

Bachelor Thesis in Physics

**Optimal Conditions for Brightness-Gated Two-Color  
Coincidence Detection:  
Chance Coincidence Limitation**

by

**Lukas Hermans**

First Examiner: Prof. Dr. Jörg Fitter  
Second Examiner: Prof. Dr. Andreas Offenhäusser

RWTH Aachen University  
Faculty of Mathematics, Computer Science and Natural Sciences  
I. Institute of Physics (IA)  
AG Biophysics

August 2020



# **Erklärung des Autors**

Ich versichere, dass ich die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, den 9. August 2020



# Abstract

Many biological processes are driven by interactions between molecules diffusing through the cytosol of the cells. Bindings between two molecule types are one sort of interaction. Here, the binding fraction is a quantity of interest. It depends on the strength of the binding and the concentration of the dissolved molecules. The measurement of binding fractions is an indispensable tool for understanding biological processes because it indicates the probability of a molecular interaction. A typical device for such measurements is a confocal fluorescence microscope. In a confocal fluorescence microscope, two types of freely diffusing dye-labeled molecules in a solution are illuminated by a light source, and the emitted fluorescence photons are detected. Different molecule types are distinguished by different dye-labelings. Thus, a molecular complex in the observation volume leads to two photon bursts of different fluorescence colors. Typically, the observation volumes for both colors do neither have the same size nor overlap completely. An approach to determine the binding fraction from such measurements is called brightness-gated two-color coincidence detection (BTCCD). It applies a brightness-gating, where only photon bursts with a minimal number of photons, called brightness threshold, are selected. Thereby, dim bursts that do not correspond to trajectories through both observation volumes, and that lead to an underestimation of the coincidence fraction are excluded.

A coincidence of two photon bursts can also occur if two independent molecules are simultaneously inside the observation volume. Such events are called chance coincidences and lead to an overestimation of the binding fraction. In this thesis, the limitation of BTCCD due to the occurrence of chance coincidences is investigated. It is shown that an existing approach to correct for chance coincidences is only applicable in a limited range of sample concentrations that depends on the diffusing properties of the dissolved molecules. Based on experimental data, a method is proposed that allows choosing a suitable sample concentration, for which the chance coincidence correction is applicable. A trial experiment validates the method.

Besides the investigation of chance coincidences, other properties of BTCCD are analyzed. First of all, it is shown that roughly 10000 photon bursts need to be recorded if the uncertainty on the coincidence fraction before the correction should be smaller than 5 %.

Relevant quantities that occur during the application of BTCCD are the average number of molecules inside the observation volume, the dwell time, and the molecular brightness. In this thesis, the dependency of these quantities on the brightness threshold is investigated. Experimental data reveal that their development is guided by two processes that take place on the time scale of the dwell time. A fast process is identified as the exclusion of dim bursts due to the application of brightness-gating. The underlying sources for a second, slower process remain unknown.



# Zusammenfassung

Viele biologische Prozesse basieren auf Wechselwirkungen zwischen Molekülen, die frei durch das Cytosol der Zellen diffundieren. Eine Art von Wechselwirkungen sind Bindungen zwischen zwei Molekültypen. Dabei ist das Bindungsverhältnis ein Untersuchungsgegenstand. Es hängt von der Bindungsstärke und der Konzentration der Moleküle ab. Die Messung von Bindungsverhältnissen hat eine hohe Bedeutung für die Untersuchung von biologischen Prozessen, weil es ein Indikator für die Wahrscheinlichkeit von molekularen Wechselwirkungen ist. Ein verbreitetes Messgerät ist das konfokale Fluoreszenzmikroskop. Dabei werden zwei Typen von frei diffundierenden Farbstoff-markierten Molekülen in einer Lösung von einer Lichtquelle beleuchtet und die emittierten Fluoreszenzphotonen detektiert. Die Molekültypen werden durch verschiedene Farbstoffmarkierungen unterschieden. Daher führt die Anwesenheit eines Molekülkomplexes im Beobachtungsvolumen zu zwei koinzidenten "Photonschwallen" verschiedener Wellenlänge. Typischerweise sind die Beobachtungsvolumen beider Farben nicht gleich groß und zudem relativ zueinander verschoben. Eine Methode zur Bestimmung des Bindungsverhältnisses aus derartigen Messungen wird helligkeitsgeleitete Zweifarben-Koinzidenzdetektion (BTCCD) genannt. Dabei wird eine Helligkeitsschwelle angewendet, durch die nur Schwalle mit einer Mindestanzahl von Photonen ausgewählt werden. Dadurch werden schwache Schwalle, die nicht aus zentralen Trajektorien durch beide Beobachtungsvolumen resultieren und zu einem unterschätzten Koinzidenzverhältnis führen, aussortiert.

Das Auftreten zweier koinzidenter Photonschwallen kann auch auf die gleichzeitige Präsenz zwei unabhängiger Moleküle zurückzuführen sein. Solche Ereignisse werden Zufallskoinzidenzen genannt. Sie führen zu einem systematisch überschätzten Bindungsverhältnis. In dieser Arbeit wird untersucht, inwiefern Zufallskoinzidenzen die Anwendung von BTCCD limitieren. Dabei wird gezeigt, dass ein existierendes Verfahren zur Korrektur von Zufallskoinzidenzen nur in einem beschränkten Bereich im Bezug auf die Probenkonzentration angewendet werden kann. Der Bereich hängt von den Diffusionseigenschaften der Probenmoleküle ab. Basierend auf experimentellen Ergebnissen wird eine Methode entwickelt, die es erlaubt eine geeignete Probenkonzentration zu wählen, für die die Korrektur von Zufallskoinzidenzen verwendbar ist. Die Methode wird durch ein Beispielexperiment validiert.

Neben der Untersuchung von Zufallskoinzidenzen werden weitere Eigenschaften von BTCCD analysiert.

Zunächst wird gezeigt, dass etwa 10000 Photonschalle registriert werden müssen, um das unkorrigierte Koinzidenzverhältnis mit einer Unsicherheit von weniger als 5 % bestimmen zu können.

Relevante Größen bei der Verwendung von BTCCD sind neben der mittleren Anzahl der Moleküle im Beobachtungsvolumen die Verweildauer der Moleküle und die molekulare Helligkeit. In der vorliegenden Arbeit wird die Abhängigkeit dieser Größen von der Helligkeitsschwelle untersucht. Experimentell wird festgestellt, dass die Entwicklung der Größen durch zwei Prozesse bestimmt wird, die auf der Zeitskala der Verweildauer stattfinden. Ein schneller Prozess wird auf den Ausschluss schwacher Photonschalle durch das Anwenden der Helligkeitsschwelle zurückgeführt. Die zugrundeliegenden Ursachen für einen zweiten, langsameren Prozess bleiben unbekannt.



# Contents

<b>Abstract</b>	<b>V</b>
<b>Zusammenfassung</b>	<b>VII</b>
<b>List of Tables</b>	<b>XI</b>
<b>List of Figures</b>	<b>XIV</b>
<b>List of Acronyms</b>	<b>XV</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theory and Methods</b>	<b>3</b>
2.1 Fluorescence . . . . .	3
2.2 Confocal Fluorescence Microscopy . . . . .	5
2.2.1 Setup . . . . .	5
2.2.2 Measurements . . . . .	6
2.3 Fluorescence Correlation Spectroscopy (FCS) . . . . .	8
2.4 Brightness-Gated Two-Color Coincidence Detection (BTCCD) . . . . .	10
2.4.1 Determination of Coincidence Fraction . . . . .	11
2.4.2 Correction for Chance Coincidences . . . . .	16
2.4.3 Dependencies on Brightness Threshold . . . . .	17
<b>3 Materials and Instrumentation</b>	<b>19</b>
3.1 Samples . . . . .	19
3.1.1 Fluorescent Dyes . . . . .	19
3.1.2 DNA . . . . .	19
3.1.3 Ribosomes . . . . .	19
3.2 Buffers . . . . .	19
3.3 Microscope Slides . . . . .	20
3.4 Confocal Fluorescence Microscope . . . . .	20
3.5 Software . . . . .	21
<b>4 Dependencies on Brightness Threshold</b>	<b>22</b>
4.1 Measurement . . . . .	22
4.2 Results . . . . .	22
4.2.1 Dwell Time . . . . .	22
4.2.2 Number of Molecules . . . . .	26
4.2.3 Molecular Brightness . . . . .	28
4.3 Two Underlying Processes . . . . .	31
<b>5 Optimal Number of Bursts</b>	<b>32</b>
5.1 Measurement . . . . .	32
5.2 Results and Discussion . . . . .	32

<b>6 Chance Coincidence Limitation</b>	<b>35</b>
6.1 Application Limits of Chance Coincidence Correction . . . . .	35
6.1.1 Measurement . . . . .	35
6.1.2 Results . . . . .	36
6.1.3 Systematic Deviation of Corrected Coincidence Fraction . . . . .	39
6.1.4 Application Range of Chance Coincidence Correction . . . . .	42
6.2 Validation of Application Limits . . . . .	44
6.2.1 Measurement . . . . .	44
6.2.2 Results and Discussion . . . . .	44
<b>7 Conclusion and Outlook</b>	<b>48</b>
<b>Bibliography</b>	<b>51</b>
<b>A Appendix: Theory and Methods</b>	<b>52</b>
A.1 Procedure of BTCCD . . . . .	52
A.2 Uncertainty on Dwell Time, Number of Molecules, and Molecular Brightness . . . . .	54
A.3 Uncertainty on Corrected Coincidence Fraction . . . . .	55
<b>B Appendix: Dependencies on Brightness Threshold</b>	<b>56</b>
B.1 Dwell Time . . . . .	56
B.2 Number of Molecules . . . . .	58
B.3 Molecular Brightness . . . . .	60
<b>C Appendix: Optimal Number of Bursts</b>	<b>62</b>
<b>D Appendix: Chance Coincidence Limitation</b>	<b>64</b>
D.1 Application Limits of Chance Coincidence Correction . . . . .	64
D.2 Validation of Application Limits . . . . .	72

# List of Tables

4.1	Burst threshold, background, and starting parameters for measurement of dual-labeled dsDNA . . . . .	22
4.2	Parameters of linear fit of dwell time . . . . .	24
4.3	Parameters of biexponential fit of dwell time . . . . .	25
4.4	Parameters of exponential fit of molecule number . . . . .	26
4.5	Parameters of biexponential fit of molecule number . . . . .	28
4.6	Parameters of exponential fit of molecular brightness . . . . .	29
4.7	Parameters of biexponential fit of molecular brightness . . . . .	30
4.8	Ranges of two underlying processes . . . . .	31
5.1	Burst threshold, background, and starting parameters for measurement of dual-labeled dsDNA, conducted by Olessya Yukhnovets (RWTH Aachen, Germany) . .	32
6.1	Measurement time, background, and starting parameters for mixture of red-labeled dsDNA and free blue dye . . . . .	36
6.2	Measurement time, burst thresholds, background, and starting parameters for mixture of red-labeled ribosomes and free blue dye . . . . .	36
6.3	Measurement time, burst thresholds, background, and starting parameters for measurements of mixture of red-labeled dsDNA and dual-labeled dsDNA . . . .	44
6.4	Corrected coincidence fraction of red channel at optimal brightness threshold for a mixture of red-labeled dsDNA and dual-labeled dsDNA . . . . .	45
6.5	Corrected coincidence fraction of red channel at $n_{br} = 6.40$ for a mixture of red-labeled dsDNA and dual-labeled dsDNA . . . . .	45
D.1	Corrected coincidence fraction of blue channel at optimal brightness threshold for a mixture of red-labeled dsDNA and dual-labeled dsDNA . . . . .	72
D.2	Corrected coincidence fraction of blue channel at $n_{br} = 3.53$ for a mixture of red-labeled dsDNA and dual-labeled dsDNA . . . . .	72



# List of Figures

2.1 Example of a Jablonski diagram . . . . .	3
2.2 Absorption and emission spectrum of Alexa 647 . . . . .	4
2.3 Sketch of used confocal fluorescence microscope . . . . .	5
2.4 Format of recorded measurement data . . . . .	7
2.5 Different size and incomplete overlap of confocal detection volumes . . . . .	10
2.6 IPL time trace and definition of bursts . . . . .	12
2.7 Determination of background in IPL time trace for red channel . . . . .	13
2.8 Coincidence fraction as a function of $n_{br}$ . . . . .	14
2.9 Search for optimal brightness threshold $n_{br,opt}$ for red channel . . . . .	15
4.1 Linear fit of dwell time for red channel . . . . .	23
4.2 Biexponential fit of dwell time for red channel . . . . .	24
4.3 Exponential fit of molecule number for red channel . . . . .	26
4.4 Biexponential fit of molecule number for red channel . . . . .	27
4.5 Exponential fit of molecular brightness for red channel . . . . .	29
4.6 Biexponential fit of molecular brightness for red channel . . . . .	30
5.1 Optimal number of selected bursts for red channel . . . . .	33
5.2 Optimal number of initial bursts for red channel . . . . .	34
6.1 Coincidence fraction of red channel for mixture of red-labeled dsDNA and free blue dye . . . . .	37
6.2 Coincidence fraction of red channel for mixture of red-labeled ribosomes and free blue dye . . . . .	38
6.3 Corrected coincidence fraction of red channel for mixture of red-labeled dsDNA and free blue dye . . . . .	38
6.4 Corrected coincidence fraction of red channel for mixture of red-labeled ribosomes and free blue dye . . . . .	39
6.5 First term of corrected coincidence fraction of red channel for mixture of red-labeled dsDNA and free blue dye . . . . .	41
6.6 First term of corrected coincidence fraction of red channel for mixture of red-labeled ribosomes and free blue dye . . . . .	41
6.7 Application range of chance coincidence correction for red channel for measurements with dsDNA and ribosomes . . . . .	43
6.8 Interpolated application range of chance coincidence correction for red channel for measurements with dsDNA and ribosomes . . . . .	43
6.9 Positions of validation measurements in application range illustration for red channel . . . . .	46
A.1 Determination of background in IPL time trace for blue channel . . . . .	52
A.2 Search for optimal brightness threshold $n_{br,opt}$ for blue channel . . . . .	53
B.1 Linear fit of dwell time for blue channel . . . . .	56
B.2 Biexponential fit of dwell time for blue channel . . . . .	57

B.3 Exponential fit of molecule number for blue channel . . . . .	58
B.4 Biexponential fit of molecule number for blue channel . . . . .	59
B.5 Exponential fit of molecular brightness for blue channel . . . . .	60
B.6 Biexponential fit of molecular brightness for blue channel . . . . .	61
C.1 Optimal number of selected bursts for blue channel . . . . .	62
C.2 Optimal number of initial bursts for blue channel . . . . .	63
D.1 Coincidence fraction of blue channel for mixture of red-labeled dsDNA and free blue dye . . . . .	64
D.2 Coincidence fraction of blue channel for mixture of red-labeled ribosomes and free blue dye . . . . .	65
D.3 Corrected coincidence fraction of blue channel for mixture of red-labeled dsDNA and free blue dye . . . . .	66
D.4 Corrected coincidence fraction of blue channel for mixture of red-labeled ribosomes and free blue dye . . . . .	67
D.5 First term of corrected coincidence fraction of blue channel for mixture of red-labeled dsDNA and free blue dye . . . . .	68
D.6 First term of corrected coincidence fraction of blue channel for mixture of red-labeled ribosomes and free blue dye . . . . .	69
D.7 Application range of chance coincidence correction for blue channel for measurements with dsDNA and ribosomes . . . . .	70
D.8 Interpolated application range of chance coincidence correction for blue channel for measurements with dsDNA and ribosomes . . . . .	71
D.9 Positions of validation measurements in application range illustration for blue channel . . . . .	73

# List of Acronyms

**APD** avalanche photodiode

**BTCCD** brightness-gated two-color coincidence detection

**dsDNA** double-stranded DNA

**FCS** fluorescence correlation spectroscopy

**IPL** inter-photon lag

**PIE** pulsed interleaved excitation

**ssDNA** single-stranded DNA

**TCCD** two-color coincidence detection



# 1 Introduction

The basic unit of every living organism is the biological cell. In the cytosol of the cell, diverse interactions between biomolecules, such as DNA, RNA, and proteins, occur. Biology investigates the connection between those molecular processes and the functionality of life. Physicists develop the tools to measure molecular processes. Thus, the application of physical methods and investigation techniques for purposes of biology is called biophysics. Considering that the measurement of molecular processes is indispensable for the understanding of living organisms, the relevance of biophysics becomes evident [1].

Bindings between two molecule types are one sort of interaction that drives biological processes. Here, the binding fraction is a quantity of interest. It depends on the strength of the binding and the concentration of the dissolved molecules. The binding fraction indicates the probability of a molecular interaction between the two molecule types. Hence, the measurement of binding fractions is essential for an understanding of biological processes [2].

Fluorescence-based techniques are one method to measure binding fractions. There, the fluorescence response of a molecule after the excitation with light is measured. Most biomolecules are not fluorescent in the visible spectrum. In this case, a typical approach is to attach a fluorescent dye to the biomolecule in order to make it observable.

