based multicolor STORM

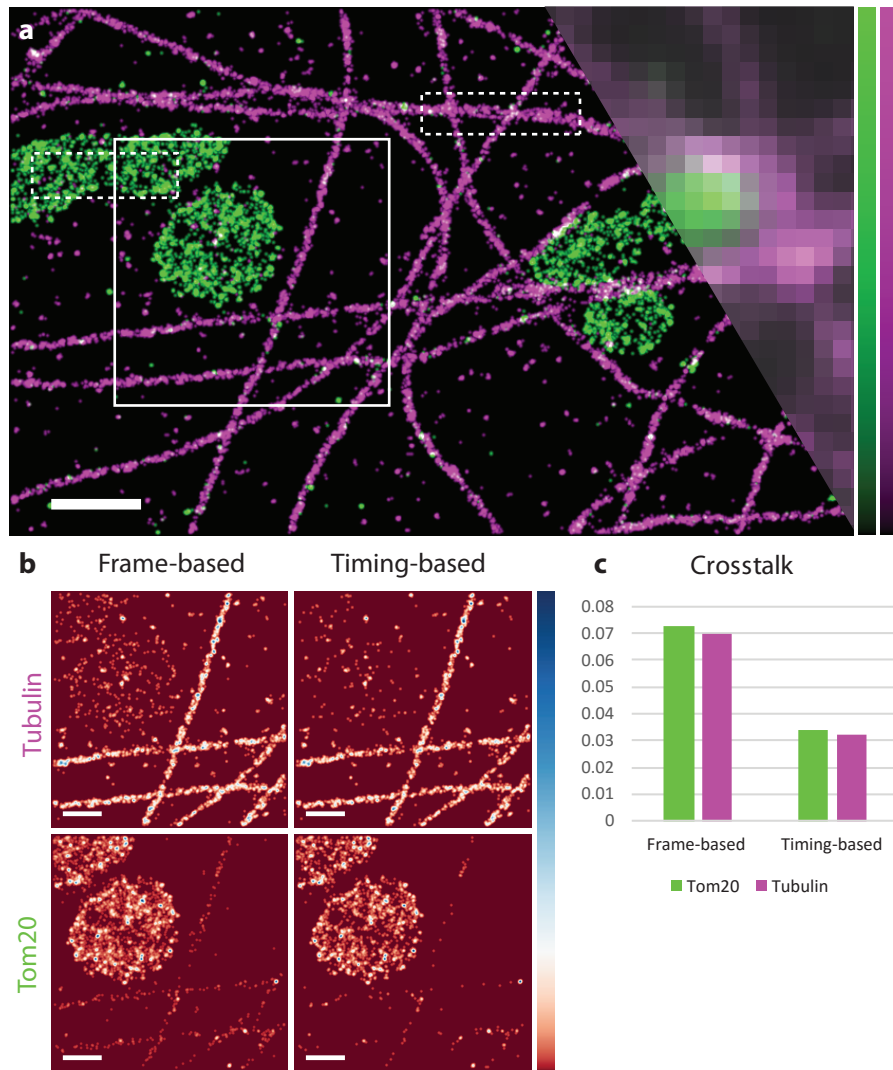
using a green or UV laser pulse, respectively. The camera exposure time was 6.8 ms. The activation pulses were given after 0.7 of the camera exposure had elapsed. The duration of the pulses was controlled using a feedback loop to keep the number of specifically activated events at a constant level. During the experiment, the pulses were applied alternately every 10 camera frames.

The data was analyzed by localizing the position of each single molecule event using a STORM image analysis algorithm [11, 65]. Then, for each event, the signal trace was extracted from the data and the starting time of the event was estimated as described in Section 2.3.2.

We compared two methods of assigning the single molecule events to the color channels. In the frame-based method, all events which were first detected in one of the frames with a pulse are assigned to the corresponding color channel. This is equivalent to the procedure described before [42, 50]. In the timing-based method, the sub-frame estimation of the on-switching time was used to assign events to each color channel which have a high probability of being correctly identified. The resulting images are compared in Figure 4.5b. In Figure 4.6, a histogram of the on-switching transition times for one of the pulses is shown. The size of the time interval used for the assignment of the color determines the number of randomly activated events assigned to the color channel. There is a trade-off between the number of incorrectly assigned random events and the number discarded events. As seen in Figure 4.6, when retaining approximately two thirds of the events, the crosstalk due to randomly activated molecules can be reduced by approximately one half by choosing a symmetric interval of length 0.2 camera exposure times, that is  $\pm 0.1$  around the time of the activation pulse.

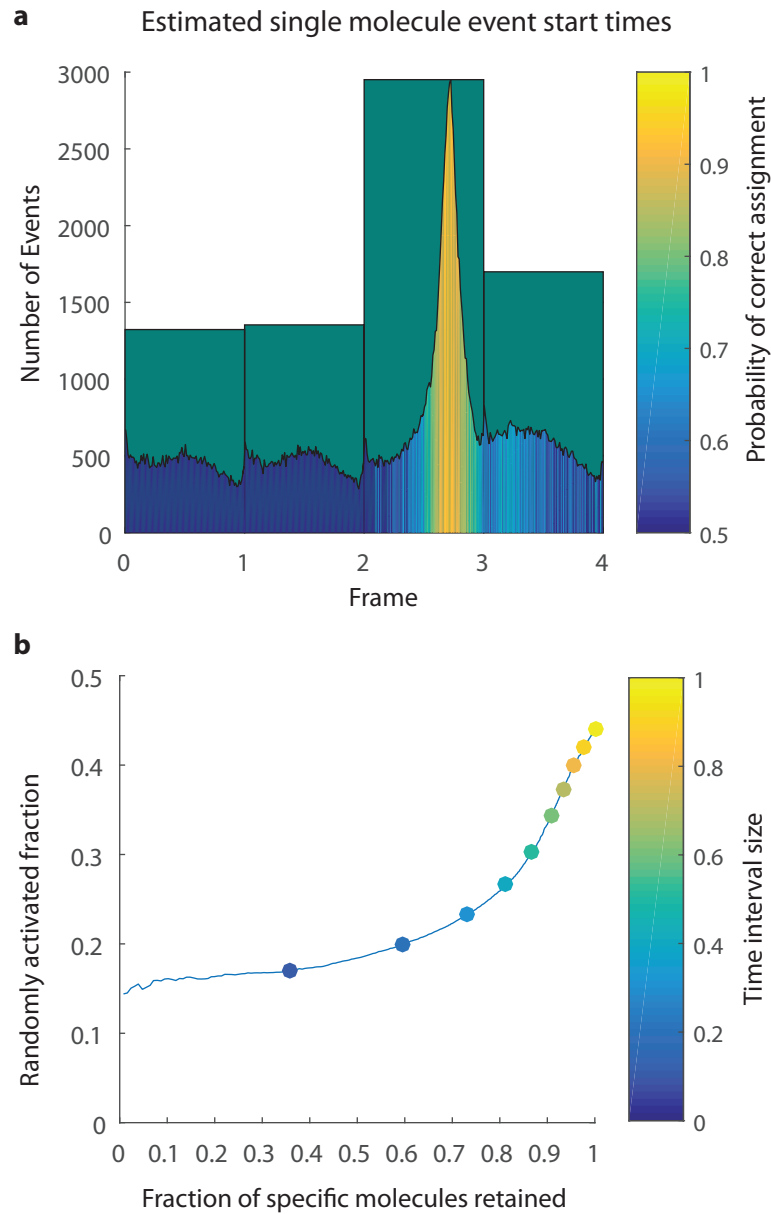
To analyze the crosstalk in the image shown in Figure 4.5, we analyzed regions in which only one of the structures were present (microtubules or Tom20). In each of these regions, the number of localizations which were assigned to each channel was counted. The crosstalk was calculated as the fraction of localizations which were assigned to the wrong color channel. For both channels, the crosstalk was about 7% using the frame-based method to assign the events to the color channels. Using the timing-based method, and a time interval of 0.2 camera exposure times to assign the color channels, the crosstalk was reduced to below 3.5%.