The advantage of fluorescence is its selectivity. Only molecules that carry a fluorescent component can be detected. Therefore, it is only sensitive for the molecules of interest, and background, e.g., of a non-fluorescent solvent, is suppressed. Since fluorescence does not alter the properties of a biomolecule, it is an appropriate technique to examine its biological functionality. Finally, fluorescence allows the investigation of dynamic processes as the fluorophore is a part of the biomolecule and moves with it [2].

An approach to determine the binding fraction between two molecule types is called brightness-gated two-color coincidence detection (BTCCD). If the two molecule types are labeled with different dyes, the simultaneous occurrence of two fluorescence colors indicates the presence of a molecular complex inside the observation volume. Typically, the observation volumes for two different wavelengths have a different size and do not overlap completely. Thus, the coincidence fraction is underestimated. BTCCD takes this systematic deviation into account and selects only molecular trajectories that lead through both observation volumes. Another phenomenon is the random presence of two different single-labeled molecules inside the observation volume at the same time. These chance coincidences do not result from an actual chemical complex, and should, thus, not increase the binding fraction. It has already been proposed an approach to correct for chance coincidences [3].

The application of BTCCD requires the detection of single molecules, and, thus, low sample concentrations. A typical criterion is that the average number of molecules per observation volume needs to be less than 0.03 to make the probability of multi-molecule events sufficiently small [4].

The experimental part of this thesis is divided into three chapters. Chapter 4 is concerned with relevant quantities that occur during the application of BTCCD. Mathematical dependencies

for those quantities, derived under several theoretical assumptions, are compared with experimental results, and, if necessary, adjusted. In Chapter 5, the required measurement size to obtain the coincidence fraction with an uncertainty of less than 5 % is investigated. The central part of the thesis is Chapter 6. There, the application limits of the existing correction for chance coincidences are examined. Phenomenologically, a method to decide prior to a measurement if the correction is suitable for particular measurement parameters is proposed. A trial measurement validates the method.

## 2 Theory and Methods

### 2.1 Fluorescence

The basis of the confocal microscopy technique applied for the experiments in this thesis is fluorescence. Fluorescence is one type of photoluminescence which, in general, describes the emission of photons from a material after the absorption of light [2]. Photoluminescence occurs mostly for molecules with aromatic compounds such as benzene [5]. A profound understanding of photoluminescence requires an extensive quantum mechanical analysis. Therefore, this section only gives a phenomenological description.

Similar to the well-known particle in a box problem, the electrons in a molecule can take different energy levels that can be visualized in a Jablonski diagram, see Figure 2.1. The electronic energy levels are associated with molecular orbitals. They are denoted by  $S_0$ ,  $S_1$ , and  $S_2$  for singlet states and by  $T_1$  for a triplet state.  $S_0$  is the highest occupied molecular orbital in the ground state. Due to rotational and vibrational degrees of freedom for molecules, there are additional energy levels that further split up the electronic states. The electron in  $S_0$  can be transferred into a higher electronic state by absorption of a photon with enough energy. The excited electron can be moved back to lower energetic states by non-radiative processes, where no photons are emitted. Another possibility is the direct transfer of the electron from  $S_1$  to  $S_0$  under the emission of a photon. This process is called fluorescence. A less likely path is the transition of the electron from  $S_1$  into the triplet state  $T_1$ , called intersystem crossing. The transition from  $T_1$  to  $S_0$  under photon emission is called phosphorescence. Since it is in principle quantum mechanically forbidden, the lifetime of an electron in a triplet state is usually around  $10^{-6}$  s. In comparison, the time that an electron remains in a particular singlet state lies at around  $10^{-9}$  s [6].

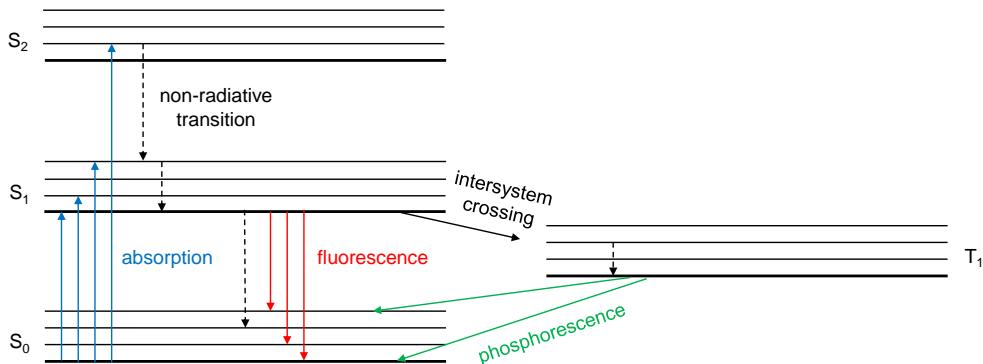


Figure 2.1: Example of a Jablonski diagram. The different energy states and transition processes of an electron in a molecule are shown.  $S_0$  is the highest occupied molecular orbital in the ground state.

A common way to classify a fluorophore is to measure its absorption and emission spectrum. The spectra are continuous due to the variety of energy states and transitions in a molecule

(Figure 2.1 shows only a few of them for illustration). An example of the red fluorescent dye Alexa 647 can be found in Figure 2.2. In the figure, typical characteristics of the spectra of a fluorophore can be observed. First of all, the emission spectrum is shifted to higher wavelengths because before fluorescence occurs, the excited electron loses energy by non-radiative transitions. This effect is called Stokes shift and is an essential requirement for the functionality of a confocal fluorescence microscope since it allows to distinguish between incoming light and fluorescence light [6]. Furthermore, the emission spectrum is a mirror image of the absorption spectrum because the vibrational and rotational energy states of the electronic states  $S_0$  and  $S_1$  have the same patterns. Finally, the emission spectrum is independent of the wavelength of the incoming light since fluorescence mostly occurs from the vibrational ground state of an energy level, e.g., from the ground state of  $S_1$  [2].

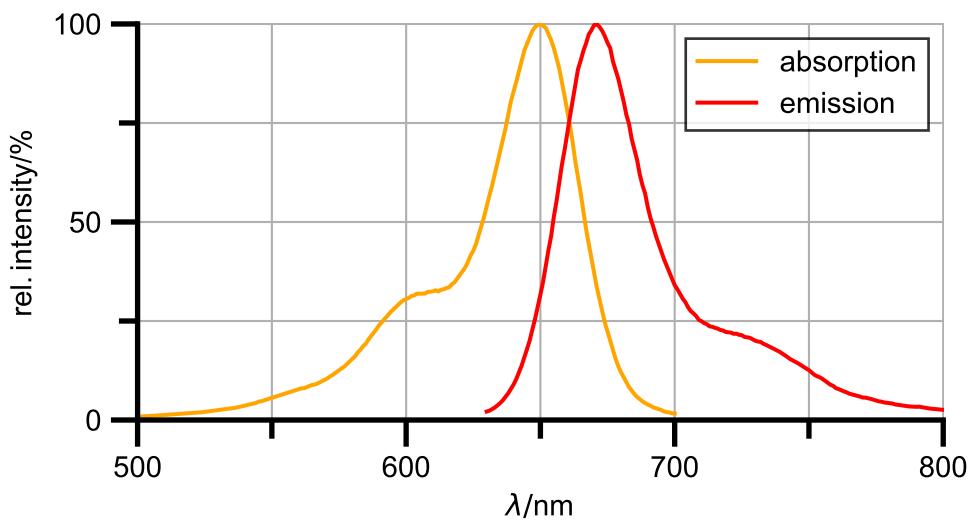


Figure 2.2: Absorption and emission spectrum of Alexa Fluor™ 647 (Thermo Fisher Scientific Inc., Waltham, USA) [7]. The intensity normalized on its maximum value is plotted against the wavelength  $\lambda$ .

## 2.2 Confocal Fluorescence Microscopy

In contrast to wide-field microscopy, a confocal microscope illuminates only a small part of the object under investigation. Thereby, background noise is suppressed, and resolution is increased [2]. In this thesis, molecules dissolved in a solution are examined. Here, confocal fluorescence microscopy allows the detection of single molecules, which is essential for the understanding of biological processes, as discussed in Chapter 1.

First, this section presents the setup and the functionality of the used confocal fluorescence microscope. After that, in Section 2.2.2, the procedure for measurements, and the format of the recorded data is discussed.

### 2.2.1 Setup

A sketch of the used confocal fluorescence microscope can be found in Figure 2.3. The objective focuses the red and blue laser beams on a small part of the solution containing the investigated samples. The samples are either intrinsic fluorophores or dye-labeled molecules that can be excited by the incoming laser light. Since the emitted fluorescence light is Stokes shifted, it can pass the dichroic mirror, while back-scattered laser light is deflected. The fluorescence light is focused on a pinhole. Thereby, it is ensured that only light from the focus of the objective is transmitted. Another dichroic mirror separates the red and blue fluorescence light. There is a detector for each color of fluorescence. For the detectors, avalanche photodiodes (APDs) were used because they offer photon detection efficiencies at around 70 % and a time resolution down to 350 ps [8, 9]. Therefore, they allow the counting of single photons, which is essential for BTCCD, see Section 2.4.

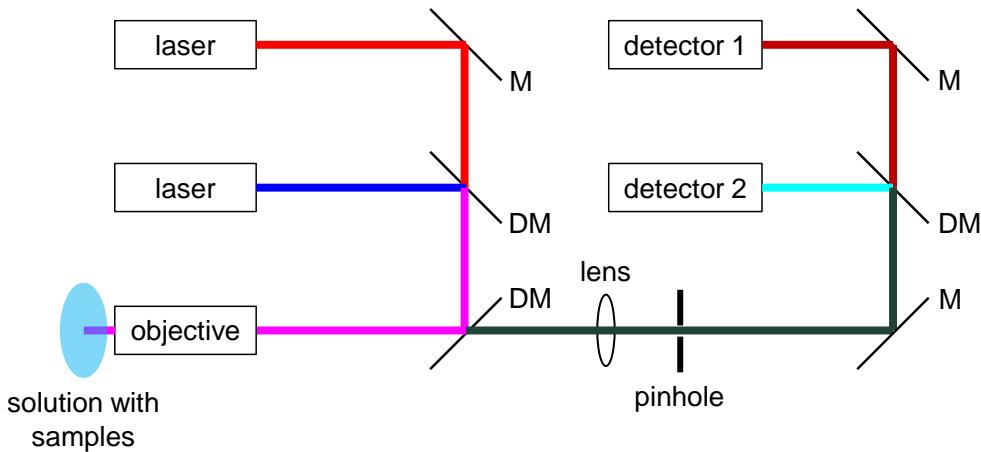


Figure 2.3: Sketch of the used confocal fluorescence microscope [10, 11]. (D)M stands for (dichroic) mirror.

The combination of objective, shape of the incoming laser beam, and pinhole leads to a spatial distribution of the fluorescence intensity in the sample solution, the so-called confocal detection volume. It can be described by a three-dimensional Gaussian

$$I(x, y, z) = I_0 \cdot \exp\left(-2 \frac{x^2 + y^2}{r_0^2}\right) \cdot \exp\left(-2 \frac{z^2}{z_0^2}\right), \quad (2.1)$$

where the  $xy$ -plane is perpendicular to the direction of the incoming laser beam and  $I_0$  is the peak intensity. Using the parameters  $r_0$  and  $z_0$ , an effective volume

$$V_{eff} = \pi^{3/2} \cdot r_0^2 \cdot z_0 \quad (2.2)$$

can be defined [11]. Since  $r_0$  and  $z_0$  depend on the properties of the objective and the wavelength, the confocal detection volumes are different for both colors. The shorter wavelength of the blue laser leads to a smaller confocal detection volume compared to the red channel. The properties of the objective depend on the wavelength, and lead to a relative shift of both volumes, see Figure 2.5.

For all measurements with the confocal fluorescence microscope, first, the focus of the objective was set on the surface of the cover glass. Then, it was raised by  $10\text{ }\mu\text{m}$  into the solution. Thereby, it is ensured that molecules that stick to the cover glass do not disturb the measurement.

### 2.2.2 Measurements

Each laser was operated at a frequency of  $10\text{ MHz}$ . Therefore, a laser pulse of the same color was emitted every  $100\text{ ns}$ . The two lasers together were used alternating. This means that the time difference between the emission of a red and a blue laser pulse was set to  $50\text{ ns}$ . This excitation scheme is called pulsed interleaved excitation (PIE). Since only one laser is powering at a time, its main advantage is the reduction of spectral cross-talk caused by non-ideal optical elements in the setup such as dichroic mirrors [12].

A measurement with the confocal fluorescence microscope consists of the detection of photons by the detectors in a certain measurement time  $T$ . For each detected photon, three parameters are saved:

- channel number: number of the detector that recorded the photon
- micro time  $\tau$ : time difference between the photon detection and the last red laser pulse
- macro time  $t$ : absolute time of the photon detection

Figure 2.4 visualizes the difference between micro time and macro time, as well as the format of the recorded data. The result of every measurement with the confocal fluorescence microscope is a list of detected photons containing the three information above. Diverse evaluation methods of the measured data have been proposed [13]. In this thesis, the methods fluorescence correlation spectroscopy (FCS) and BTCCD are applied, further discussed in Sections 2.3 and 2.4. Both of them only use the channel number and the macro time while ignoring the micro time.

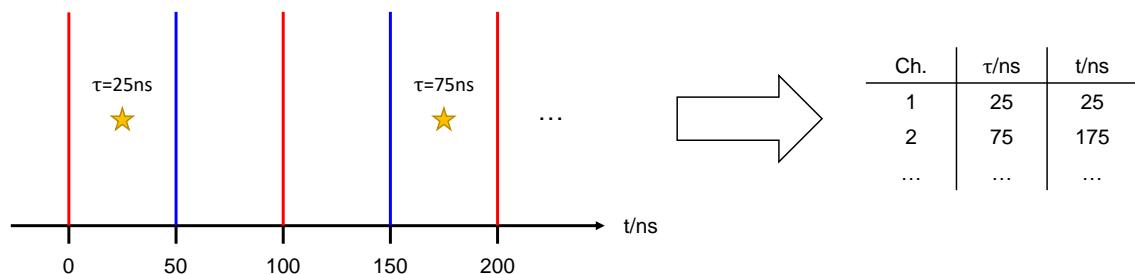


Figure 2.4: Format of the recorded measurement data. The red and blue lines symbolize the laser pulses, whereas the yellow stars indicate the detection of a photon. For each photon, channel number, micro time  $\tau$ , and macro time  $t$  are recorded.

## 2.3 Fluorescence Correlation Spectroscopy (FCS)

One method to process the data acquired with the confocal microscope is fluorescence correlation spectroscopy (FCS). In contrast to BTCCD, it does not require that the probability for more than one molecule inside the confocal detection volume is negligible. Thus, it can also be used for higher concentrated sample solutions.

FCS is based on the normalized auto-correlation function that is in general defined as

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t + \tau) \rangle_t}{\langle F(t) \rangle_t^2}, \quad (2.3)$$

where  $\delta F(t) = F(t) - \langle F(t) \rangle_t$ . Here,  $F$  is the fluorescence intensity of one channel that depends on the macro time  $t$ .  $\langle \cdot \rangle_t$  denotes the time average.  $G$  is a measure for the self-similarity of the fluorescence signal after a lag time  $\tau$ . Note, that the lag time should not be confused with the micro time defined in the previous section.

On the one hand,  $G$  can be calculated directly from the measurement data since the detected number of photons per time and, thus, the fluorescence intensity  $F$  is known. On the other hand, a theoretical expression for  $G$  can be derived. Its specific form depends on the properties of the dissolved sample. If a three-dimensional Gaussian describes the confocal detection volume, see Equation (2.1), the normalized auto-correlation function for freely diffusing particles with triplet states of lifetime  $\tau_{triplet}$  is given by

$$G(\tau) = G(0) \cdot \left( 1 + \frac{T}{1-T} \cdot e^{-\tau/\tau_{triplet}} \right) \cdot \frac{1}{1 + \frac{4D\tau}{r_0^2}} \cdot \frac{1}{\sqrt{1 + \frac{4D\tau}{\kappa^2 r_0^2}}}. \quad (2.4)$$

Here,  $T$  is the average fraction of particles in a triplet state,  $D$  the diffusion coefficient of the sample, and  $\kappa = z_0/r_0$  [11].