#### 4. Activation-based multicolor STORM



**Figure 4.5. Multicolor STORM image of Tom20 and Tubulin.** **a:** Multicolor STORM image of Tom20 and Tubulin. Scale bar, 1  $\mu\text{m}$  **b:** Inset region in **a** shown with a high contrast colormap. The frame- and timing-based methods are compared for the individual color channels. Scale bar, 500 nm. **c** Quantified crosstalk values in the dashed regions in **a**. Sample preparation and measurement was carried out by Michael Weber and Mark Bates.

#### 4. Activation-based multicolor STORM



**Figure 4.6. On-switching transition time and color assignment. a:** Green: Histogram of the number of events detected in each frame, according to the frame-based method. Color coded: Histogram of estimated on-switching transition times color coded with the estimated probability of correct assignment. **b:** Trade-off between crosstalk and fraction of specifically activated events retained in the image. The size of the chosen time interval for the assignment of the color channel is shown in color code. The choice of the timing window determines both the fraction of specifically activated molecules retained in the image and the fraction of randomly activated events which cause the crosstalk.

#### 4.4.1. Quantification of the sources of crosstalk

In Section 2.3.3, we described two sources of crosstalk. The cross-activation due to the spectral overlap of the two activator dyes and the crosstalk due to random activation of the reporter fluorophores. By estimating the on-switching time for each single molecule event, we can reduce the crosstalk due to the randomly activated reporter dyes. To find the relative contributions of the crosstalk due to randomly activated reporter dyes and cross-activation of the dyes, control experiments were conducted.

First, two-color STORM images of the two nuclear pore complex proteins, gp210 and panFG in A6 cells were taken shown in Figure 4.7. The structures were labeled with primary antibodies against gp210 and panFG, and secondary antibodies labeled with either the Alexa 647-Alexa 405 dye pair, or the Alexa 647-Cy3 dye pair.

In the control experiments, the samples were only labeled with one of the secondary antibodies. The same intensity profile to which the feedback loop was adjusted in the two-color experiment was applied for the activation pulses of both colors.

The contributions of the two sources of crosstalk depending on the chosen time interval  $\Delta t_{\text{timing}}$  can be quantified as follows. We denote the number of activations in a time interval with a specific activation pulse for the species  $i$  as  $N_i$ . This is a mixture of the specifically activated events  $S_i$  of the species associated with the channel and the cross-activated events from the other channels:

$$N_i = S_i + r_i S_i + \sum_{j \neq i} (x_j + r_j) S_j. \quad (4.5)$$

Here, the fraction of randomly activated events for each species is given by  $r_i$  and the fraction of cross-activated events are given by  $x_i$ . In a time interval of the same proportions during a frame without a specific activation pulse, only randomly activated events occur:

$$N_0 = \sum_j r_j S_j. \quad (4.6)$$

The measured crosstalk in the final image is given by the contributions from the other color channels:

$$\begin{aligned} X_i &= \frac{\sum_{j \neq i} (x_j + r_j) S_j}{N_i} & (4.7) \\ &= X_i^{\text{cross-activation}} + X_i^{\text{random}}. & (4.8) \end{aligned}$$

In the control images, shown in Figures 4.8 and 4.9, the fractions  $x_i$  and  $r_i$  can be measured directly. Using these ratios, we can estimate the relative contribution

#### 4. Activation-based multicolor STORM

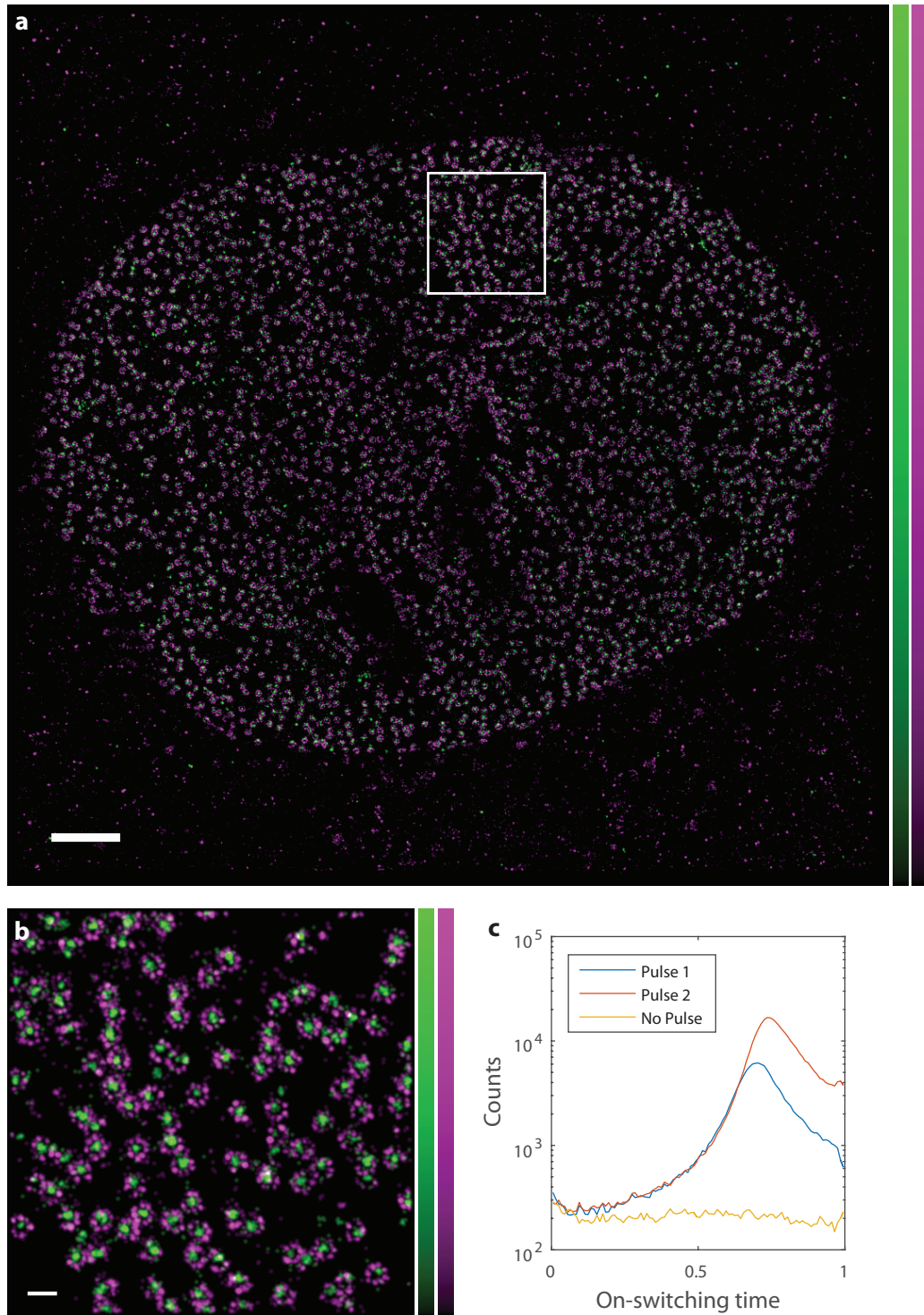
of the cross-activation and the random activation on the crosstalk in the two color image for the frame-based and the timing-based analysis.

To this end, we measure the numbers  $N_i$  from the estimated on-switching times of the two color dataset shown in Figure 4.7. By solving the system of linear equations (4.5), the number of specific activations in the color channels is estimated. The crosstalk for both color channels can be calculated according to Equation (4.7), while also finding the contributions for the two sources of crosstalk. The results of the crosstalk estimation are shown in Table 4.2. The total crosstalk is again reduced by approximately one half by using the timing based method by choosing a symmetric interval of 0.2 camera exposure times around the pulse time. The improvement is mainly due to the reduction of randomly activated events which are assigned to each color channel.

**Table 4.2.** Estimated crosstalk for random activation and cross-activation.

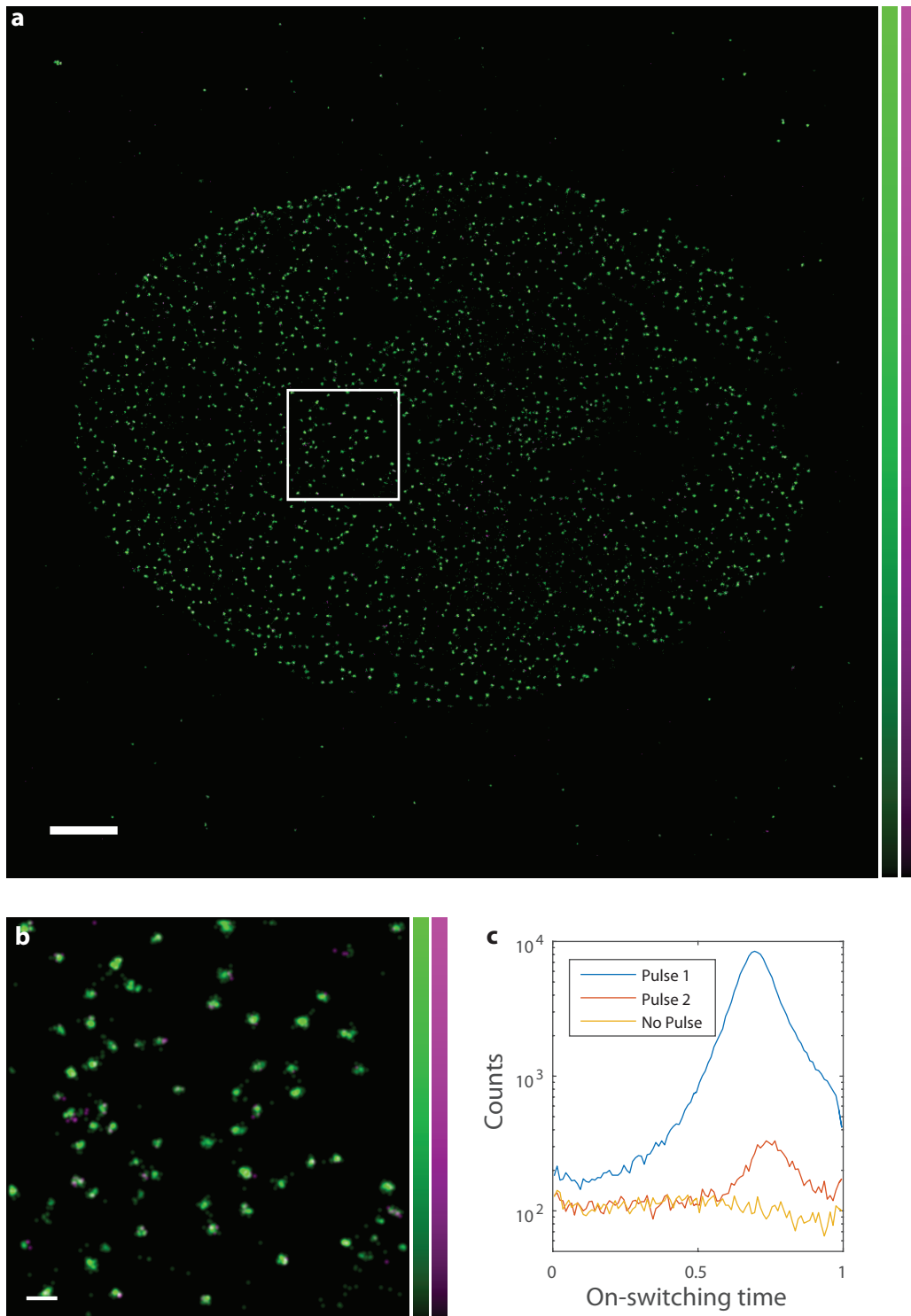
Species	Method	Total crosstalk	Cross-activation	Random activation
gp210	Frame-based	3.32%	0.99%	2.33%
	Timing-based	1.62%	0.65%	0.97%
panFG	Frame-based	10.34%	0.85%	9.49%
	Timing-based	5.08%	1.8%	3.28%

#### 4. Activation-based multicolor STORM



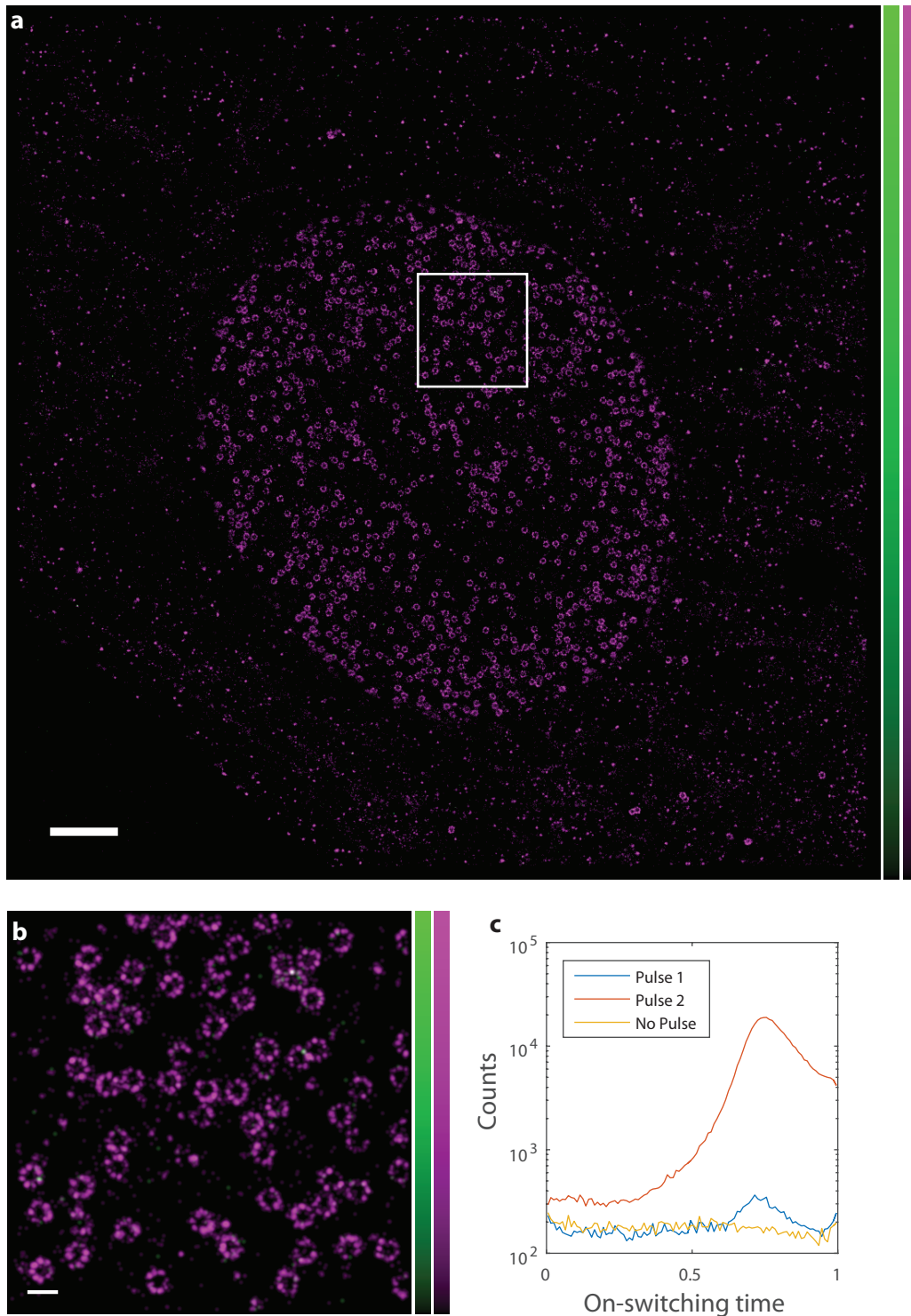
**Figure 4.7.** Multicolor STORM image of gp210 and panFG. **a:** Whole field of view. Scale bar,  $2 \mu\text{m}$ . **b:** Zoomed region. Scale bar,  $250 \text{ nm}$ . **c:** Histogram of measured on-switching times shown on a logarithmic scale for the frames with specific activation pulses and a frame without a specific activation pulse. Sample preparation and measurement was carried out by Michael Weber and Mark Bates.