FCS can be used to determine the effective volume  $V_{eff}$ , defined by Equation (2.2). For this purpose, the diffusion coefficient  $D$  has to be known. Its value for typical reference samples and its temperature dependency can be found in [14]. Then, the measurement data can be fitted with Equation (2.4), where the fit parameters are  $G(0)$ ,  $T$ ,  $\tau_{triplet}$ ,  $r_0$ , and  $\kappa$ . Finally, the effective volume is [15]

$$V_{eff} = \pi^{3/2} \cdot r_0^3 \cdot \kappa. \quad (2.5)$$

Once having determined the effective volume for both channels of the setup, FCS was mainly applied to identify the concentration  $C$  of a sample. To do so, the measurement results were again fitted with Equation (2.4). Then, the relation

$$G(0) = \frac{1}{C \cdot V_{eff}} \quad (2.6)$$

allows to calculate the sample concentration. In some cases, the average number of molecules  $\langle N(t) \rangle_t$  inside the effective volume instead of the concentration is the quantity of interest. The

fluorescence intensity  $F$  is proportional to  $N$ . Taking into account that  $N$  is Poisson distributed, yields

$$G(0) = \frac{\langle (\delta F(t))^2 \rangle_t}{\langle F(t) \rangle_t^2} = \frac{\langle (\delta N(t))^2 \rangle_t}{\langle N(t) \rangle_t^2} = \frac{\langle N(t) \rangle_t}{\langle N(t) \rangle_t^2} = \frac{1}{\langle N(t) \rangle_t}. \quad (2.7)$$

Finally, the comparison of Equations (2.6) and (2.7) leads to

$$\langle N(t) \rangle_t = C \cdot V_{eff} \quad (2.8)$$

so that FCS allows the determination of both the sample concentration  $C$  as well as the average number of molecules  $\langle N(t) \rangle_t$  inside the confocal detection volume [11].

Furthermore, a fit of  $G$  gives a value for the diffusion coefficient  $D$  of the sample, so that the diffusion properties can be checked. There are diverse reasons for possible deviations of the measured diffusion coefficient from an expectation. If the measured value for  $D$  is lower, the sample molecules can, for instance, aggregate to larger complexes. If it is higher, the molecules may be partially dissociated into smaller components, or free fluorescent dyes could be present.

## 2.4 Brightness-Gated Two-Color Coincidence Detection (BTCCD)

A major biological application of confocal fluorescence microscopy is the determination of binding fractions between two types of molecules. One method to do so is called two-color coincidence detection (TCCD). The idea of TCCD is that a dye-labeled molecule diffusing through the confocal detection volume causes an intensity peak in the corresponding channel. If there are observed two peaks in both channels at the same time, it can be concluded that there were both fluorescent dyes present in the confocal detection volume. This can be caused by two separate molecules inside the confocal detection volume at the same time, which is called a chance coincidence. However, if the sample concentration is low enough, it is more likely that the coinciding peaks were caused by a chemical complex of the two molecule types. Since the number of non-coincident peaks for both channels is known, TCCD allows the calculation of the binding fraction for both molecule types [16].

Typically, TCCD underestimates the binding fraction because the confocal detection volumes for both channels have a different size and do not overlap completely. Usually, the red confocal detection volume is larger than the blue one. This situation is depicted in Figure 2.5. Three types of trajectories through the confocal detection volumes can be distinguished. If the path of the diffusing molecule only goes through the edge of the red confocal detection volume, it is a borderline trajectory (I). A path through the center of the red confocal detection volume but only through the edge of the blue one is called peripheral trajectory (II). A central trajectory (III) is present if the path goes through the center of both confocal detection volumes. Because of this situation, a complex of both molecule types can pass the confocal detection volume of one channel without causing an intensity peak in the other channel, which leads to an underestimation of the binding fraction. An approach to take the different confocal detection volumes into account is called brightness-gated two-color coincidence detection (BTCCD) [3].

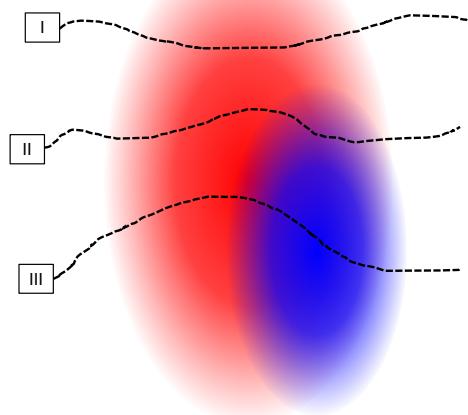


Figure 2.5: Red and blue confocal detection volume. Typically, the confocal detection volumes have a different size and do not overlap completely. Therefore, borderline (I), peripheral (II), and central trajectories (III) can be distinguished [3].

This section is dedicated to the principles of BTCCD. First of all, in Section 2.4.1, the procedure to determine the coincidence fraction is described in detail. Since chance coincidences become relevant for BTCCD even for low concentrations, in Section 2.4.2 an existing analytic approach to correct for them is presented. Finally, in Section 2.4.3, mathematical expressions for three important quantities that occur during the application of BTCCD are derived.

### 2.4.1 Determination of Coincidence Fraction

The procedure to determine the coincidence fraction with BTCCD taking the different trajectory types into account consists of four steps:

#### 1 Calculate IPL Time Trace

By using the inter-photon lag (IPL) time trace instead of the intensity trace, BTCCD takes advantage of the single-photon detection of the APDs. The IPL time trace can be calculated from the macro time  $t(n)$ , where  $n$  indicates the number of the detected photon. It is defined by

$$IPL(n) = t(n) - t(n-1), \quad (2.9)$$

such that for every detected photon (except the first one) the time difference to the occurrence of the previous photon is calculated. To stabilize the IPL time trace, a moving average

$$IPL_m(n) = \sum_{i=n-m}^{n+m} \frac{IPL(i)}{2m+1} \quad (2.10)$$

is applied. Throughout this thesis,  $m$  was set to 2. It has been shown that the results of BTCCD do not depend significantly on the choice of  $m$  [3]. An excerpt of an IPL time trace can be found in Figure 2.6. There, the raw IPL time trace as well as the smoothed one is shown.

#### 2 Set Burst Threshold

For the determination of the coincidence fraction, it is necessary to define what is considered as a photon burst, corresponding to a peak in the intensity trace. For this purpose, burst thresholds  $IPL^{tre}$  for both channels are defined. Every time the IPL time trace falls below this threshold, a photon burst is present. As illustrated in Figure 2.5 above, three types of trajectories through the confocal detection volume can be distinguished. They are also identifiable in the IPL time trace as shown in Figure 2.6. A borderline trajectory (I) leads to a dim burst in the red channel, while there is no burst in the blue one. If the blue IPL time trace drops slightly below the burst threshold, the molecule complex followed a peripheral trajectory (II). A central trajectory (III) corresponds to bright bursts in both channels.

Substantial quantities that follow from the definition of a photon burst are the molecular brightness  $MB$  and the dwell time  $\tau_d$ . The dwell time is the width of a burst, while the molecular brightness is given by

$$MB = \frac{N_\gamma}{\tau_d}, \quad (2.11)$$

where  $N_\gamma$  is the number of photons in the burst with dwell time  $\tau_d$ . Both quantities can be calculated for every single burst. In practice, their average values for all bursts in one channel are of interest. Therefore, in the following,  $\langle MB \rangle$  and  $\langle \tau_d \rangle$  denote these average values, and it

is not always explicitly stated that they are calculated from a certain number of bursts. For the analysis of the measurements, the brightness thresholds were set such that  $\langle MB \rangle$  and  $\langle \tau_d \rangle$  are in rough agreement with the values obtained by FCS.<sup>1</sup>

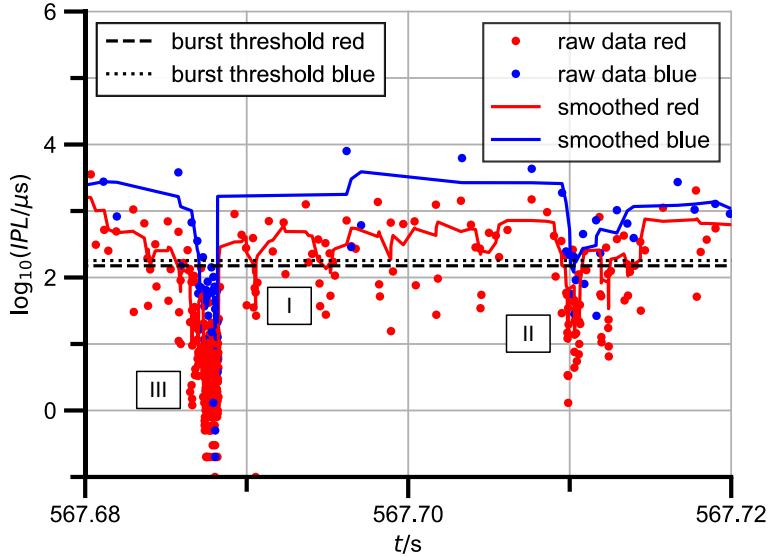


Figure 2.6: IPL time trace calculated from macro time trace. Photon bursts are distinguished from the background by a burst threshold. Bursts resulting from borderline (I), peripheral (II), and central trajectories (III) can be identified, corresponding to Figure 2.5. The data is taken from a measurement of dual-labeled dsDNA, see Section 4.1.

Every measurement with the confocal microscope is, to a certain degree, disturbed by external sources, e.g., dark currents in the detection electronics. If the effect is not too large, the associated IPL values are larger than those caused by a molecule passing through the confocal detection volume. The average background  $\langle IPL^{bg} \rangle$  can be estimated by plotting a histogram of the measured IPL values. An example of the red channel can be found in Figure 2.7. The histogram reveals two separate populations. The right population corresponds to higher IPL values and reflects thus the background. Since the right population is almost symmetrical, its maximum can be taken as a value for  $\langle IPL^{bg} \rangle$ . The corresponding background characterization for the blue channel is illustrated in Figure A.1. The burst threshold  $IPL^{tre}$  should be smaller than the average background to prevent the misidentification of the background as bursts.

<sup>1</sup>Although FCS uses the intensity time trace, an average dwell time can be derived from the measured diffusion coefficient:  $\langle \tau_d \rangle = r_0^2/(4D)$ . Then, the molecular brightness  $\langle MB \rangle$  is simply the average number of photons in an intensity peak  $\langle N_\gamma \rangle$  divided by the average dwell time  $\langle \tau_d \rangle$  as well as by the average number of molecules inside the confocal detection volume  $\langle N \rangle$ :  $\langle MB \rangle = \langle N_\gamma \rangle / (\langle \tau_d \rangle \langle N \rangle)$ . For FCS, the single quantities are averaged before calculating the molecular brightness, whereas BTCCD allows the calculation of the molecular brightness for every burst [11].

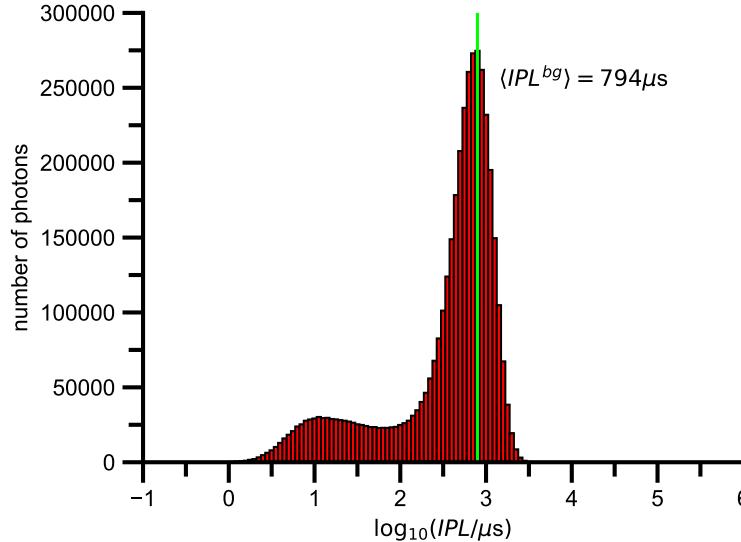


Figure 2.7: Histogram of IPL values for the red channel. The right population reflects the background. In this case, it is characterized by  $\langle IPL^{bg} \rangle = 794\mu\text{s}$  as indicated by the vertical green line. The data is taken from a measurement of dual-labeled dsDNA, see Section 4.1.

### 3 Apply Different Brightness Thresholds

Every burst defined by the burst threshold in the IPL time trace consists of a certain number of photons. Dim bursts of type I and II contain only a small number of photons. The scope of BTCCD is to exclude these bursts in order to prevent an underestimation of the coincidence fraction. To do so, the bursts are further selected by choosing only bursts with more than a minimum number of photons  $N_{\gamma,min}$ , called brightness threshold. To ensure the comparability between both channels, the brightness threshold is usually normalized on the average number of photons  $\langle N_{\gamma,0} \rangle$  in the initial bursts for each channel. In the following, the normalized brightness threshold is called  $n_{br}$  and is given by

$$n_{br} = \frac{N_{\gamma,min}}{\langle N_{\gamma,0} \rangle}. \quad (2.12)$$

To calculate the coincidence fraction for the bursts selected by  $n_{br}$  in the red channel, the total number of selected red bursts  $B_R$ , as well as the number of coincident bursts  $B_{RB}$  has to be counted. A burst in the red channel is coincident if its start or end time falls between the start and end time of a blue burst. Here, all initial blue bursts are considered since the brightness threshold for both channels is applied separately. Then, the coincidence fraction for the red channel is given by

$$f_{RB}(n_{br}) = \frac{B_{RB}(n_{br})}{B_R(n_{br})}. \quad (2.13)$$

For the blue channel, an analogous consideration yields

$$f_{BR}(n_{br}) = \frac{B_{BR}(n_{br})}{B_B(n_{br})}. \quad (2.14)$$

Since a count of bursts is a Poisson distributed number, the uncertainty on these quantities are given by

$$\sigma_{f_{RB}}(n_{br}) = f_{RB}(n_{br}) \cdot \sqrt{\frac{1}{B_{RB}(n_{br})} + \frac{1}{B_R(n_{br})}} \quad (2.15)$$

and

$$\sigma_{f_{BR}}(n_{br}) = f_{BR}(n_{br}) \cdot \sqrt{\frac{1}{B_{BR}(n_{br})} + \frac{1}{B_B(n_{br})}}. \quad (2.16)$$

An example for the calculation of the coincidence fraction as a function of  $n_{br}$  is illustrated in Figure 2.8. There, the exclusion of borderline (I) and peripheral trajectories (II) can be observed by the increase of the coincidence fractions for lower brightness thresholds. For higher brightness thresholds, the curves saturate since only central trajectories (III) remain.

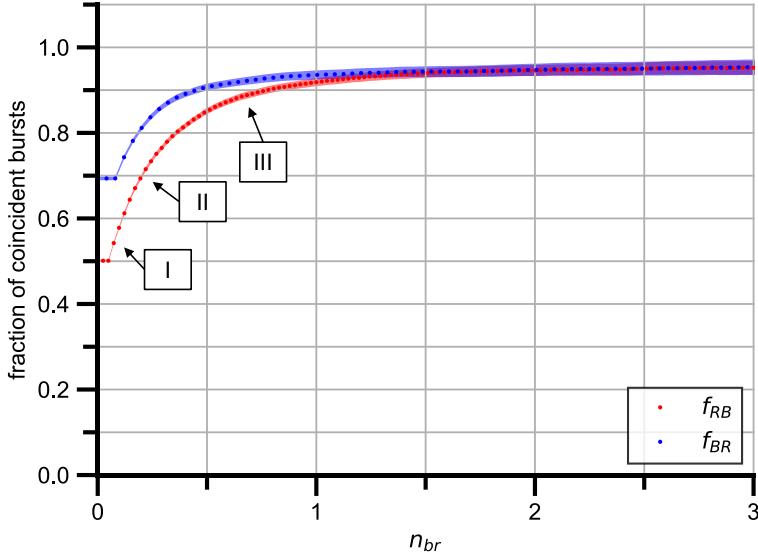


Figure 2.8: Fraction of coincident bursts for both channels as a function of  $n_{br}$ . For large values of  $n_{br}$ , the curves saturate since borderline (I) and peripheral trajectories (II) have been excluded, compare to Figures 2.5 and 2.6. The data is taken from a measurement of dual-labeled dsDNA, see Section 5.1.

The application of BTCCD effectively decreases the confocal detection volume. For a given brightness threshold, the average number of molecules inside the confocal detection volume of one channel can be calculated from the number of selected bursts  $B$ , the total measurement time  $T$ , and the dwell time  $\langle \tau_d \rangle$  according to [3]

$$\langle N \rangle(n_{br}) = -\ln \left( 1 - \frac{B(n_{br}) \cdot \langle \tau_d \rangle(n_{br})}{T} \right). \quad (2.17)$$

In Section A.2, the uncertainties on the relevant quantities  $\langle \tau_d \rangle$ ,  $\langle N \rangle$ , and  $\langle MB \rangle$  are derived.

#### 4 Find Optimal Brightness Threshold

Having determined the coincidence fraction for different burst thresholds, the question remains which coincidence fraction should be taken as the ideal one. For larger  $n_{br}$ , the uncertainty on the coincidence fraction increases significantly since the number of bursts decreases. However, for lower  $n_{br}$  not all trajectories of type I and II have been excluded. Therefore, a compromise between precision and accuracy has to be found. For the red channel, the precision is defined as the relative uncertainty:

$$\frac{\sigma_{f_{RB}}(n_{br})}{f_{RB}(n_{br})} = \sqrt{\frac{1}{B_{RB}(n_{br})} + \frac{1}{B_R(n_{br})}}. \quad (2.18)$$

The accuracy is the deviation of the coincidence fraction at a certain  $n_{br}$  from its value  $f_{RB,high}$  at the highest brightness threshold. It is given by

$$\frac{\Delta f_{RB}(n_{br})}{f_{RB}(n_{br})} = \frac{f_{RB,high} - f_{RB}(n_{br})}{f_{RB}(n_{br})}. \quad (2.19)$$

For the blue channel, similar expressions are obtained. To find the optimal brightness threshold  $n_{br,opt}$ , the intersection between precision and accuracy has to be determined. An example for the red channel can be found in Figure 2.9. The search for the optimal brightness threshold for the blue channel is illustrated in Figure A.2.