#### 4. Activation-based multicolor STORM



**Figure 4.8.** Multicolor STORM image of panFG control sample. **a:** Whole field of view. Scale bar,  $2 \mu\text{m}$ . **b:** Zoomed region. Scale bar,  $250 \text{ nm}$ . **c:** Histogram of measured on-switching times shown on a logarithmic scale for the frames with specific activation pulses and a frame without a specific activation pulse. Sample preparation and measurement was carried out by Michael Weber and Mark Bates.

#### 4. Activation-based multicolor STORM



**Figure 4.9.** Multicolor STORM image of gp210 control sample. **a:** Whole field of view. Scale bar,  $2 \mu\text{m}$ . **b:** Zoomed region. Scale bar,  $250 \text{ nm}$ . **c:** Histogram of measured on-switching times shown on a logarithmic scale for the frames with specific activation pulses and a frame without a specific activation pulse. Sample preparation and measurement was carried out by Michael Weber and Mark Bates.



## 5. Discussion

In this work, we quantitatively modeled superresolution microscopy based on photoswitchable emitters using a simple two state switching model. For RESOLFT microscopy based on reversible switchable fluorescent proteins, we developed a method to calibrate the brightness per fluorophore from the image data, based on the modeling and estimation of the excess variance due to the stochastic switching of the fluorophores. This was applied to estimate the number density of fluorophores in the image. For STORM microscopy, we introduced a method to measure the on-switching transition time of the single molecule events with sub-frame timing precision. This was applied to reduce the crosstalk due to random activation in activation-based multicolor STORM microscopy.

### 5.1. Two state model

Both imaging modalities considered in this thesis were modeled based on the two state Markov chain. As described in Section 2.1.1, the switching behavior for a single realization of the Markov chain is characterized by its exponentially distributed state dwell times. The single molecule data of Alexa 647 we describe in Section 4.1 show, that both the on- and the off-state dwell times are distributed exponentially under the controlled conditions in the single molecule experiment. In another work, up to four states have been considered to model the switching behavior of Alexa 647 [76]. However, on the timescales of the camera exposure time (5-50 ms) used and the duration of our experiments (5 minutes), we find that two states are sufficient to describe the switching behavior relevant to SMLM.

The switching curves measured for rsEGFP and rsEGFP2 shown in Section 3.1 could be approximated closely with a superposition of exponential decay functions, weighted by a gamma distribution. Thus, the decay law describing the off-switching signal in these proteins is not described fully using the simple two state Markov model. This is not surprising, as different processes can be involved in the photo-switching process for reversibly fluorescent proteins, like isomerization and protonation [59].

## 5. Discussion

### 5.1.1. States on shorter timescales

On timescales shorter than those considered in this work, fluorophores can undergo transitions between the excited and the ground state, which happens on the order of nanoseconds, and the triplet state, which can have lifetimes of a few microseconds to a few milliseconds [69].

In the RESOLFT measurements, the dwell times were chosen to be  $30 \mu\text{s}$ , which is longer than the reported lifetimes of the triplet state dynamics rsEGFP2, which were reported on the order of  $1 \mu\text{s}$  [18]. As the signal does not show strong saturation effects, as seen in Figure 3.2, metastable triplet state dynamics are probably not limiting the signal from the proteins.

The switching transitions on shorter timescales do play a role for the measurement of the transition time in Alexa 647, as they could increase the variance of the measured fluorescence signal. An increased variance may lead to a lower timing precision than expected for Poisson noise. These processes may explain why the timing precision measured for the single molecule data is worse than expected for Poisson noise, as seen in Section 4.2.2.

### 5.1.2. States on longer timescales

On longer timescales, most fluorophores lose the ability to fluoresce due to photobleaching. The fluorescent protein rsEGFP2 was shown to last more than 2000 switching cycles on average before the signal drops to half of the initial value [17]. For the *Drosophila* embryo muscle cells used for the quantification of the brightness in Section 3.3, it was possible to take movies of 20 frames before the fluorescence signal dropped to two thirds of the initial value [20]. Assuming a linear photobleaching process depending on the applied light dose, this would amount to a loss of fluorescence signal of about 2% per recorded image. Photobleaching is thus not expected to strongly influence the result of the quantification.

In the single molecule experiments for Alexa 647, we were not able to observe the photobleaching process, as the duration of the experiments was too short. The average number of switching cycles Alexa 647 can undergo before photobleaching was previously reported to be approximately 11 [27]. The average off-state dwell times we measured in Section 4.1 are approximately one quarter of the total duration of the experiment even for the highest intensity. To measure the total number of photoswitching cycles under these conditions, longer experiments would be needed.

However, in the multicolor measurements described in Section 4.4, photobleaching can occur on the timescale of the experiment. Photobleaching decreases the density

## 5. Discussion

of fluorophores which are able to fluoresce over time. For the crosstalk in activation-based multicolor STORM, a lower density of fluorophores is desirable, as more single molecule events can be activated specifically. However, due to a decreasing labeling density, the structure may not be represented well by the remaining labels after a large fraction of the fluorophores is bleached [72].

### 5.2. Quantitative RESOLFT microscopy

The stochastic image formation model for RESOLFT microscopy described in Section 2.2.1 was used to predict the mean and the variance of the measured signal. For the switching kinetics with a linear intensity dependence of the on- and off-switching rates over a large range of intensities, we calculated the specific shape of the effective RESOLFT PSF. We validated this switching model in Section 3.1. In the general case, to calculate the shape of the PSF, the intensity-dependent switching kinetics must be known.

The effective RESOLFT PSF depends not only on the on- and off-switching rates, but also on the focal intensity distribution which is used to switch the fluorophores. We showed in Section 2.2.2, that the approximations used fit well with numerically calculated intensity profiles. Due to aberrations, the intensity distributions in the experiment may have slightly different shapes, such that the theoretical model which consists of two Gaussian PSFs may not be appropriate [77].

However, by estimating the PSF directly from the image data, as described in Section 2.2.4, it should be possible to detect strong deviations from the expected shapes. The described procedure does not assume any model for the shape of the PSF. In Section 3.3.1, the shape of the estimated PSF was seen to be well described by the theoretical model.

The filtering method used to estimate the object heuristically may fail in some cases, for example if the PSF has more than one peak. For dense regions in the image, like at the crossing of two lines, the algorithm fails to identify the lines correctly. Due to the deconvolution procedure, the shape of the PSF is optimized over the whole image, such that small regions of higher density should not influence the result strongly.

### 5.2.1. Blind deconvolution and regularization

The procedure sketched in Section 2.2.4 for estimating the PSF is a blind deconvolution method [78–80]. Blind deconvolution algorithms are used if both the PSF and the image are unknown. In this case, an unconstrained maximum likelihood approach would fail, because the trivial solution, in which the underlying structure is equal to the image and the PSF is a delta peak, maximizes the likelihood. This is also called overfitting [66]. Overfitting can happen if there are more flexible parameters than independent data points. As the number of parameters in blind deconvolution is much larger than the number of image pixels, the method is prone to overfitting.

Regularization methods can be used to reduce the effect of overfitting in these cases. Some regularization methods add terms to the likelihood function, which are designed to reduce the likelihood for complex solutions [66]. In the method for estimating the effective PSF, two implicit regularization effects are at work. First, the object is fixed to the result of the heuristic maximum line extraction. Therefore the parameters of the object are not optimized. The second regularization effect is the number of iterations in the Richardson-Lucy deconvolution. By stopping the algorithm early without converging to the maximum likelihood solution, the result is regularized implicitly [81]. In practice, this can reduce typical artifacts which occur, like the separation of structures into alternating pixels with high and low signals, respectively.

The Richardson-Lucy deconvolution is also used for the estimation of the brightness, as described in Section 2.2.5. Here, given the effective PSF, the object is estimated. In a second step, the excess variance in the image is used to quantify the brightness. This procedure depends strongly on the number of iterations used. Each iteration of the Richardson-Lucy deconvolution increases the likelihood of the model given the data. This also decreases the square deviation of the data to the model, thus decreasing the excess variance which can be measured. During the first iterations, the estimate of the object typically improves, at some point however, artifacts are introduced, which increase the value of the likelihood function, but may not represent the object accurately. It is not clear how to determine exactly the transition between these two regimes [81]. The number of iterations thus has to be chosen with care.

### 5.2.2. Counting error

The final result of the estimated number density depends critically on the correct calibration of the on-switching probability. The total number of molecules in the RESOLFT image is basically given by the total number of photon counts in the image divided by the brightness and by the on-switching probability. The on-switching probability is calibrated indirectly, using the shape of the PSF and the equilibrium level for the switching process, which was calibrated in a separate experiment. Any errors in the calibration of the brightness and the on-switching probability will lead to an error in the estimation of the number of molecules in the image. In simulated data with a known on-switching probability, the estimated brightness showed a small bias and was typically estimated too low. This would amount to systematically over-estimating the number of molecules in the image. Assuming that the calibration of the brightness and the on-switching probability can be performed accurately, the error in determining the number is limited by the variance of the signal.

In all counting methods, the single molecule behavior is modeled. The signal of, for example a cluster of molecules is then assumed to be the superposition of the independent signals of many single molecules. When counting molecules in superresolution microscopy, the number of fluorophores in sub-diffraction volumes is estimated. When fluorophores are closer together than about 10 nm, they may interact by fluorescence resonance energy transfer (FRET) between an excited donor and an acceptor [69, 82]. This process is used to study interactions in biological systems, by using fluorophores of different species with overlapping spectra. If the absorption and emission spectra of a single fluorophore overlap, FRET between fluorophores of the same species can occur [53, 83]. These interactions may invalidate the assumption of the independent switching and emission in the fluorophore models used for counting molecules. For the microtubule sample studied in Section 3.3, the linear density of fluorophores along a microtubule was estimated to be around 400 per  $\mu\text{m}$ . For a random distribution of 400 fluorescent proteins in a cylindrical volume given by the diameter of 25 nm of the microtubule and its length of 1  $\mu\text{m}$ , the average distance between proteins is approximately 15 nm. Thus, interactions between some of the fluorophores in the sample cannot be ruled out. However, as the fluorophores are prepared in a sub-diffraction volume in the RESOLFT microscope, not all of the fluorophores in the detection volume are in the fluorescent state at the time of detection. As estimated in Section 3.3, approximately 14% of the fluorophores at the focus remain in the on-state before the detection step, such that the effective intermolecular distance between fluorophores in the on-state may be larger.

### 5.3. Single molecule event timing and multicolor STORM

By estimating the switching transition time of single molecule events with a sub-frame time precision, we were able to reduce the crosstalk in activation-based multicolor STORM microscopy by about a factor of two, as shown in Section 4.4.

#### 5.3.1. Single molecule event timing

A method for estimating transition times with a higher timing precision for camera-based data has been used in FRET studies [84]. The FRET signal was modeled as a step function, similar to the model for the single molecule events in Section 2.3.2. However, instead of estimating the transition time for each event individually, the signals from many individual events were aligned by correlating upsampled signals for each event. Another method for increasing the time resolution of a camera detector is stroboscopic illumination, where the illumination light is only applied for a short duration during the camera exposure [85].

The problem of estimating a change point in the time-resolved signal of a Poisson process is important for the analysis of single molecule signals [86, 87]. However, typically it is assumed that the time resolution is high, for example for data taken with an avalanche photodiode, in which the time of each single photon arrival is measured. The determination of the change point is then limited by the noise of the signal. In this case, the time of the change point can be estimated using the binary segmentation technique [88]. The dataset is split in two parts. The null hypothesis of no change in the Poisson rate is compared to the hypothesis of a change point at the position at which the dataset was split using a likelihood ratio test.

Due to the limited time resolution of the camera, this method cannot be applied to the single molecule signals considered in this work. Using camera detectors, however, has the advantage that the signals from many single molecules can be measured in parallel. This is crucial for the reconstruction of an image in SMLM.

The transition time is estimated by taking the ratio of the signal in the first frame of a single molecule trace to the average brightness in fully exposed frames. For this method to work, the single molecule trace has to be long enough, such that there is at least one full frame during which the molecule was in the on-state. As the duration of the single molecule events is exponentially distributed, there can be a large fraction of events which are too short for average on-state dwell times of two to four camera exposure times. The first frame of each single molecule event is detected using an HMM. Due to the noise in the signal, it can be misidentified in some cases, especially for transition times which are close to the boundary between two frames.

### 5.3.2. Crosstalk in multicolor SMLM

Multicolor methods for SMLM in which all channels are imaged during the same experiment [39–42,50], typically assign each single molecule event to a color channel. Independent of the method used, there is typically a small probability of assigning an event to a wrong channel. These incorrectly assigned events determine the crosstalk in multicolor SMLM. If one of the species is more abundant than the other ones, the crosstalk from this channel into the other channels will be much higher. The crosstalk in the final image can be reduced by discarding events which cannot be assigned with a high certainty [39]. With a higher number of distinct color species in the image, the crosstalk is given by the sum of the contributions of each species. Thus, it is especially important to reduce the crosstalk if more than for example two colors are imaged.

In multicolor SMLM using spectrally separated fluorophores, the crosstalk which can be achieved is typically lower than in activation-based multicolor STORM. Using two color channels for the detection of the fluorescence, a crosstalk below 10% was achieved [39,40]. Using an extra objective to measure the spectrum for each single molecule event on a second camera, the crosstalk could be reduced to below 2% [41].

In activation-based multicolor STORM, the crosstalk was typically reported to be between 10 and 20% [42]. Using the switching time analysis, we were able to reduce the crosstalk by a factor of two. In the measurements shown in Section 4.4, a crosstalk of between 3% and 5% was achieved. This is comparable with other methods for multicolor SMLM, however, the problem of image registration described in Section 1.2 is avoided.

Another aspect related to image registration in SMLM is drift correction. As typical SMLM experiments take a few minutes or even up to hours, drift on the scale of the resolution can typically not be avoided and has to be corrected in post-processing [6,89].

## 6. Outlook

The stochastic behavior of photoswitchable fluorophores was modeled by a simple two state Markov chain with a fluorescent on- and a non-fluorescent off-state. This model was applied to describe the stochastic image formation process in RESOLFT microscopy based on reversibly switchable proteins. Due to the stochastic switching behavior of the fluorophores in the focal region of a RESOLFT microscope, the variance of the measured signal is higher than for non-switching fluorophores. This excess variance was estimated from a single image to calibrate the brightness per fluorophore, with the goal of estimating the number density of fluorophores in the image.

Instead of a confocal excitation and detection, other modes of readout can be considered. The signal from the fluorescent proteins can be read out with an additional off-switching step using STED [90]. This can be included into the model by using the effective STED PSF to model the readout step.

The image formation process can be readily extended to different intensity patterns for the confinement of the on-state population to sub-diffraction volumes [91]. By taking into account the specific noise of the camera, the image formation model can be applied to in parallelized RESOLFT microscopy [92].

The calibration of the brightness from the image data may be improved by using a specific deconvolution approach with an adjusted noise model, by taking into account the joint likelihood function for the number density and the brightness. This may solve some of the problems of the deconvolution algorithm used.

If more than one image can be taken of the structure, this can improve the estimate of the brightness in several ways. First, more data can in principle increase the precision of any estimate, although photobleaching may have to be taken into account. Second, the variance of the signal can be estimated from several images without the need of the deconvolution method. These approaches could reduce the bias in the estimation of the brightness.