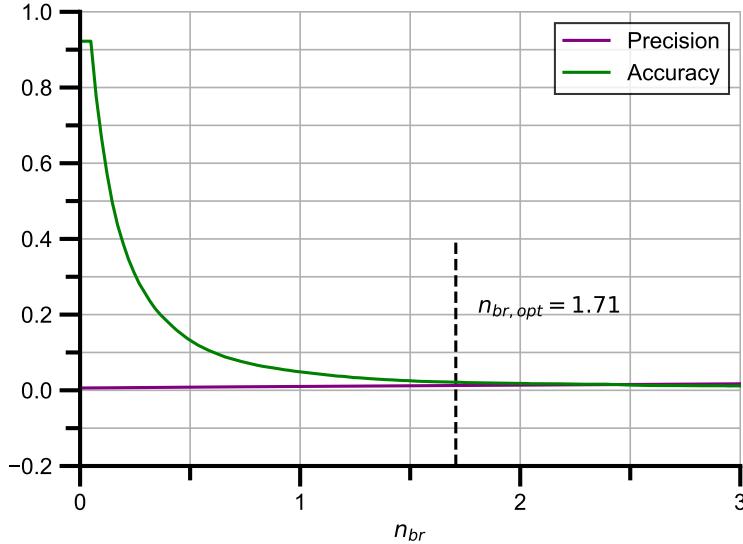


Figure 2.9: Search for the optimal brightness threshold  $n_{br,opt}$  for the red channel.  $n_{br,opt}$  is given by the intersection of precision and accuracy. The data is taken from a measurement of dual-labeled dsDNA, see Section 5.1. In this case,  $n_{br,opt} = 1.71$  leads to a coincidence fraction of  $f_{RB}(n_{br,opt}) = (94.3 \pm 1.2)\%$ .

### 2.4.2 Correction for Chance Coincidences

A coincidence will be counted by the procedure presented so far if two or more molecules with a different dye-labeling are simultaneously inside the confocal detection volume. This leads to an overestimation of the binding fraction since the coincidence may not be caused by an actual binding between different molecule types. The probability of the occurrence of such chance coincidences depends on the sample concentration. If the average number of molecules inside the confocal detection volume for each channel is lower than 0.03, the probability for more than one molecule present in the confocal detection volume is smaller than 0.1 % [3]. Therefore, chance coincidences do not need to be considered for TCCD. However, by applying BTCCD, dim bursts are excluded. For the remaining subset of bright bursts, the molecules stay for a relatively long time in the confocal detection volume. Thus, the probability that another molecule enters the confocal detection volume increases. Hence, for BTCCD, even for low sample concentrations, the effect of chance coincidences cannot be neglected.

In [3], an approach to correct for chance coincidences has been proposed. To estimate the chance coincidence fraction, the average dwell time  $\langle \tau_d \rangle$  as well as the average number of molecules inside the confocal detection volume  $\langle N \rangle$  have to be known for both channels.  $\langle \tau_d \rangle$  is directly accessible, while  $\langle N \rangle$  is given by Equation 2.17. If the probability for more than two molecules of different colors inside the confocal detection volume is neglected, the corrected coincidence fraction for the red channel can be expressed as

$$f_{RB}^{cor}(n_{br}) = \frac{f_{RB}(n_{br}) - f_{RB,0}^{chance}(n_{br})}{1 - f_{RB,0}^{chance}(n_{br})}, \quad (2.20)$$

where

$$f_{RB,0}^{chance}(n_{br}) = 1 - \exp \left\{ -\langle N_B \rangle(0) \cdot \left( \frac{\langle \tau_d^R \rangle(n_{br})}{\langle \tau_d^B \rangle(0)} + 1 \right) \right\}. \quad (2.21)$$

Both  $\langle N_B \rangle$  and  $\langle \tau_d^B \rangle$  do not depend on the brightness threshold since the procedure of BTCCD is applied separately for both channels. The corrected coincidence fraction for the blue channel is analogously given by

$$f_{BR}^{cor}(n_{br}) = \frac{f_{BR}(n_{br}) - f_{BR,0}^{chance}(n_{br})}{1 - f_{BR,0}^{chance}(n_{br})}, \quad (2.22)$$

where

$$f_{BR,0}^{chance}(n_{br}) = 1 - \exp \left\{ -\langle N_R \rangle(0) \cdot \left( \frac{\langle \tau_d^B \rangle(n_{br})}{\langle \tau_d^R \rangle(0)} + 1 \right) \right\}. \quad (2.23)$$

A detailed derivation of the uncertainties on  $f_{RB}^{cor}$  and  $f_{BR}^{cor}$  can be found in Section A.3.

### 2.4.3 Dependencies on Brightness Threshold

In this section, the dependencies of  $\langle \tau_d \rangle$ ,  $\langle N \rangle$ , and  $\langle MB \rangle$  on the brightness threshold  $n_{br}$  are derived. The obtained theoretical results are compared with experimental data in Chapter 4.

#### Dwell Time

A certain number of photons is emitted if a molecule passes through the confocal detection volume. On average, this number is proportional to the minimum number of photons  $N_{\gamma,min}$  defined by the burst threshold. Since the emission of photons by a fluorophore is a statistical process, an average time  $\Delta t'_\gamma$  between the emission of two photons can be defined. Let  $\tau_d^0$  be the average duration of a burst for  $N_{\gamma,min} = 0$ . Then, the average dwell time as a function of  $N_{\gamma,min}$  can be expressed by

$$\langle \tau_d \rangle(N_{\gamma,min}) = \Delta t'_\gamma(N_{\gamma,min}) \cdot n_{br} + \tau_d^0 \quad (2.24)$$

since, on average, every photon increases the dwell time by  $\Delta t'_\gamma$ . In general,  $\Delta t'_\gamma$  depends on  $N_{\gamma,min}$  since, for dim bursts, the time between the emission of two photons is larger than for bright bursts. However, for larger  $N_{\gamma,min}$ , most dim bursts are excluded, and only bright bursts remain. Since bright bursts correspond to central trajectories through the confocal detection volume, a further increase of  $N_{\gamma,min}$  does not change the time between the emission of two photons significantly.

The normalized brightness threshold is given by  $n_{br} = N_{\gamma,min}/\langle N_{\gamma,0} \rangle$ , where  $\langle N_{\gamma,0} \rangle$  is the average number of photons in the initial bursts. Hence, the dependency of the average dwell time on  $n_{br}$  is described by

$$\langle \tau_d \rangle(n_{br}) = \Delta t_\gamma(n_{br}) \cdot n_{br} + \tau_d^0, \quad (2.25)$$

where  $\Delta t_\gamma(n_{br}) = \Delta t'_\gamma(N_{\gamma,min}) \cdot \langle N_{\gamma,0} \rangle$ . All in all, it is expected that, for large enough  $n_{br}$ ,  $\Delta t_\gamma$  is constant, and a linear function with constant slope describes the dwell time.

#### Number of Molecules

An increase of  $n_{br}$  effectively decreases the confocal detection volume, and leads, thus, also to a decrease of  $\langle N \rangle$ . To examine the specific form of decrease, suppose that a certain increase of  $n_{br}$  decreases the confocal detection volume by  $\Delta V$ . If  $\langle N \rangle$  is large, many molecules are in the volume element  $\Delta V$ . In contrast, for a small  $\langle N \rangle$ , only a few molecules are in the exact same volume element. The number of molecules that is excluded by decreasing the confocal detection volume depends, thus, on  $\langle N \rangle$  itself. This situation can be described by the differential equation

$$\frac{d\langle N \rangle}{dn_{br}}(n_{br}) = -\frac{1}{n} \langle N \rangle(n_{br}), \quad (2.26)$$

where  $1/n$  is a proportionality constant. Its solution is given by

$$\langle N \rangle(n_{br}) = N^0 e^{-n_{br}/n}, \quad (2.27)$$

where  $N^0$  is the average number of molecules for  $n_{br} = 0$ .

### Molecular Brightness

The molecular brightness is a relative quantity since it describes the number of photons in a burst in relation to the dwell time. For large  $n_{br}$ , the number of photons in a burst, and the dwell time is proportional to  $n_{br}$ , see above. Therefore, the molecular brightness saturates at some value  $MB^{max}$ . The molecular brightness  $MB^0$  for  $n_{br} = 0$  is smaller than  $MB^{max}$  since dim bursts decrease the molecular brightness. Furthermore,  $\langle MB \rangle$  should be a smooth function that connects  $MB^0$  with  $MB^{max}$  since there is no particular  $n_{br}$  for which suddenly all dim bursts are excluded. This means that the slope of  $\langle MB \rangle$  depends on the difference between  $MB^{max}$  and the current value of  $\langle MB \rangle$  because the derivative of  $\langle MB \rangle$  has to approach 0. Thus,  $\langle MB \rangle$  is described by the non-homogeneous differential equation

$$\frac{d\langle MB \rangle}{dn_{br}}(n_{br}) = \frac{1}{n} \cdot (MB^{max} - \langle MB \rangle(n_{br})), \quad (2.28)$$

where  $1/n$  is a proportionality constant. Its solution is given by

$$\langle MB \rangle(n_{br}) = MB^0 + \Delta MB(1 - e^{-n_{br}/n}), \quad (2.29)$$

where  $MB^{max} - MB^0$ .

# 3 Materials and Instrumentation

## 3.1 Samples

### 3.1.1 Fluorescent Dyes

The following fluorescent dyes were used:

- Alexa Fluor™ 488 NHS-Ester (Thermo Fisher Scientific Inc., Waltham, USA)
- Cy5® NHS-Ester (Thermo Fisher Scientific Inc., Waltham, USA)
- Atto 647N NHS-Ester (ATTO-TEC GmbH, Siegen, Germany)

For long storage times, the dyes were lyophilized and stored at –20 °C. For timely usage, they were dissolved in PBS buffer and stored at room temperature.

### 3.1.2 DNA

The base sequence of the dye-labeled single-stranded DNA (ssDNA) was:

5'-GGA CTA GTC TAG GCG AAC GTT TAA GGC GAT CTC TGT TTA CAA CTC CGA-3'.

The following modifications were used:

- dual-labeled: 5'-modified with Alexa 488, 3'-modified with Atto 647N
- single-labeled: 3'-modified with Atto 647N

To synthesize the double-stranded DNA (dsDNA) samples, the dye-labeled ssDNA was hybridized with the complementary unlabeled ssDNA. All DNA strands were acquired from IBA Lifesciences (Göttingen, Germany). For long storage times, the DNA samples were dissolved in TE buffer and stored at –20 °C. For timely usage, they were dissolved in DNA buffer and stored at 4 °C.

### 3.1.3 Ribosomes

Ribosomes from *E. coli* unspecifically labeled with Cy5 were used. Detailed information on the preparation of the ribosome samples can be found in [17]. For timely usage, the ribosomes were dissolved in DNA buffer and stored at 4 °C.

## 3.2 Buffers

The chemical compositions of the buffers were the following:

### DNA buffer (pH of 7.5)

- 20 mmol/L tris(hydroxymethyl)aminomethane (TRIS, C<sub>4</sub>H<sub>11</sub>NO<sub>3</sub>)
- 100 mmol/L sodium chloride (NaCl)
- 10 mmol/L magnesium chloride (MgCl<sub>2</sub>)

**PBS buffer (pH of 7.2)**

- 14 mmol/L monopotassium phosphate ( $\text{KH}_2\text{PO}_4$ )
- 36 mmol/L dipotassium phosphate ( $\text{K}_2\text{HPO}_4$ )
- 150 mmol/L sodium chloride ( $\text{NaCl}$ )

**TE buffer (pH of 7.5)**

- 10 mmol/L tris(hydroxymethyl)aminomethane (TRIS,  $\text{C}_4\text{H}_{11}\text{NO}_3$ )
- 1 mmol/L ethylenediaminetetraacetic acid (EDTA,  $\text{C}_{10}\text{H}_{16}\text{N}_2\text{O}_8$ )

Prior to use, the fluorescent background of the buffers was measured. If lower concentrations of the samples were needed, they were diluted with DNA buffer. For measurements with free dyes and DNA samples, 5 vol% of Tween 20 (Sigma-Aldrich, St. Louis, USA) was added to increase the solubility.

### 3.3 Microscope Slides

Microscope slides of size 22 mm × 22 mm with a thickness of  $(170 \pm 5)\text{ }\mu\text{m}$  (“No. 1.5H”, Paul Marienfeld GmbH & Co. KG, Lauda-Königshofen, Germany) were used. For experiments with DNA and ribosomes, the microscope slides were prepared additionally to prevent the sticking of the samples to the surface of the microscope slide. The preparation procedure of the microscope slides for those samples consisted of the following steps:

1. Rinse for a few seconds with isopropanol, then acetone, and finally distilled water.
2. Dry with nitrogen.
3. Perform plasma cleaning (“Zepto”, Diener electronic GmbH + Co. KG, Ebhausen, Germany) for at least 20 minutes.

To prevent the evaporation of the sample during a measurement, the microscope slides were sealed with a paraffin foil.

### 3.4 Confocal Fluorescence Microscope

For all experiments, the confocal fluorescence microscope “MicroTime 200” (PicoQuant GmbH, Berlin, Germany) was used. The following list gives an overview of its main parts:

**objective**

- “UPLSAPO60XW” (Olympus Deutschland GmbH, Hamburg, Germany), magnification: 60x, numerical aperture: 1.2, water immersion, correction collar set to 0.17

**lasers**

- 481 nm: “LDH-D-C-485” (PicoQuant GmbH, Berlin, Germany)
- 633 nm: “LDH-D-C-640” (PicoQuant GmbH, Berlin, Germany)

**detectors**

- channel 1: “ $\tau$ -SPAD” (PicoQuant GmbH, Berlin, Germany)
- channel 2: “COUNT-T100” (Laser Components GmbH, Olching, Germany)

**emission bandpass filters**

- blue emission filter: “FF01-530/55-25” (Semrock Inc., Rochester, USA)
- red emission filter: “ET685/80m” (Chroma Technology Corp., Rockingham, USA)

## 3.5 Software

The following list gives an overview of the software that was used for this thesis. For every software, the main purpose, the name, and the used version is given:

- data acquisition with confocal microscope, and FCS analysis: “SymPhoTime 64”, version 2.4 (PicoQuant GmbH, Berlin, Germany)
- BTCCD analysis: “BTCCD”, written by Henning Höfig (RWTH Aachen, Germany) in “MATLAB”, version R2015b [18]
- self-written code for data analysis: “MATLAB”, version R2015b (The MathWorks Inc., Natick, USA)
- self-written code for data analysis, and visualization: “Python”, version 3.7.7 (Python Software Foundation, Wilmington, USA)
- Python distribution: “Anaconda Software Distribution”, version 2020.02 [19]

# 4 Dependencies on Brightness Threshold

The scope of the following measurement with dual-labeled dsDNA is the analysis of the dependency of  $\langle \tau_d \rangle$ ,  $\langle N \rangle$ , and  $\langle MB \rangle$  on the brightness threshold  $n_{br}$ . The theoretically derived mathematical expressions from Section 2.4.3 are compared with the experimental findings.

## 4.1 Measurement

A measurement of 40 min with dual-labeled dsDNA was conducted. The recorded macro time trace was transformed into an IPL time trace, and BTCCD was applied according to the procedure described in Section 2.4.1. Information on the burst threshold, the background, and on the starting parameters can be found in Table 4.1. Since bursts are defined using the smoothed IPL time trace, a burst may consist of only zero or one detected photon. For those bursts, no physically meaningful dwell time exists. To characterize the measurement only with stable bursts, the given values for  $\langle \tau_d \rangle$ ,  $\langle N \rangle$ , and  $\langle MB \rangle$  in the table are calculated from all bursts that contain at least three photons. This convention is used for the rest of the thesis, every time general information on a measurement is given.

ch.	$IPL^{thr}$ [μs]	$\langle IPL^{bg} \rangle$ [μs]	$\langle N \rangle [10^{-3}]$	$\langle \tau_d \rangle$ [μs]	$\langle MB \rangle$ [kHz]
red	150	794	$10.19 \pm 0.10$	$1123.6 \pm 8.1$	$29.8 \pm 4.1$
blue	180	1995	$2.469 \pm 0.040$	$844.7 \pm 8.9$	$23.1 \pm 7.8$

Table 4.1: Burst threshold, background, and starting parameters for a measurement of 40 min with dual-labeled dsDNA.