We introduced a method for estimating the switching transition time for single molecule events in SMLM with a sub-frame timing precision. This method was used to reduce the crosstalk due to randomly activated fluorophores in activation-based



## 6. Outlook

multicolor STORM. With a reduced crosstalk, it is possible to increase the number of color species in the sample. Without the need to register the images, activation-based STORM can be applied with a high resolution in three dimensions, using 4Pi detection [93].

The estimation of the switching time can be applied to other kinds of single molecule data, for example single molecule FRET, where the additional time resolution may be helpful.

The precision of the switching time estimate depends strongly on the stability of the signal. For single cyanine dyes, the stability depends strongly on the used buffer conditions [69]. Optimizing these buffer conditions will increase the precision in transition time measurements, and potentially lead to a further decrease of the crosstalk in activation-based multicolor STORM. To this end, the tools we developed for the measurement of the single molecule switching kinetics will be helpful.

## A. Bibliography

- [1] E. Abbe, “Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung,” *Archiv für mikroskopische Anatomie*, vol. 9, no. 1, pp. 413–418, 1873.
- [2] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics Letters*, vol. 19, no. 11, pp. 780–782, 1994.
- [3] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, “Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission,” *Proceedings of the National Academy of Sciences*, vol. 97, no. 15, pp. 8206–8210, 2000.
- [4] S. W. Hell, “Far-field optical nanoscopy,” *Science*, vol. 316, no. 5828, pp. 1153–1158, 2007.
- [5] S. W. Hell, “Microscopy and its focal switch,” *Nature Methods*, vol. 6, no. 1, pp. 24–32, 2009.
- [6] W. Vandenberg, M. Leutenegger, T. Lasser, J. Hofkens, and P. Dedecker, “Diffraction-unlimited imaging: from pretty pictures to hard numbers,” *Cell and Tissue Research*, vol. 360, no. 1, pp. 151–178, 2015.
- [7] S. W. Hell, M. Dyba, and S. Jakobs, “Concepts for nanoscale resolution in fluorescence microscopy,” *Current Opinion in Neurobiology*, vol. 14, no. 5, pp. 599–609, 2004.
- [8] R. E. Thompson, D. R. Larson, and W. W. Webb, “Precise nanometer localization analysis for individual fluorescent probes,” *Biophysical Journal*, vol. 82, no. 5, pp. 2775–2783, 2002.
- [9] K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, “Optimized localization analysis for single-molecule tracking and super-resolution microscopy,” *Nature Methods*, vol. 7, no. 5, pp. 377–381, 2010.

## A. Bibliography

- [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*, vol. 313, no. 5793, pp. 1642–1645, 2006.
- [11] M. J. Rust, M. Bates, and X. Zhuang, “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM),” *Nature Methods*, vol. 3, no. 10, pp. 793–796, 2006.
- [12] S. T. Hess, T. P. Girirajan, and M. D. Mason, “Ultra-high resolution imaging by fluorescence photoactivation localization microscopy,” *Biophysical Journal*, vol. 91, no. 11, pp. 4258–4272, 2006.
- [13] C. Geisler, A. Schönle, C. Von Middendorff, H. Bock, C. Eggeling, A. Egner, and S. Hell, “Resolution of  $\lambda/10$  in fluorescence microscopy using fast single molecule photo-switching,” *Applied Physics A*, vol. 88, no. 2, pp. 223–226, 2007.
- [14] M. Heilemann, S. Van De Linde, M. Schüttpeitz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, “Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes,” *Angewandte Chemie International Edition*, vol. 47, no. 33, pp. 6172–6176, 2008.
- [15] J. Fölling, M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling, and S. W. Hell, “Fluorescence nanoscopy by ground-state depletion and single-molecule return,” *Nature Methods*, vol. 5, no. 11, pp. 943–945, 2008.
- [16] M. Hofmann, C. Eggeling, S. Jakobs, and S. W. Hell, “Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 49, pp. 17565–17569.
- [17] T. Grotjohann, I. Testa, M. Leutenegger, H. Bock, N. T. Urban, F. Lavoie-Cardinal, K. I. Willig, C. Eggeling, S. Jakobs, and S. W. Hell, “Diffraction-unlimited all-optical imaging and writing with a photochromic GFP,” *Nature*, vol. 478, no. 7368, pp. 204–208, 2011.
- [18] T. Grotjohann, I. Testa, M. Reuss, T. Brakemann, C. Eggeling, S. W. Hell, and S. Jakobs, “rsEGFP2 enables fast RESOLFT nanoscopy of living cells,” *eLife*, vol. 1, p. e00248, 2012.

## A. Bibliography

- [19] M. Ratz, I. Testa, S. W. Hell, and S. Jakobs, “CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells,” *Scientific Reports*, vol. 5, 2015.
- [20] S. Schnorrenberg, T. Grotjohann, G. Vorbrüggen, A. Herzig, S. W. Hell, and S. Jakobs, “In vivo super-resolution RESOLFT microscopy of *Drosophila melanogaster*,” *eLife*, vol. 5, p. e15567, jun 2016.
- [21] V. C. Coffman and J.-Q. Wu, “Counting protein molecules using quantitative fluorescence microscopy,” *Trends in Biochemical Sciences*, vol. 37, no. 11, pp. 499–506, 2012.
- [22] M. H. Ulbrich and E. Y. Isacoff, “Subunit counting in membrane-bound proteins,” *Nature Methods*, vol. 4, no. 4, pp. 319–321, 2007.
- [23] H. Ta, A. Kiel, M. Wahl, and D.-P. Herten, “Experimental approach to extend the range for counting fluorescent molecules based on photon-antibunching,” *Physical Chemistry Chemical Physics*, vol. 12, no. 35, pp. 10295–10300, 2010.
- [24] H. Ta, J. Keller, M. Haltmeier, S. K. Saka, J. Schmied, F. Opazo, P. Tinnefeld, A. Munk, and S. W. Hell, “Mapping molecules in scanning far-field fluorescence nanoscopy,” *Nature Communications*, vol. 6, 2015.
- [25] P. Annibale, S. Vanni, M. Scarselli, U. Rothlisberger, and A. Radenovic, “Quantitative photo activated localization microscopy: unraveling the effects of photoblinking,” *PloS ONE*, vol. 6, no. 7, p. e22678, 2011.
- [26] S.-H. Lee, J. Y. Shin, A. Lee, and C. Bustamante, “Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM),” *Proceedings of the National Academy of Sciences*, vol. 109, no. 43, pp. 17436–17441, 2012.
- [27] R. P. Nieuwenhuizen, M. Bates, A. Szymborska, K. A. Lidke, B. Rieger, and S. Stallinga, “Quantitative Localization Microscopy: Effects of Photophysics and Labeling Stoichiometry,” *PloS ONE*, vol. 10, no. 5, p. e0127989, 2015.
- [28] F. Fricke, J. Beaudouin, R. Eils, and M. Heilemann, “One, two or three? probing the stoichiometry of membrane proteins by single-molecule localization microscopy,” *Scientific Reports*, vol. 5, 2015.

## A. Bibliography

- [29] N. Durisic, L. Laparra-Cuervo, Á. Sandoval-Álvarez, J. S. Borbely, and M. Lakadamyali, “Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate,” *Nature Methods*, vol. 11, no. 2, pp. 156–162, 2014.
- [30] G. C. Rollins, J. Y. Shin, C. Bustamante, and S. Pressé, “Stochastic approach to the molecular counting problem in superresolution microscopy,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 2, pp. E110–E118, 2015.
- [31] R. Jungmann, M. S. Avendaño, M. Dai, J. B. Woehrstein, S. S. Agasti, Z. Feiger, A. Rodal, and P. Yin, “Quantitative super-resolution imaging with qPAINT,” *Nature Methods*, 2016.
- [32] Y. Chen, J. D. Müller, P. T. So, and E. Gratton, “The photon counting histogram in fluorescence fluctuation spectroscopy,” *Biophysical Journal*, vol. 77, no. 1, pp. 553–567, 1999.
- [33] P. Kask, K. Palo, D. Ullmann, and K. Gall, “Fluorescence-intensity distribution analysis and its application in biomolecular detection technology,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 24, pp. 13756–13761, 1999.
- [34] M. A. Digman, R. Dalal, A. F. Horwitz, and E. Gratton, “Mapping the number of molecules and brightness in the laser scanning microscope,” *Biophysical Journal*, vol. 94, no. 6, pp. 2320 – 2332, 2008.
- [35] L. Kastrop, H. Blom, C. Eggeling, and S. W. Hell, “Fluorescence fluctuation spectroscopy in subdiffraction focal volumes,” *Physical Review Letters*, vol. 94, no. 17, p. 178104, 2005.
- [36] S. Geissbuehler, N. L. Bocchio, C. Dellagiacoma, C. Berclaz, M. Leutenegger, and T. Lasser, “Mapping molecular statistics with balanced super-resolution optical fluctuation imaging (bSOFI),” *Optical Nanoscopy*, vol. 1, no. 1, p. 4, 2012.
- [37] H. Bock, C. Geisler, C. A. Wurm, C. Von Middendorff, S. Jakobs, A. Schönle, A. Egner, S. Hell, and C. Eggeling, “Two-color far-field fluorescence nanoscopy based on photoswitchable emitters,” *Applied Physics B*, vol. 88, no. 2, pp. 161–165, 2007.
- [38] H. Shroff, C. G. Galbraith, J. A. Galbraith, H. White, J. Gillette, S. Olenych, M. W. Davidson, and E. Betzig, “Dual-color superresolution imaging of genet-

## A. Bibliography

- ically expressed probes within individual adhesion complexes,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 51, pp. 20308–20313, 2007.
- [39] M. Bossi, J. Fölling, V. N. Belov, V. P. Boyarskiy, R. Medda, A. Egner, C. Eggeling, A. Schönle, and S. W. Hell, “Multicolor far-field fluorescence nanoscopy through isolated detection of distinct molecular species,” *Nano Letters*, vol. 8, no. 8, pp. 2463–2468, 2008.
- [40] I. Testa, C. A. Wurm, R. Medda, E. Rothermel, C. Von Middendorf, J. Fölling, S. Jakobs, A. Schönle, S. W. Hell, and C. Eggeling, “Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength,” *Biophysical Journal*, vol. 99, no. 8, pp. 2686–2694, 2010.
- [41] Z. Zhang, S. J. Kenny, M. Hauser, W. Li, and K. Xu, “Ultrahigh-throughput single-molecule spectroscopy and spectrally resolved super-resolution microscopy,” *Nature Methods*, vol. 12, no. 10, pp. 935–938, 2015.
- [42] M. Bates, G. T. Dempsey, K. H. Chen, and X. Zhuang, “Multicolor super-resolution fluorescence imaging via multi-parameter fluorophore detection,” *ChemPhysChem*, vol. 13, no. 1, pp. 99–107, 2012.
- [43] A. Gahlmann, J. L. Ptacin, G. Grover, S. Quirin, A. R. von Diezmann, M. K. Lee, M. P. Backlund, L. Shapiro, R. Piestun, and W. Moerner, “Quantitative multicolor subdiffraction imaging of bacterial protein ultrastructures in three dimensions,” *Nano Letters*, vol. 13, no. 3, pp. 987–993, 2013.
- [44] M. Erdelyi, E. Rees, D. Metcalf, G. S. K. Schierle, L. Dudas, J. Sinko, A. E. Knight, and C. F. Kaminski, “Correcting chromatic offset in multicolor super-resolution localization microscopy,” *Optics Express*, vol. 21, no. 9, pp. 10978–10988, 2013.
- [45] R. Jungmann, M. S. Avendaño, J. B. Woehrstein, M. Dai, W. M. Shih, and P. Yin, “Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT,” *Nature Methods*, vol. 11, no. 3, p. 313, 2014.
- [46] J. Tam, G. A. Cordier, J. S. Borbely, Á. S. Álvarez, and M. Lakadamyali, “Cross-talk-free multi-color STORM imaging using a single fluorophore,” *PLoS ONE*, vol. 9, no. 7, p. e101772, 2014.
- [47] J. C. Vaughan, S. Jia, and X. Zhuang, “Ultrabright photoactivatable fluorophores created by reductive caging,” *Nature Methods*, vol. 9, no. 12, pp. 1181–1184, 2012.

## A. Bibliography

- [48] M. Dai, R. Jungmann, and P. Yin, “Optical imaging of individual biomolecules in densely packed clusters,” *Nature Nanotechnology*, 2016.
- [49] P. Annibale, M. Scarselli, M. Greco, and A. Radenovic, “Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy,” *Optical Nanoscopy*, vol. 1, no. 1, p. 9, 2012.
- [50] M. Bates, B. Huang, G. T. Dempsey, and X. Zhuang, “Multicolor super-resolution imaging with photo-switchable fluorescent probes,” *Science*, vol. 317, no. 5845, pp. 1749–1753, 2007.
- [51] G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, and X. Zhuang, “Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging,” *Nature Methods*, vol. 8, no. 12, pp. 1027–1036, 2011.
- [52] W. Moerner and D. P. Fromm, “Methods of single-molecule fluorescence spectroscopy and microscopy,” *Review of Scientific Instruments*, vol. 74, no. 8, pp. 3597–3619, 2003.
- [53] J. R. Lakowicz, *Principles of fluorescence spectroscopy*. Springer Science & Business Media, 2013.
- [54] G. T. Dempsey, M. Bates, W. E. Kowtoniuk, D. R. Liu, R. Y. Tsien, and X. Zhuang, “Photoswitching mechanism of cyanine dyes,” *Journal of the American Chemical Society*, vol. 131, no. 51, pp. 18192–18193, 2009.
- [55] J. R. Norris, *Markov chains*. No. 2, Cambridge University Press, 1998.
- [56] D. J. C. MacKay, *Information Theory, Inference & Learning Algorithms*. New York, NY, USA: Cambridge University Press, 2002.
- [57] E. T. Jaynes, *Probability theory: The logic of science*. Cambridge University Press, 2003.
- [58] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA, 2014.
- [59] D. Bourgeois and V. Adam, “Reversible photoswitching in fluorescent proteins: a mechanistic view,” *IUBMB life*, vol. 64, no. 6, pp. 482–491, 2012.
- [60] J. Keller, A. Schönle, and S. W. Hell, “Efficient fluorescence inhibition patterns for RESOLFT microscopy,” *Optics Express*, vol. 15, no. 6, pp. 3361–3371, 2007.

## A. Bibliography

- [61] B. Harke, J. Keller, C. K. Ullal, V. Westphal, A. Schönle, S. W. Hell, *et al.*, “Resolution scaling in STED microscopy,” *Optics Express*, vol. 16, no. 6, pp. 4154–4162, 2008.
- [62] B. Richards and E. Wolf, “Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system,” in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 253, pp. 358–379, The Royal Society, 1959.
- [63] W. H. Richardson, “Bayesian-based iterative method of image restoration,” *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55–59, 1972.
- [64] L. B. Lucy, “An iterative technique for the rectification of observed distributions,” *The Astronomical Journal*, vol. 79, p. 745, 1974.
- [65] P. B. Stetson, “Daophot: A computer program for crowded-field stellar photometry,” *Publications of the Astronomical Society of the Pacific*, vol. 99, no. 613, p. 191, 1987.
- [66] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [67] L. R. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [68] L. E. Baum and T. Petrie, “Statistical inference for probabilistic functions of finite state Markov chains,” *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554–1563, 1966.
- [69] T. Ha and P. Tinnefeld, “Photophysics of fluorescence probes for single molecule biophysics and super-resolution imaging,” *Annual Review of Physical Chemistry*, vol. 63, p. 595, 2012.
- [70] A. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.
- [71] R. Durbin, S. R. Eddy, A. Krogh, and G. Mitchison, *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press, 1998.