## 4.2 Results

### 4.2.1 Dwell Time

A theoretical consideration came to the result that the dwell time  $\langle \tau_d \rangle$  depends linearly on the brightness threshold  $n_{br}$ . Thus, the starting point of the analysis of the dwell time is a linear function

$$\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta t_\gamma \cdot n_{br}, \quad (4.1)$$

where  $\Delta t_\gamma$  and  $\tau_d^0$  are fit parameters. The red dots in the upper graph of Figure 4.1 show the raw data for the red channel as well as a linear fit, illustrated by a green line. Underneath, in the lower graph, the corresponding residuals are plotted. As expected, the part of the raw data for smaller  $n_{br}$  is not linear. Thus, the vertical dashed line indicates the interval that was used for fitting. For higher  $n_{br}$ , the linear function reflects the general increasing trend of the dwell time. Nonetheless, the residual plot reveals a structured deviation of the fit from the raw data.

In the figure, the values of the fit parameters are given. Additionally, the statistical measures

$\chi^2/n_{dof}$  and  $p$  are stated. In general,  $\chi^2$  is defined as

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - f(x_i))^2}{\sigma_{y_i}^2}, \quad (4.2)$$

where  $n$  is the number of raw data points  $(x_i, y_i)$  with uncertainty  $\sigma_{y_i}$ , and  $f$  is the fitted model function.  $\chi^2$  quantifies the deviation of the fit from the raw data in relation to the uncertainty. The value of  $\chi^2$  follows the  $\chi^2$  distribution that has the expectation value  $n_{dof}$ , where  $n_{dof}$  is the difference between the number of measurement points  $n$  and the parameters of the fit model. The normalization of  $\chi^2$  on  $n_{dof}$  allows an interpretation of the calculated value, and an evaluation of the fit quality.  $p$  is the probability to obtain a value of  $\chi^2$  that is larger than the calculated one.

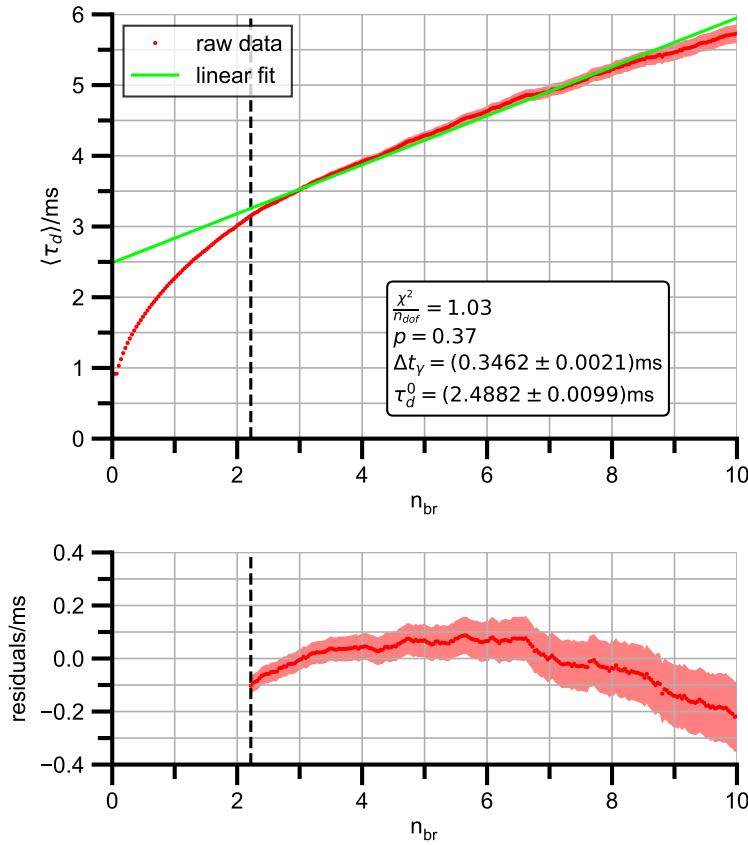


Figure 4.1: Linear fit according to  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta t_y \cdot n_{br}$  for the red channel. The vertical dashed line indicates the range of  $n_{br}$  that was used for fitting. For smaller  $n_{br}$ , the dependency is not linear. For higher  $n_{br}$ , the residual plot reveals a structure in the deviation of the fitted model from the raw data.

The linear fit of the dwell time for the blue channel can be found in Figure B.1. Table 4.2 gives an overview of the fit parameters of the linear model for both channels.

ch.	$\tau_d^0$ [μs]	$\Delta t_\gamma$ [μs]
red	$2488.2 \pm 9.9$	$346.2 \pm 2.1$
blue	$1221 \pm 17$	$381.7 \pm 4.4$

Table 4.2: Parameters of the fit of  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta t_\gamma \cdot n_{br}$  for both channels.

The part for smaller  $n_{br}$  and the deviation revealed by the residual plot motivate a biexponential approach

$$\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta\tau_1(1 - e^{-n_{br}/n_1}) + \Delta\tau_2(1 - e^{-n_{br}/n_2}), \quad (4.3)$$

where  $\tau_d^0$ ,  $\Delta\tau_1$ ,  $n_1$ ,  $\Delta\tau_2$ , and  $n_2$  are fit parameters. The fit of this model for the red channel can be found in Figure 4.2. The biexponential function takes two different saturation processes into account that describe the dependency for the complete range of  $n_{br}$  accurately. The fit for the blue channel is illustrated in Figure B.2.

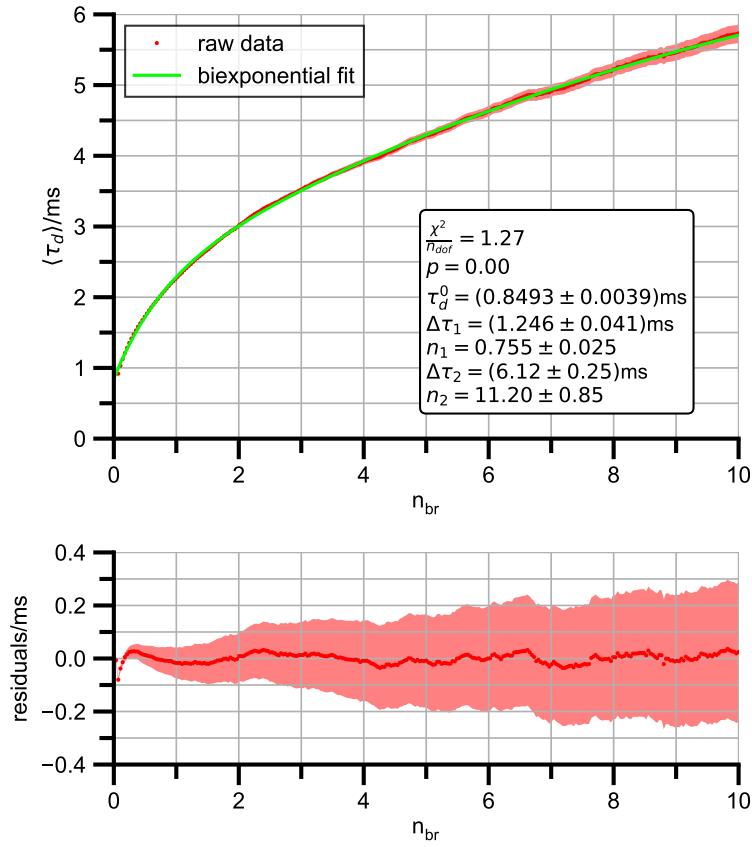


Figure 4.2: Biexponential fit according to  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta\tau_1(1 - e^{-n_{br}/n_1}) + \Delta\tau_2(1 - e^{-n_{br}/n_2})$  for the red channel. The biexponential function takes two different saturation processes into account. Thus, the residual plot does not show a structure anymore.

In Table 4.3, the fit parameters for both channels are stated. They reveal that the exponential term modulated by  $\Delta\tau_1$  and  $n_1$  describes the initial saturation for smaller  $n_{br}$ , wherein a small range of  $n_1$  the initial dwell time  $\tau_d^0$  is increased by  $\Delta\tau_1$ . The second exponential term modulates an increase of the dwell time over a larger range of  $n_2$ . For the red channel, both exponential terms describe a saturation with notable effects in the relevant range of  $n_{br}$  from 0 to 10. For the blue channel, the second exponential saturation takes place over an extremely large range of  $n_2 = (9.7 \pm 1.1) \times 10^6$ . Here, in fact, the second term describes a linear increase. Mathematically, this becomes obvious by rewriting it as

$$\Delta\tau_2(1 - e^{-n_{br}/n_2}) \approx \Delta\tau_2(1 - (1 + n_{br}/n_2)) = \frac{\Delta\tau_2}{n_2} n_{br}, \quad (4.4)$$

where  $e^x \approx 1 + x$  for  $x \ll 1$  was used. The slope is given by

$$m_2 = \frac{\Delta\tau_2}{n_2} = (282 \pm 49) \mu\text{s}, \quad (4.5)$$

where the uncertainty was propagated according to

$$\sigma_{m_2} = m_2 \cdot \sqrt{\left(\frac{\sigma_{\Delta\tau_2}}{\Delta\tau_2}\right)^2 + \left(\frac{\sigma_{n_2}}{n_2}\right)^2}. \quad (4.6)$$

$m_2$  can be compared with the results from the linear fit for the blue channel, where the slope was given by  $\Delta t_\gamma = (381.7 \pm 4.4) \mu\text{s}$ , see Table 4.2. Both values are compatible with regard to their uncertainties:

$$\frac{\Delta t_\gamma - m_2}{\sqrt{\sigma_{m_2}^2 + \sigma_{\Delta t_\gamma}^2}} \approx 2.04. \quad (4.7)$$

The difference in the absolute values is a consequence of the different approaches. For the linear fit, the range of  $n_{br}$  that was supposed to be linear was manually specified. By applying a biexponential fit, the range is mathematically optimized. Thus, the structure in the residuals, for the blue channel, reduces, compare Figures B.1 and B.2.

ch.	$\tau_d^0$ [μs]	$\Delta\tau_1$ [μs]	$n_1$	$\Delta\tau_2$ [μs]	$n_2$
red	$849.3 \pm 3.9$	$1246 \pm 41$	$0.755 \pm 0.025$	$6120 \pm 250$	$11.20 \pm 0.85$
blue	$602.9 \pm 4.8$	$1350 \pm 140$	$2.47 \pm 0.22$	$(2.74 \pm 0.35) \times 10^9$	$(9.7 \pm 1.1) \times 10^6$

Table 4.3: Parameters of the fit of  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta\tau_1(1 - e^{-n_{br}/n_1}) + \Delta\tau_2(1 - e^{-n_{br}/n_2})$  for both channels.

To sum it up, the theoretically derived linear expression allows a description of the increasing trend of  $\langle \tau_d \rangle$  for larger values of  $n_{br}$ . However, a biexponential approach modulates the development of the dwell time over the complete range of  $n_{br}$  accurately. The two exponential terms reveal two underlying processes. The first exponential saturation on a small range of  $n_1$  can be explained by the exclusion of dim bursts, as already discussed in the theoretical derivation. In contrast, the second saturation on a larger scale  $n_2$  does not necessarily take place in the relevant range of  $n_{br}$ , as observed for the blue channel. Its specific source is unknown.

### 4.2.2 Number of Molecules

The result of the theoretical derivation was that the average number of molecules inside the confocal detection volume  $\langle N \rangle$  can be described by an exponential decay

$$\langle N \rangle (n_{br}) = N^0 e^{-n_{br}/n}, \quad (4.8)$$

where  $N^0$  and  $n$  are fit parameters. Figure 4.3 shows the raw data and a fit of this model for the red channel. The fit reflects the general decreasing trend of  $\langle N \rangle$ , but deviates significantly from its specific behavior. A similar effect can be observed for the blue channel, see Figure B.3. An overview of the fit parameters for both channels is given in Table 4.4.

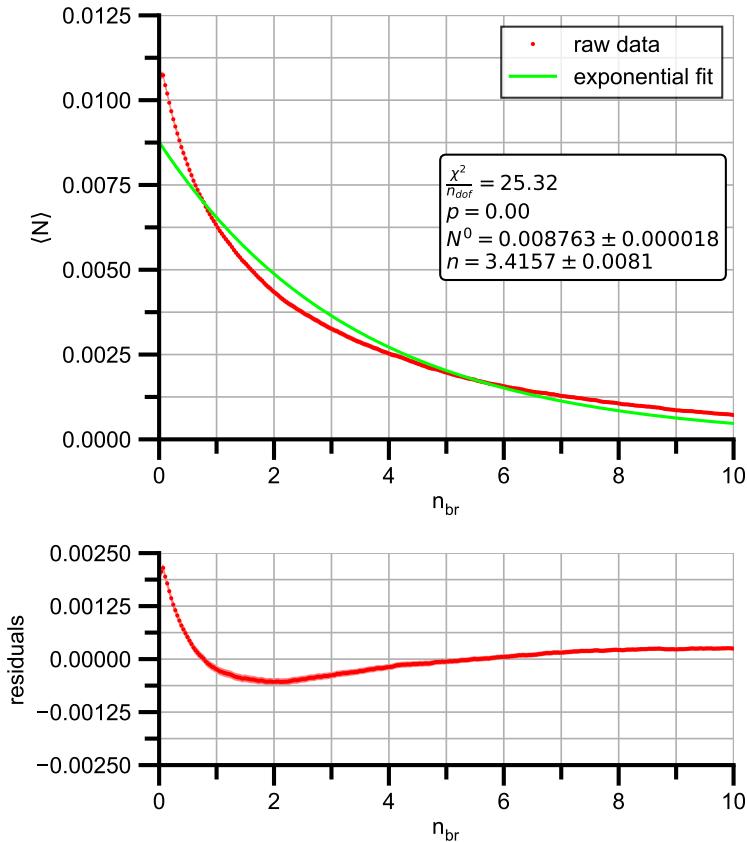


Figure 4.3: Exponential fit according to  $\langle N \rangle (n_{br}) = N^0 e^{-n_{br}/n}$  for the red channel. The fit reflects the general decreasing trend of  $\langle N \rangle$ , but deviates clearly from the specific behavior.

ch.	$N^0 [10^{-3}]$	$n$
red	$8.763 \pm 0.018$	$3.4157 \pm 0.0081$
blue	$2.680 \pm 0.013$	$2.178 \pm 0.011$

Table 4.4: Parameters of the fit of  $\langle N \rangle (n_{br}) = N^0 e^{-n_{br}/n}$  for both channels.

An improved model has to take the deviation between the theoretically derived expression and the raw data into account. The average number of molecules inside the confocal detection volume is directly calculated from the dwell time, see Equation (2.17). Thus, inspired by the two processes observed for the dwell time above, a biexponential decay

$$\langle N \rangle(n_{br}) = ae^{-n_{br}/n_1} + be^{-n_{br}/n_2} \quad (4.9)$$

is proposed. Here,  $a$ ,  $n_1$ ,  $b$ , and  $n_2$  are the fit parameters. A fit of this function for the red channel can be found in Figure 4.4. The biexponential model describes the data more accurately than the monoexponential function. This is also reflected in the residual plot. The results for the blue channel show the same improvement and can be found in Figure B.4. Table 4.5 contains the fit parameters for both the blue and the red channel. The parameters reveal again that the first exponential decay modulates a decrease of  $\langle N \rangle$  by  $a$  over a relative short range  $n_1$ . The second exponential term describes a decay by  $b$  over a larger range  $n_2$ . In contrast to the dwell time in the previous section, here, the second exponential decay with reasonable values for  $b$  and  $n_2$  can be observed in both channels.

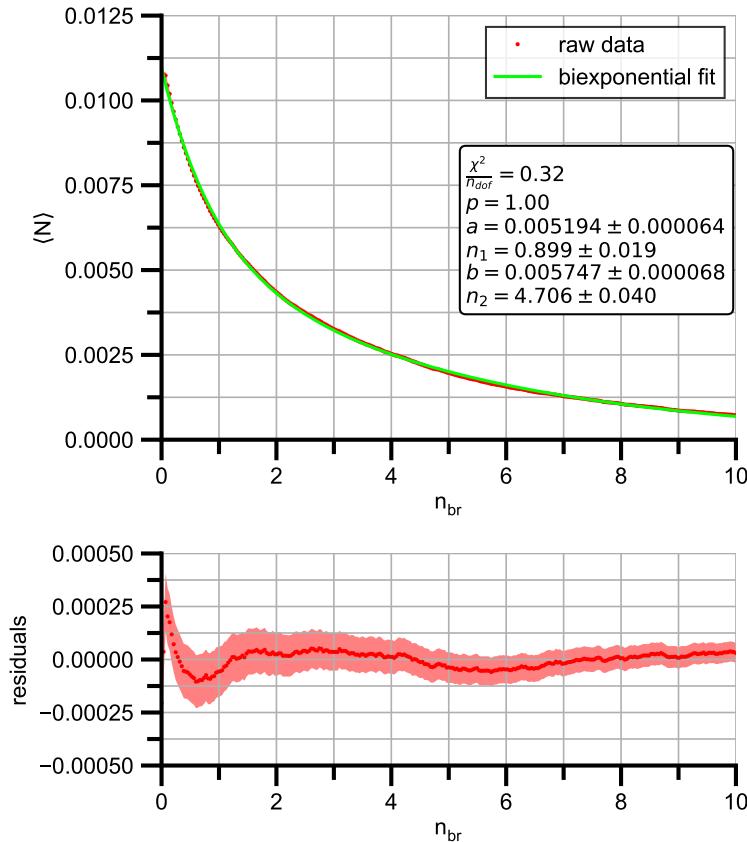


Figure 4.4: Biexponential fit according to  $\langle N \rangle(n_{br}) = ae^{-n_{br}/n_1} + be^{-n_{br}/n_2}$  for the red channel. The two decay processes of the biexponential function modulate the measurement more accurately than the monoexponential model.

ch.	$a [10^{-3}]$	$n_1$	$b [10^{-3}]$	$n_2$
red	$5.194 \pm 0.064$	$0.899 \pm 0.019$	$5.747 \pm 0.068$	$4.706 \pm 0.040$
blue	$1.38 \pm 0.15$	$1.163 \pm 0.094$	$1.53 \pm 0.16$	$2.82 \pm 0.11$

Table 4.5: Parameters of the fit of  $\langle N \rangle (n_{br}) = ae^{-n_{br}/n_1} + be^{-n_{br}/n_2}$  for both channels.