## A. Bibliography

- [72] S. van de Linde, S. Wolter, M. Heilemann, and M. Sauer, “The effect of photo-switching kinetics and labeling densities on super-resolution fluorescence imaging,” *Journal of Biotechnology*, vol. 149, no. 4, pp. 260 – 266, 2010. BioImaging - Contributions from Biology, Physics and Informatics.
- [73] H. Babcock, Y. M. Sigal, and X. Zhuang, “A high-density 3D localization algorithm for stochastic optical reconstruction microscopy,” *Optical Nanoscopy*, vol. 1, no. 1, p. 1, 2012.
- [74] L. A. Amos and A. Klug, “Arrangement of subunits in flagellar microtubules,” *Journal of Cell Science*, vol. 14, no. 3, pp. 523–549, 1974.
- [75] O. Valiron, N. Caudron, and D. Job, “Microtubule dynamics,” *Cellular and Molecular Life Sciences*, vol. 58, no. 14, pp. 2069–2084, 2001.
- [76] Y. Lin, J. J. Long, F. Huang, W. C. Duim, S. Kirschbaum, Y. Zhang, L. K. Schroeder, A. A. Rebane, M. G. M. Velasco, A. Virrueta, *et al.*, “Quantifying and optimizing single-molecule switching nanoscopy at high speeds,” *PloS ONE*, vol. 10, no. 5, p. e0128135, 2015.
- [77] T. J. Gould, D. Burke, J. Bewersdorf, and M. J. Booth, “Adaptive optics enables 3D STED microscopy in aberrating specimens,” *Optics Express*, vol. 20, no. 19, pp. 20998–21009, 2012.
- [78] J.-B. Sibarita, “Deconvolution microscopy,” in *Microscopy Techniques*, pp. 201–243, Springer.
- [79] M. Bertero, P. Boccacci, G. Desiderà, and G. Vicidomini, “Image deblurring with Poisson data: from cells to galaxies,” *Inverse Problems*, vol. 25, no. 12, p. 123006, 2009.
- [80] D. Fish, J. Walker, A. Brinicombe, and E. Pike, “Blind deconvolution by means of the Richardson–Lucy algorithm,” *Journal of the Optical Society of America A*, vol. 12, no. 1, pp. 58–65, 1995.
- [81] A. Munk and M. Pricop, “On the self-regularization property of the EM algorithm for Poisson inverse problems,” in *Statistical Modelling and Regression Structures: Festschrift in Honour of Ludwig Fahrmeir*, pp. 431–448, 2010.
- [82] T. Förster, “Zwischenmolekulare Energiewanderung und Fluoreszenz,” *Annalen der Physik*, vol. 437, no. 1-2, pp. 55–75, 1948.

## A. Bibliography

- [83] A. N. Bader, E. G. Hofman, J. Voortman, P. M. v. B. en Henegouwen, and H. C. Gerritsen, “Homo-FRET imaging enables quantification of protein cluster sizes with subcellular resolution,” *Biophysical Journal*, vol. 97, no. 9, pp. 2613–2622, 2009.
- [84] Y. Zhao, Q. Fang, A. D. Herbst, K. N. Berberian, W. Almers, and M. Lindau, “Rapid structural change in synaptosomal-associated protein 25 (SNAP25) precedes the fusion of single vesicles with the plasma membrane in live chromaffin cells,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 35, pp. 14249–14254, 2013.
- [85] S. Farooq and J. Hohlbein, “Camera-based single-molecule FRET detection with improved time resolution,” *Physical Chemistry Chemical Physics*, vol. 17, no. 41, pp. 27862–27872, 2015.
- [86] L. P. Watkins and H. Yang, “Detection of intensity change points in time-resolved single-molecule measurements,” *The Journal of Physical Chemistry B*, vol. 109, no. 1, pp. 617–628, 2005.
- [87] H. Yang, “Model-free statistical reduction of single-molecule time series,” *Theory and Evaluation of Single-Molecule Signals edited by E. Barkai, F. Brown, M. Orrit, and H. Yang, World Scientific publishing*, 2008.
- [88] J. Chen and A. K. Gupta, *Parametric statistical change point analysis: with applications to genetics, medicine, and finance*. Springer Science & Business Media, 2011.
- [89] C. Geisler, T. Hotz, A. Schönle, S. W. Hell, A. Munk, and A. Egner, “Drift estimation for single marker switching based imaging schemes,” *Optics Express*, vol. 20, no. 7, pp. 7274–7289, 2012.
- [90] J. G. Danzl, S. C. Sidenstein, C. Gregor, N. T. Urban, P. Ilgen, S. Jakobs, and S. W. Hell, “Coordinate-targeted fluorescence nanoscopy with multiple off states,” *Nature Photonics*, 2016.
- [91] U. Böhm, S. W. Hell, and R. Schmidt, “4Pi-RESOLFT nanoscopy,” *Nature Communications*, vol. 7, 2016.
- [92] A. Chmyrov, J. Keller, T. Grotjohann, M. Ratz, E. d’Este, S. Jakobs, C. Eggeling, and S. W. Hell, “Nanoscopy with more than 100,000 ‘doughnuts’,” *Nature Methods*, vol. 10, no. 8, pp. 737–740.

## A. Bibliography

- [93] D. Aquino, A. Schönle, C. Geisler, C. v Middendorff, C. A. Wurm, Y. Okamura, T. Lang, S. W. Hell, and A. Egner, “Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores,” *Nature Methods*, vol. 8, no. 4, pp. 353–359, 2011.

## B. Acknowledgements

I would like to thank Prof. Stefan W. Hell for the great opportunity to work in the Department of NanoBiophotonics. The work environment created in the Max Planck Institute for Biophysical Chemistry is ideal for great science.

The thesis committee meetings together with Prof. Axel Munk and Prof. Helmut Grubmüller were a great help to organize my thoughts about the project. I would thus also like to thank the organizers of the International Max Planck Research School, Antje Erdmann and Frauke Bergmann for the support and the structure provided by the program.

The scientific discussions with Prof. Axel Munk, Dr. Timo Aspelmeier and Miguel Del Alamo from the Institute of Mathematical Stochastics were very helpful for me, and I am looking forward to further interesting collaboration.

Dr. Jan Keller was an invaluable help during the writing process. He read the whole manuscript repeatedly and his many suggestions were very useful.

Many thanks to Dr. Mark Bates and Michael Weber for providing me with excellent STORM data to analyze, sometimes more than I could initially handle.

Philipp Alt and Sebastian Schnorrenberg have measured more switching proteins than I can count. Thank you for the great data and the great scientific discussions on the roof.

I would also like to thank Dr. Francisco Balzarotti, Dr. Ulrike Böhm and Dr. Marcel Leutenegger for reading parts of the thesis and the helpful suggestions and great discussions.

Thanks to all my colleagues at the Department of NanoBiophotonics.

Finally I would like to thank my family, my parents Karin and Holger Frahm, for the great support during my time in Göttingen and my sister Maren Frahm. Most of all I have to thank my wife Sophie Frahm for everything. I am looking forward to seeing our family grow.

## C. Curriculum vitae

### Personal information

Lars Frahm, born on 26.1.1988 in Bremen

### Education

**Max Planck Institute for Biophysical Chemistry, Göttingen**  
Doctoral Student, 2013-2016.

### Leibniz Universität Hannover

Master of Science, Physics, 2013.

Title of the thesis: “High harmonic generation with elliptically polarized laser pulses”

Bachelor of Science, Physics, 2010.

Title of the thesis: “Monte Carlo Simulation von Gittermodellen”

### Contributions to Conferences

**Single molecule localization microscopy symposium**  
Lausanne, 2016.

Poster presentation: “Multicolor STORM with low crosstalk using single molecule event timing”

### Focus on microscopy

Göttingen, 2014.

Poster presentation: “Counting molecules in RESOLFT nanoscopy”

### Labeling and Nanoscopy

Heidelberg, 2013.

Poster presentation: “Uncertainty in counting molecules in fluorescence microscopy”

### Publications

Lars Frahm & Jan Keller

“Polarization modulation adds little additional information to super-resolution fluorescence microscopy”

**Nature Methods**, vol.13, no. 1, pp. 7-8, 2016.