In conclusion, a biexponential decay describes the average number of molecules inside the confocal detection volume  $\langle N \rangle$  accurately. The analysis yields again two processes that lead to a decrease of  $\langle N \rangle$ . In contrast to the results for the dwell time above, they are observable for both channels.

#### 4.2.3 Molecular Brightness

The theoretical consideration obtained that an exponential saturation

$$\langle MB \rangle (n_{br}) = MB^0 + \Delta MB(1 - e^{-n_{br}/n}), \quad (4.10)$$

where  $MB^0$ ,  $\Delta MB$ , and  $n$  are fit parameters, describes the molecular brightness. The raw data and a fit for the red channel can be found in Figure 4.5. Over large parts, the monoexponential model describes the raw data correctly. However, for  $n_{br} < 1$ , it underestimates the increasing trend of the raw data. The fit for the blue channel can be found in Figure B.5. For the blue channel, such a deviation is not observed. The low  $\chi^2/n_{dof}$  for both channels is a result of the relatively large uncertainty on the molecular brightness. Since  $\langle MB \rangle$  is an average value calculated from all bursts, it is a sensitive quantity. The extreme uncertainty for  $n_{br} \rightarrow 0$  comes from bursts with only a tiny number of photons. This is another reason, why, for the characterization of the measurement above, a minimal number of at least three photons was required. The fit parameters for both channels are stated in Table 4.6.

Despite the broad agreement of the monoexponential fit with the raw data, a biexponential model

$$\langle MB \rangle (n_{br}) = MB^0 + \Delta MB_1(1 - e^{-n_{br}/n_1}) + \Delta MB_2(1 - e^{-n_{br}/n_2}), \quad (4.11)$$

where  $MB^0$ ,  $\Delta MB_1$ ,  $n_1$ ,  $\Delta MB_2$ , and  $n_2$  are the parameters, is fitted. Hereby, it is investigated if the two processes that were observed for the dwell time and the number of molecules are also present in the molecular brightness. Figure 4.6 illustrates the fit for the red channel. The fit for the blue channel is depicted in Figure B.6. As expected, the value of  $\chi^2/n_{dof}$  becomes even lower. However, for the red channel, the biexponential fit modulates the part of the raw data for  $n_{br} < 1$  more precisely than the monoexponential saturation. The fit parameters reveal two distinct processes described by the two exponential functions, as can be seen in Table 4.7. For the blue channel, the parameters of the second exponential saturation with regard to their large uncertainties are in agreement with zero. Therefore, the molecular brightness is effectively described by a monoexponential saturation. The situation for the blue channel is similar to the results for the dwell time, where the second process is not as much pronounced as in the red channel.

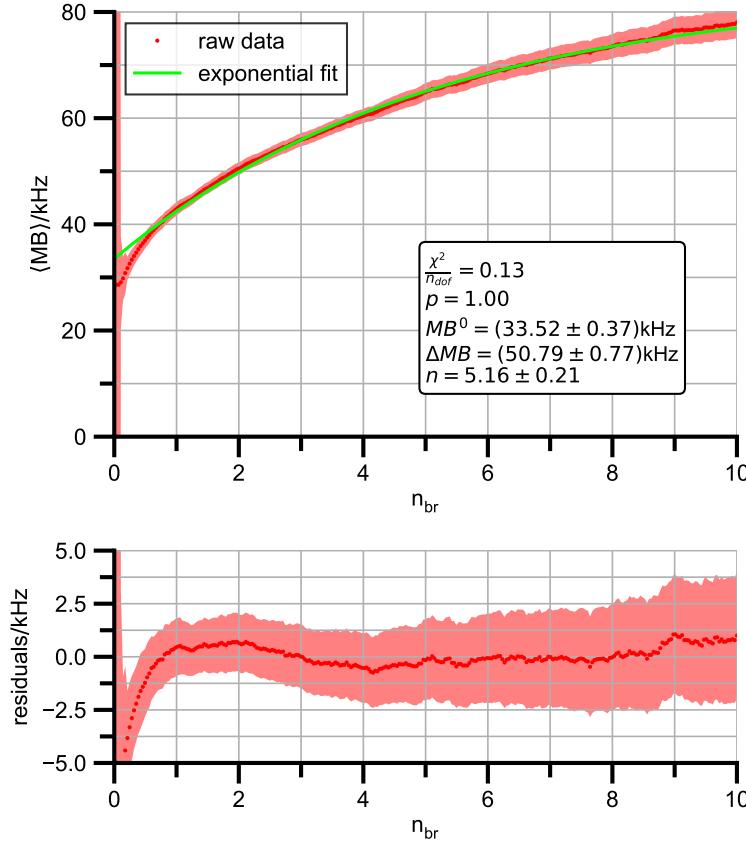


Figure 4.5: Exponential fit according to  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB(1 - e^{-n_{br}/n})$  for the red channel. For  $n_{br} < 1$ , the model slightly underestimates the increasing trend of the raw data. The comparably large uncertainty on  $\langle MB \rangle$  explains the low value for  $\chi^2/n_{dof}$  for the fits.

ch.	$MB^0$ [kHz]	$\Delta MB$ [kHz]	$n$
red	$33.52 \pm 0.37$	$50.79 \pm 0.77$	5.16
blue	$23.39 \pm 0.63$	$12.8 \pm 2.3$	$7.53 \pm 2.87$

Table 4.6: Parameters of the fit of  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB(1 - e^{-n_{br}/n})$  for both channels.

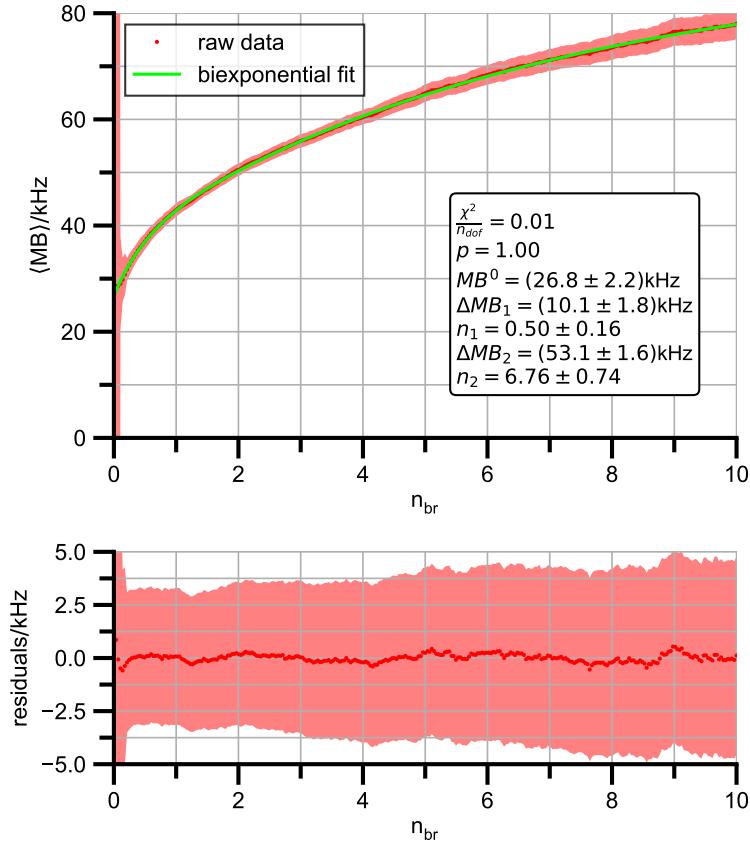


Figure 4.6: Fit according to  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB_1(1 - e^{-n_{br}/n_1}) + \Delta MB_2(1 - e^{-n_{br}/n_2})$  for the red channel. Information on the measurement can be found in Section 4.1. The bi-exponential model describes the part of the raw data for  $n_{br} < 1$  with a higher slope more precisely than the monoexponential approach.

ch.	$MB^0$ [kHz]	$\Delta MB_1$ [kHz]	$n_1$	$\Delta MB_2$ [kHz]	$n_2$
red	$26.8 \pm 2.2$	$10.1 \pm 1.8$	$0.50 \pm 0.16$	$53.1 \pm 1.6$	$6.76 \pm 0.74$
blue	$19.7 \pm 3.6$	$6.0 \pm 3.2$	$0.84 \pm 0.50$	$(-1.3 \pm 2.1) \times 10^7$	$(-1.7 \pm 2.7) \times 10^7$

Table 4.7: Parameters of a fit of  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB_1(1 - e^{-n_{br}/n_1}) + \Delta MB_2(1 - e^{-n_{br}/n_2})$  for both channels.

Overall, for the red channel, two processes are observed. Hence, a biexponential saturation describes the raw data more accurately than a monoexponential model. For the blue channel, the second process is not very pronounced. The uncertainty on  $\langle MB \rangle$  complicates the specification of the raw data for higher values of  $n_{br}$ .

### 4.3 Two Underlying Processes

In the previous section, it was shown phenomenologically that the dependency of  $\langle \tau_d \rangle$ ,  $\langle N \rangle$ , and  $\langle MB \rangle$  on the brightness threshold  $n_{br}$  can be modulated by two distinct exponential terms. For  $\langle \tau_d \rangle$  and  $\langle MB \rangle$ , a biexponential saturation takes place, while  $\langle N \rangle$  is described by a biexponential decay. The two exponential terms of the form  $e^{-n_{br}/n_1}$  and  $e^{-n_{br}/n_2}$  represent two processes that have an influence on the BTCCD analysis. Table 4.8 gives an overview of the ranges  $n_1$  and  $n_2$  for all analyzed quantities. The first exponential term describes a process over a small range of  $n_{br}$ , while the second process takes place over a larger range. The specific range varies for every quantity, and, for the blue channel, the second process is less pronounced.

ch.	quantity	$n_1$	$n_2$
red	$\langle \tau_d \rangle$	$0.755 \pm 0.025$	$11.20 \pm 0.85$
	$\langle N \rangle$	$0.899 \pm 0.019$	$4.706 \pm 0.040$
	$\langle MB \rangle$	$0.50 \pm 0.16$	$6.76 \pm 0.74$
blue	$\langle \tau_d \rangle$	$2.47 \pm 0.22$	$(9.7 \pm 1.1) \times 10^6$
	$\langle N \rangle$	$1.163 \pm 0.094$	$2.82 \pm 0.11$
	$\langle MB \rangle$	$0.84 \pm 0.50$	$(-1.7 \pm 2.7) \times 10^7$

Table 4.8: Overview of the ranges of the exponential terms of the form  $e^{-n_{br}/n_1}$  and  $e^{-n_{br}/n_2}$  that describe the analyzed quantities.

The first process is the exclusion of dim bursts connected with peripheral and borderline trajectories.

The second process allows the conclusion that another phenomenon in the order of the dwell time takes place during the measurement. This process changes the fluorescence characteristic. Transitions to triplet states, see Section 2.1, increase the average time that an excited electron takes to relax into the ground state and to emit a photon. The occurrence of triplet state transitions depends on the chemical structure of the dye and the intensity of the incoming laser light [2]. Thus, the variation of the intensity is a way to investigate the origin of the second term further. Another aspect is the negative charge of DNA. In combination with the overall positive charge of Atto 647N and the negative charge of Alexa 488, Coulomb interactions could influence the fluorescence emission of the dyes [20].

There is a variety of other processes that change the emission characteristics of a fluorophore. Hence, the specific nature of the second process is an object of further research [21].

# 5 Optimal Number of Bursts

The coincidence fraction  $f$  at the optimal brightness threshold  $n_{br,opt}$  is a measure for the binding fraction of the investigated sample. Its uncertainty  $\sigma_f$  depends on the recorded number of bursts. For BTCCD, an uncertainty lower than 5 % is achievable [17]. Thus, the scope of the following measurement is to determine the required number of bursts to obtain an uncertainty lower than 5 %.

## 5.1 Measurement

A measurement of 40 min with dual-labeled dsDNA was conducted by Olessya Yukhnovets (RWTH Aachen, Germany). Table 5.1 includes information on the burst threshold, the background, and on the starting parameters.

ch.	$IPL^{thr}$ [\mu s]	$\langle IPL^{bg} \rangle$ [\mu s]	$\langle N \rangle [10^{-3}]$	$\langle \tau_d \rangle$ [\mu s]	$\langle MB \rangle$ [kHz]
red	150	1000	$44.46 \pm 0.33$	$1565.3 \pm 9.6$	$27.8 \pm 2.5$
blue	180	891	$27.80 \pm 0.18$	$1332.9 \pm 5.8$	$22.7 \pm 2.6$

Table 5.1: Burst threshold, background, and starting parameters for a measurement of 40 min with dual-labeled dsDNA.

## 5.2 Results and Discussion

For the red channel, the coincidence fraction  $f_{RB}$  at the optimal brightness threshold is given by

$$f_{RB}(n_{br,opt}) = \frac{B_{RB}(n_{br,opt})}{B_R(n_{br,opt})}, \quad (5.1)$$

where  $B_{RB}$  and  $B_R$  denote the number of coincidence bursts, and the total number of selected bursts, respectively. According to Equation (2.15), the uncertainty  $\sigma_{f_{RB}}$  is

$$\sigma_{f_{RB}}(n_{br,opt}) = f_{RB}(n_{br,opt}) \sqrt{\frac{1}{B_R(n_{br,opt})} + \frac{1}{B_{RB}(n_{br,opt})}}. \quad (5.2)$$

Substituting  $B_{RB}(n_{br,opt}) = f_{RB}(n_{br,opt})B_R(n_{br,opt})$ , yields

$$\sigma_{f_{RB}}(n_{br,opt}) = \sqrt{(f_{RB}(n_{br,opt}))^2 + f_{RB}(n_{br,opt})} \cdot \frac{1}{\sqrt{B_R(n_{br,opt})}}. \quad (5.3)$$

Thus, the uncertainty on  $f_{RB}$  depends on the number of selected bursts. The pre-factor is determined by the coincidence fraction at the optimal brightness threshold, and takes values between 0 and  $\sqrt{2}$ . For the blue channel, the same equation is valid by substituting  $RB \rightarrow BR$ .

For a measurement of solely dual-labeled dsDNA, a coincidence fraction of 100 % is expected for both channels. Therefore, the pre-factor takes its highest possible value of  $\sqrt{2}$ . If an uncertainty of 5 % is desired, the required minimum number of selected bursts for the red channel can be calculated by rearranging Equation (5.3)

$$B_R(n_{br,opt}) = \frac{2}{(\sigma_{f_{RB}}(n_{br,opt}))^2} = \frac{2}{0.05^2} = 800. \quad (5.4)$$

For the blue channel, the same number is obtained.

Figure 5.1 illustrates the dependency of the uncertainty on the selected bursts. The data points were obtained by splitting the IPL time trace into several parts with increasing lengths. Each part contains an increasing number of bursts because of the corresponding increasing measurement time. For each part, the coincidence fraction and its uncertainty at the optimal brightness threshold is determined. The data follows Equation (5.3). For approximately  $B_R(n_{br,opt}) = 800$ , the uncertainty falls below 5 %. For the blue channel, a similar behavior is observed, see Figure C.1. There, the uncertainty drops below 5 % slightly before  $B_B(n_{br,opt}) = 800$  because the experimentally determined coincidence fraction is not exactly 100 % but slightly lower. Thus, the pre-factor of Equation (5.3) is a bit smaller.

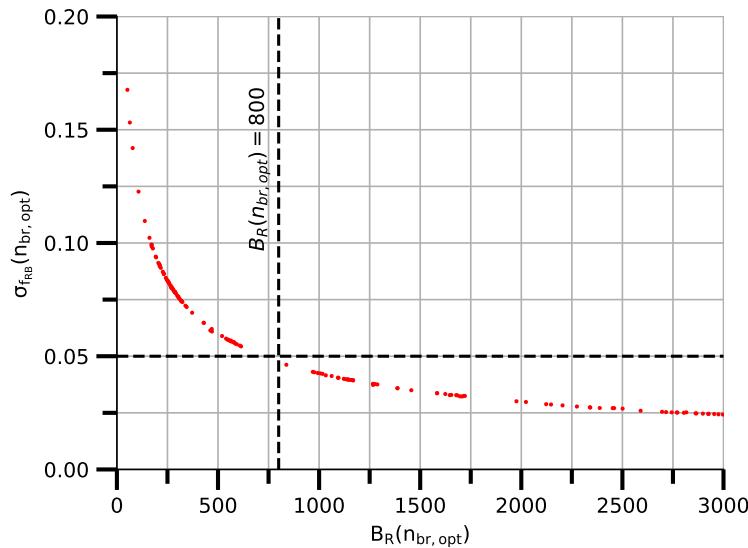


Figure 5.1: Uncertainty on the coincidence fraction  $f_{RB}$  at the optimal brightness threshold  $n_{br,opt}$  as a function of the selected number of bursts  $B_R(n_{br,opt})$  for the red channel. The optimal number of selected bursts, for which the uncertainty falls below 5 %, is approximately  $B_R(n_{br,opt}) = 800$ .

The pre-factor and thus the uncertainty take their largest values for a coincidence fraction of 100 %. If the expected coincidence fraction for a measurement is lower, the pre-factor can be adjusted. Nonetheless, it has to be paid attention to chance coincidences that increase the coincidence fraction  $f_{RB}$ . The effect of chance coincidences on the coincidence fraction is further discussed in Chapter 6.

Another way to illustrate the uncertainty on the coincidence fraction at the optimal brightness threshold is a plot against the initial number of bursts  $B_R(0)$ . The advantage is that this number is known before the application of BTCCD, and can be used to estimate the required measurement time. Figure 5.2 illustrates such a plot for the red channel. The initial number of bursts  $B_R(0)$  does not determine the uncertainty completely because  $\sigma_{f_{RB}}(n_{br,opt})$  depends only on the selected number of bursts  $B_R(n_{br,opt})$ .  $B_R(0)$  and  $B_R(n_{br,opt})$  are correlated as more initial bursts include most likely more selected bursts, but there is no deterministic relationship. Nonetheless, it can be seen that, for around 10 000 initial bursts, the uncertainty is most probably lower than 5 %. For the blue channel, even around 7500 initial bursts are sufficient, see Figure C.2.

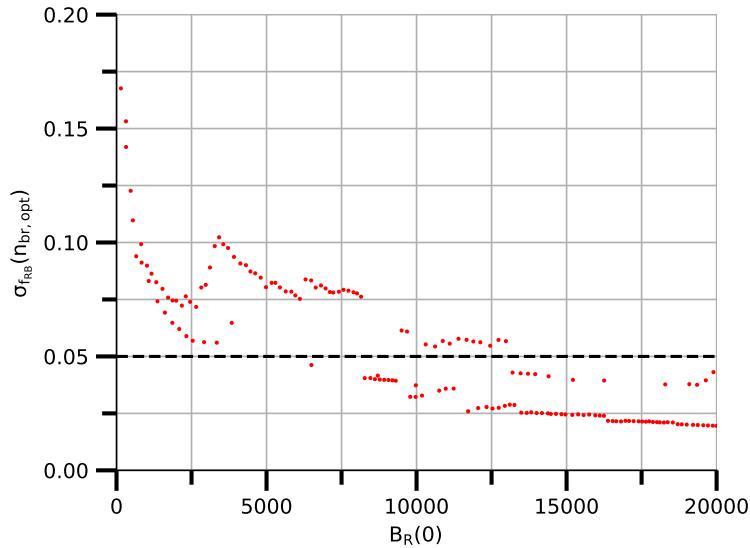


Figure 5.2: Uncertainty on the coincidence fraction  $f_{RB}$  at the optimal brightness threshold  $n_{br,opt}$  as a function of the initial number of bursts  $B_R(0)$  for the red channel. The graph does not reveal a completely deterministic behavior. However, for a number of about 10 000 initial bursts, the uncertainty is typically lower than 5 %.

To sum it up, the uncertainty at the optimal brightness threshold depends on the value of the coincidence fraction. If the coincidence fraction is almost 100 %, 800 selected bursts for both channels are sufficient to obtain an uncertainty of less than 5 %. There is no strict relation between the uncertainty and the initial number of bursts. Nonetheless, about 10 000 initial bursts lead most likely to an uncertainty below 5 %. For smaller coincidence fractions, the required number of bursts becomes less. Hence, 10 000 recorded initial bursts, or 800 selected bursts guarantee an uncertainty of less than 5 % for both channels.

# 6 Chance Coincidence Limitation

In the previous chapter, the required measurement size to obtain an uncertainty of less than 5 % on the coincidence fraction  $f$  was investigated.  $f$  is directly accessible from a BTCCD analysis. As discussed in Section 2.4.2,  $f$  systematically overestimates the binding fraction due to the occurrence of chance coincidence events. In the same section, an existing approach to correct for chance coincidences was presented. For the red channel, the corrected coincidence fraction is given by

$$f_{RB}^{cor}(n_{br}) = \frac{f_{RB}(n_{br}) - f_{RB,0}^{chance}(n_{br})}{1 - f_{RB,0}^{chance}(n_{br})}, \quad (6.1)$$

where

$$f_{RB,0}^{chance}(n_{br}) = 1 - \exp \left\{ -\langle N_B \rangle(0) \cdot \left( \frac{\langle \tau_d^R \rangle(n_{br})}{\langle \tau_d^B \rangle(0)} + 1 \right) \right\}. \quad (6.2)$$

The chance coincidence correction depends on the initial number of blue molecules  $\langle N_B \rangle(0)$  and the dwell time ratio  $\langle \tau_d^R \rangle(n_{br}) / \langle \tau_d^B \rangle(0)$ . For the blue channel, the correction has a similar form and depends on  $\langle N_R \rangle(0)$  and  $\langle \tau_d^B \rangle(n_{br}) / \langle \tau_d^R \rangle(0)$ . In the following, those two parameters are simply called molecule number and dwell time ratio.

In the first part of this chapter, see Section 6.1, it is investigated, for which molecule numbers and dwell time ratios the chance coincidence correction is applicable. Based on experimental data, a method is proposed that allows choosing a suitable molecule number for the measurement of a given sample. Then, in Section 6.2, a trial experiment validates the proposed method.

## 6.1 Application Limits of Chance Coincidence Correction

### 6.1.1 Measurement

To investigate the application limits of the existing chance coincidence correction, two measurement sets were recorded. Table 6.1 contains information on measurement time, burst threshold, background, and starting parameters on seven measurement rows with a mixture of red-labeled dsDNA and free Alexa 488. Alexa 488 is a dye that emits blue fluorescence light. It is further referred to as free blue dye. Using FCS, the measurement rows were performed with a sample concentration ranged between 20 pmol/L and 600 pmol/L. The second measurement set was conducted with a mixture of red-labeled ribosomes and free blue dye. Information on measurement time, burst threshold, background, and starting parameters for the six measurement rows can be found in Table 6.2. Here, the sample concentration lied between 40 pmol/L to 1400 pmol/L.

For all measurement rows, using FCS, it was determined that the mixture ratio was almost 1 : 1.

ch.	$T$ [min]	$\langle IPL^{bg} \rangle$ [ $\mu\text{s}$ ]	$\langle N \rangle [10^{-3}]$	$\langle \tau_d \rangle$ [ $\mu\text{s}$ ]	$\langle MB \rangle$ [kHz]
red blue	60	1585	$15.853 \pm 0.086$	$776.4 \pm 3.1$	$47.7 \pm 4.0$
		1413	$3.111 \pm 0.019$	$271.4 \pm 1.0$	$52.6 \pm 5.4$
red blue	33	282	$78.15 \pm 0.29$	$878.8 \pm 2.3$	$44.1 \pm 3.2$
		562	$16.102 \pm 0.064$	$302.22 \pm 0.73$	$49.7 \pm 2.6$
red blue	20	251	$79.28 \pm 0.38$	$877.9 \pm 3.0$	$43.8 \pm 1.7$
		562	$16.167 \pm 0.082$	$303.54 \pm 0.94$	$49.1 \pm 3.6$
red blue	60	794	$103.76 \pm 0.12$	$834.4 \pm 2.4$	$45.8 \pm 1.3$
		355	$23.903 \pm 0.060$	$321.67 \pm 0.49$	$47.0 \pm 1.8$
red blue	20	89	$126.312 \pm 0.051$	$924.1 \pm 2.6$	$41.5 \pm 1.3$
		251	$46.89 \pm 0.16$	$360.43 \pm 0.72$	$42.8 \pm 2.2$
red blue	18	89	$472.2 \pm 1.7$	$1190.2 \pm 2.7$	$36.61 \pm 0.70$
		141	$104.04 \pm 0.28$	$424.87 \pm 0.69$	$39.6 \pm 2.9$
red blue	20	89	$556.6 \pm 2.0$	$1259.1 \pm 2.7$	$36.3 \pm 1.3$
		141	$107.99 \pm 0.27$	$429.64 \pm 0.65$	$39.4 \pm 1.2$

Table 6.1: Measurement time, background, and starting parameters for measurements with a mixture of red-labeled dsDNA and free blue dye. For all measurements, the burst thresholds were set to  $IPL_R^{tre} = 75\mu\text{s}$  and  $IPL_R^{tre} = 80\mu\text{s}$ . The measurement rows are ordered with increasing sample concentration.

ch.	$T$ [min]	$IPL^{thr}$ [ $\mu\text{s}$ ]	$\langle IPL^{bg} \rangle$ [ $\mu\text{s}$ ]	$\langle N \rangle [10^{-3}]$	$\langle \tau_d \rangle$ [ $\mu\text{s}$ ]	$\langle MB \rangle$ [kHz]
red blue	60	280	1413	$29.89 \pm 0.15$	$1193.3 \pm 4.2$	$24.3 \pm 4.6$
		100	1259	$4.012 \pm 0.024$	$326.1 \pm 1.2$	$46.1 \pm 3.8$
red blue	53	280	1413	$30.33 \pm 0.15$	$1171.5 \pm 4.0$	$23.9 \pm 4.7$
		100	1413	$3.063 \pm 0.022$	$324.7 \pm 1.5$	$45.4 \pm 7.7$
red blue	20	150	178	$214.23 \pm 0.76$	$1119.0 \pm 2.6$	$24.1 \pm 2.3$
		80	447	$13.643 \pm 0.076$	$318.2 \pm 1.1$	$45.2 \pm 3.2$
red blue	20	85	89	$382.2 \pm 1.0$	$820.4 \pm 1.3$	$33.2 \pm 1.2$
		70	282	$29.54 \pm 0.11$	$306.36 \pm 0.69$	$48.0 \pm 8.5$
red blue	20	85	89	$638.0 \pm 1.8$	$1050.0 \pm 1.6$	$31.44 \pm 0.92$
		70	224	$43.04 \pm 0.14$	$325.26 \pm 0.64$	$45.4 \pm 1.8$
red blue	20	60	89	$1304.9 \pm 5.3$	$1313.9 \pm 2.0$	$42.33 \pm 0.83$
		60	112	$90.71 \pm 0.21$	$325.18 \pm 0.46$	$47.5 \pm 1.2$

Table 6.2: Measurement time, burst thresholds, background, and starting parameters for measurements with a mixture of red-labeled ribosomes and free blue dye. The measurement rows are ordered with increasing sample concentration.

### 6.1.2 Results

As explained above, the coincidence fraction  $f$  is directly obtained from the BTCCD analysis. For the red channel,  $f_{RB}$  is illustrated in Figures 6.1 and 6.2. The coincidence fraction is depicted for a small, intermediate, and high molecule number. Theoretically, a coincidence fraction of 0 % is expected because the sample solution contains only single-labeled molecules. Thus, the observed coincidence fraction is assumably only caused by chance coincidences. The

chance coincidence fraction becomes more pronounced for a larger molecule number and increases with  $n_{br}$ . The corresponding results for the blue channel are illustrated in Figures D.1 and D.2. For the blue channel, the same dependencies are observed.

For the conducted measurements, it is not possible to determine an optimal brightness threshold  $n_{br, opt}$  since a coincidence fraction of 0 % is expected. Therefore, the BTCCD analysis was applied for values of  $n_{br}$  between 0 and 10 to cover a reasonable range, in which the optimal brightness threshold typically lies.

Likewise, the optimal number of bursts criterion from the previous chapter cannot be applied because it refers to the uncertainty of the coincidence fraction at the optimal brightness threshold. Hence, although more than 10 000 initial bursts were recorded for every measurement, the uncertainty on the blue coincidence fraction becomes notably large for high brightness thresholds. Typically, there are fewer molecules in the blue channel than in the red channel, see e.g. Tables 6.1 and 6.2, because the red confocal detection volume is larger than the blue one. Therefore, the uncertainty on the blue coincidence fraction is more significant than on the red one.

To correct for the observed chance coincidences, the existing approach is applied. Figures 6.3 and 6.4 illustrate the corrected coincidence fraction for the red channel. A systematical deviation of the corrected coincidence fraction from the expected value of 0 % is observed. The deviation increases with the molecule number and with  $n_{br}$ . The results for the blue channel can be found in Figures D.3 and D.4. For the blue channel, the deviation from the expected coincidence fraction is not as expressed as in the red channel. More pronounced is the huge uncertainty on the corrected coincidence fraction for higher brightness thresholds. It is the result of the propagation of the uncertainty on the uncorrected coincidence fraction, which, as discussed above, reveals considerable uncertainties.

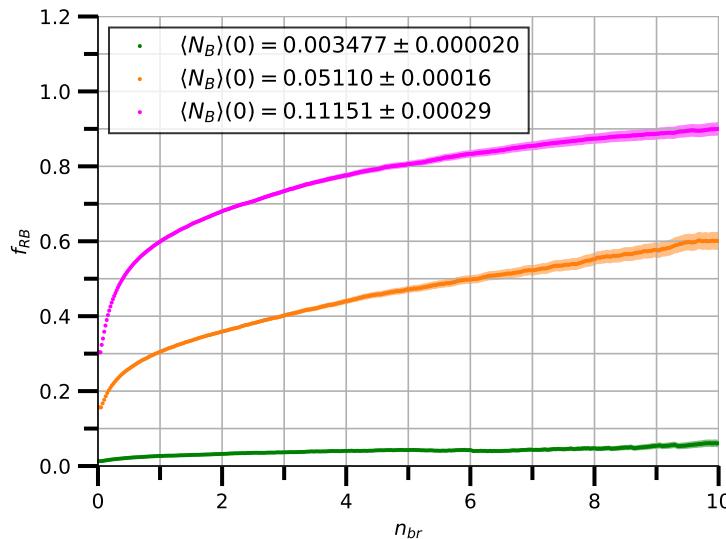


Figure 6.1: Coincidence fraction of red channel  $f_{RB}$  for the mixture of red-labeled dsDNA and free blue dye for a small, intermediate, and high molecule number. The observed coincidence fraction is primarily caused by chance coincidences and increases with  $\langle N_B \rangle (0)$  and  $n_{br}$ .

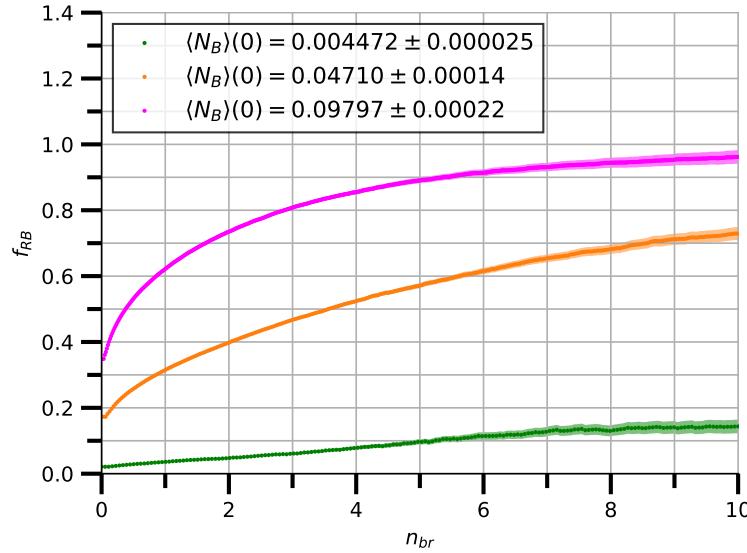


Figure 6.2: Coincidence fraction of red channel  $f_{RB}$  for the mixture of red-labeled ribosomes and free blue dye for a small, intermediate, and high molecule number. The observed coincidence fraction is primarily caused by chance coincidences and increases with  $\langle N_B \rangle(0)$  and  $n_{br}$ .

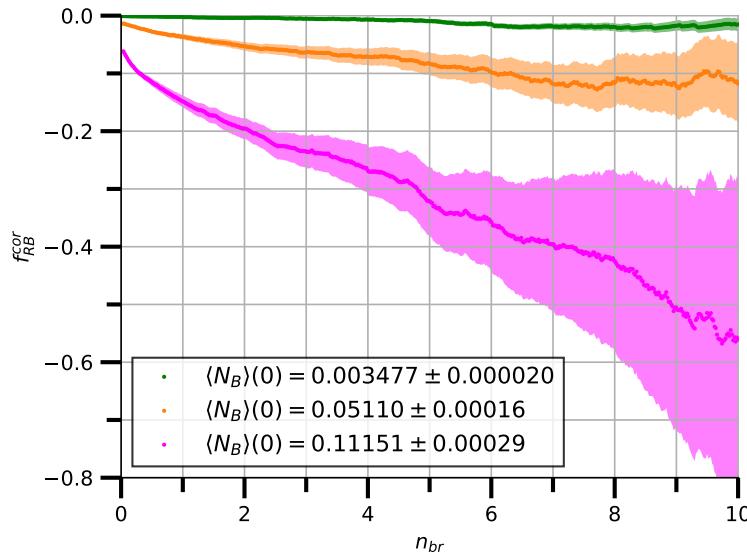


Figure 6.3: Corrected coincidence fraction of red channel  $f_{RB}^{cor}$  for the mixture of red-labeled dsDNA and free blue dye for a low, intermediate, and high molecule number. A deviation of the corrected coincidence fraction from the expected value of 0 % is observed. The deviation increases with  $\langle N_B \rangle(0)$  and  $n_{br}$ .

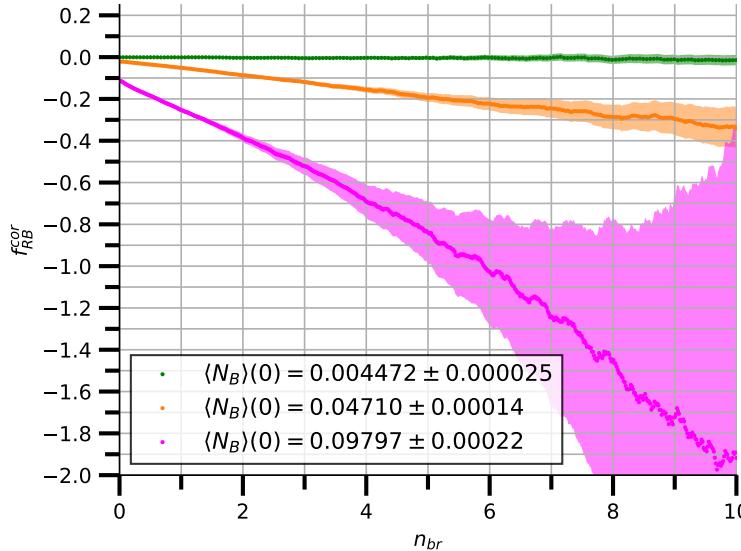


Figure 6.4: Corrected coincidence fraction of red channel  $f_{RB}^{cor}$  for the mixture of red-labeled ribosomes and free blue dye for a low, intermediate, and high molecule number. A deviation of the corrected coincidence fraction from the expected value of 0 % is observed. The deviation increases with  $\langle N_B \rangle (0)$  and  $n_{br}$ .

### 6.1.3 Systematic Deviation of Corrected Coincidence Fraction

For the conducted measurements, the theoretical coincidence fraction of 0 % is known. Therefore, the observed coincidence fraction is mostly caused by chance coincidences. According to Equation (6.2), the chance coincidence fraction increases with the molecule number and the dwell time ratio. In the measurement results, the dependency on the molecule number was directly observed. To understand the dependency on the dwell time ratio, the knowledge from Section 2.4.3 that the dwell time ratio is a strictly monotonically increasing function of  $n_{br}$  can be used. Then, the observed increase of the coincidence fraction with  $n_{br}$  can be explained.

It remains the question, why the corrected coincidence fraction deviates clearly from the expected value of 0 %. For this purpose, it is useful to rewrite the expression for the corrected coincidence fraction, see Equation (6.1), as

$$f_{RB}^{cor}(n_{br}) = \left( f_{RB}(n_{br}) - f_{RB,0}^{chance}(n_{br}) \right) \cdot \exp \left\{ \langle N_B \rangle (0) \cdot \left( \frac{\langle \tau_d^R \rangle (n_{br})}{\langle \tau_d^B \rangle (0)} + 1 \right) \right\}. \quad (6.3)$$

The first term describes the deviation of the theoretical expression for the chance coincidence fraction  $f_{RB,0}^{chance}$  from the measured coincidence fraction  $f_{RB}$ . For the conducted measurements, the first term should be approximately 0 % because the observed coincidence fraction was mostly caused by chance coincidences. The second term describes an exponential increase in terms of the molecule number and the dwell time ratio. It is required to rescale the corrected coincidence fraction because  $f_{RB,0}^{chance}$  is derived under the assumption that no real coincidences are present. However, if molecular complexes occur, the effective molecule number that contribute to chance coincidences decreases. The second term takes this phenomenon

into account [3]. For the blue channel, the expression for the corrected coincidence fraction can be rewritten similarly.

First, the deviation of the first term from 0 % is analyzed. For the red channel, Figures 6.5 and 6.6 illustrate the first term for a small, intermediate, and large molecule number. The deviation from the expected value lies roughly between 0 % and 5 % for the dsDNA measurement, and between 0 % and 7 % for the ribosome measurement. The smallest deviation is obtained for a small molecule number. Nonetheless, the deviation does not depend deterministically on the molecule number nor on  $n_{br}$ . The deviation for the blue channel is shown in Figures D.5 and D.6. For the blue channel, the deviation of the first term from the expected value is not as much pronounced as for the red channel. The deviation for small brightness thresholds is of the same order as for the red channel. However, for large  $n_{br}$ , the deviation becomes huge and takes even values of up to 40 %. As explained above, only a small number of blue bursts remains for these regions, and explains the deviation.

So, the first term shows a slight deviation from the expected value of 0 %. With this knowledge, formally, the first term can be rewritten as

$$f_{RB}(n_{br}) - f_{RB,0}^{chance}(n_{br}) = f_{RB}^{true}(n_{br}) - f_{RB,0}^{chance,true}(n_{br}) + \Delta f_{RB}(n_{br}). \quad (6.4)$$

Here,  $f_{RB}^{true}$  denotes the true coincidence fraction that would be measured if there were no uncertainty.  $f_{RB,0}^{chance,true}$  describes an ideal estimation of the chance coincidence fraction. Both quantities are unknown. The deviation between the true and the actual values is absorbed into  $\Delta f_{RB}$ . Inserting this formulation into Equation (6.3), yields

$$f_{RB}^{cor}(n_{br}) = \left( f_{RB}^{true}(n_{br}) - f_{RB,0}^{chance,true}(n_{br}) + \Delta f_{RB}(n_{br}) \right) \cdot \exp \left\{ \langle N_B \rangle(0) \cdot \left( \frac{\langle \tau_d^R \rangle(n_{br})}{\langle \tau_d^B \rangle(0)} + 1 \right) \right\}. \quad (6.5)$$

Thus, if the first term deviates slightly by  $\Delta f_{RB}$  from its true value, the deviation is exponentially amplified by the second term. The specific amplification depends on the molecule number and on the dwell time ratio. Since the dwell time ratio is a function of  $n_{br}$ , this explains the increasing deviation observed for the corrected coincidence fraction. For the blue channel, the dwell time ratio takes smaller values than for the red channel because the dwell time of blue dye is shorter than that of dsDNA. Hence, the systematic deviation in the blue channel is less pronounced.

To sum it up, the observed deviation of the corrected coincidence fraction from the expected value of 0 % is the combination of a slight deviation of the first term and a systematic exponential amplification of the second term. Nonetheless, the latter term is necessary if the coincidence fraction is not expected to be 0 %, as discussed above. Therefore, a method is required that determines if the existing chance coincidence correction is applicable for a particular molecule number and the corresponding range of the dwell time ratio.

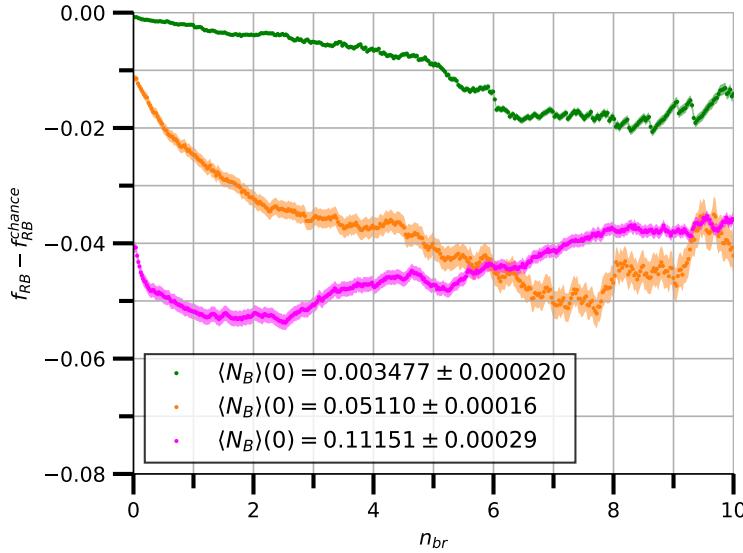


Figure 6.5: First term of corrected coincidence fraction of red channel  $f_{RB} - f_{RB,0}^{chance}$  for the mixture of red-labeled dsDNA and free blue dye for a small, intermediate, and high molecule number. The expected value of the first term is 0 %. The deviation from that value lies roughly between 0 % and 5 %.

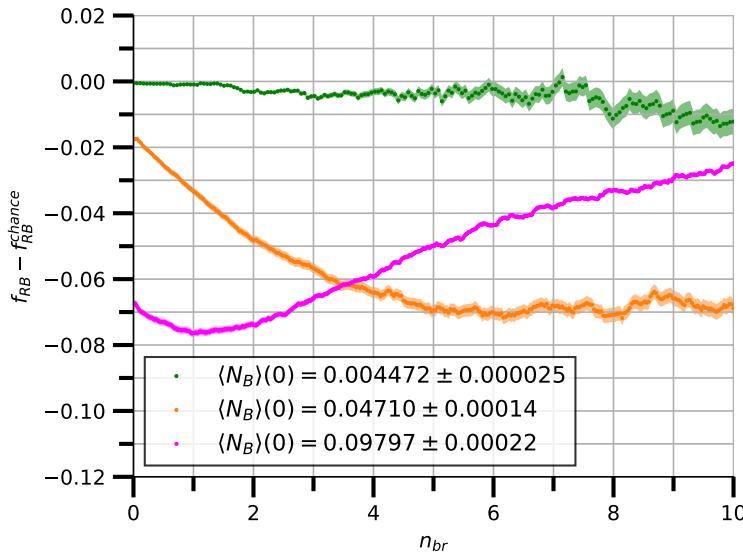


Figure 6.6: First term of corrected coincidence fraction of red channel  $f_{RB} - f_{RB,0}^{chance}$  for the mixture of red-labeled ribosomes and free blue dye for a small, intermediate, and high molecule number. The expected value of the first term is 0 %. The deviation from that value lies roughly between 0 % and 7 %.

#### 6.1.4 Application Range of Chance Coincidence Correction

The application range of the existing chance coincidence correction can be derived from the conducted measurements. For this purpose, two criteria are defined, to decide, whether the chance coincidence correction leads to consistent results. Both the deviation of the corrected coincidence fraction and its statistical uncertainty should not exceed 5 %. The motivation for the second criterion is that it does not make sense to search for a systematic deviation if the statistical uncertainty is already of the same order. The concrete threshold of 5 % is based on typical requirements on the accuracy of the binding fraction.

Figure 6.7 illustrates the application range for the red channel. For every measured molecule number, the range of the dwell time ratio for which the criteria are fulfilled is plotted as a vertical line. Remember, that the dwell time ratio is a strictly monotonically increasing function of  $n_{br}$ . Hence, the lowest dwell time ratio corresponds to  $n_{br} = 0$ . There are two possible reasons why the vertical lines end at a particular dwell time ratio. One reason is that the criteria are not anymore fulfilled for larger dwell time ratios. In the figure, this case is marked with a short horizontal line at the highest dwell time ratio. If the criterion holds for the whole range of brightness thresholds from 0 to 10, the ending is not specifically marked. A simple dot denotes that the criteria are not fulfilled for the whole range of dwell time ratios. As expected, the application range of the chance coincidence correction in terms of the dwell time ratio decreases for larger molecule numbers.

The next step is to interpolate the obtained results for the application range, see Figure 6.8. The interpolation allows to determine the application range for an arbitrary molecule number. Note that the interpolation serves only for illustration and is not calculated from a theoretical relation. It is solely based on the obtained measurement results.

So far, the application range in terms of the dwell time ratio can be determined for every molecule number. The sample concentration roughly specifies the molecule number. In contrast, to obtain knowledge on the range of the dwell time ratio, an accelerated BTCCD can be applied. There, the dwell time ratio for  $n_{br} = 0$  and for a larger brightness threshold is calculated. Prior to a precise measurement, the optimal brightness threshold  $n_{br,opt}$  is unknown. Therefore, it is advised to consider a range of  $n_{br}$  in which  $n_{br,opt}$  typically lies, and to calculate the dwell time ratio for the high brightness threshold. If the range of dwell time ratios does not lie completely in the application range of the chance coincidence correction, the molecule number needs to be decreased. Otherwise, it is not ensured that the chance coincidence correction is applicable at the optimal brightness threshold.

For the blue channel, the raw data for the application range of the chance coincidence correction and the interpolation can be found in Figures D.7 and D.8. The main limit for the blue channel is that the statistical uncertainty on the corrected coincidence fraction exceeds 5 %. Due to the relative low number of bursts for high brightness thresholds, it becomes considerable large for large dwell time ratios.

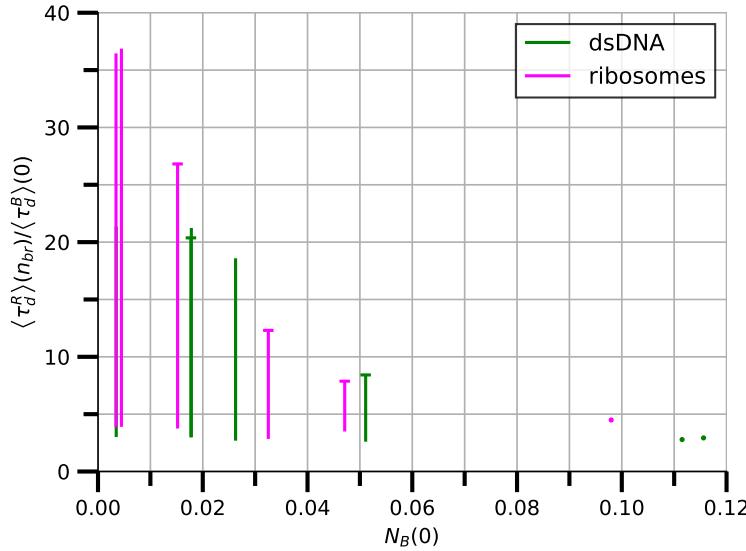


Figure 6.7: Application range in terms of dwell time ratio as a function of the molecule number for the red channel. Every vertical line represents one measurement. A short horizontal line at the highest dwell time ratio indicates that the deviation of the corrected coincidence fraction or its statistical uncertainty becomes more than 5 % for higher dwell time ratios. A simple dot denotes that the chance coincidence correction is unsuitable for the whole measurement.

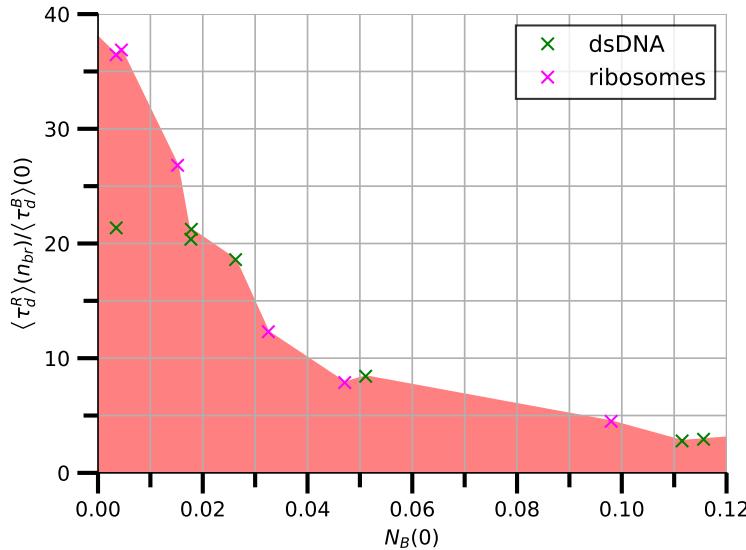


Figure 6.8: Interpolated application range in terms of dwell time ratio as a function of the molecule number for the red channel. The interpolation serves only for illustration and is not derived from theoretical knowledge. It allows to determine the application range for an arbitrary molecule number.

## 6.2 Validation of Application Limits

### 6.2.1 Measurement

Six measurement rows with a mixture of red-labeled dsDNA and dual-labeled dsDNA were conducted. Using FCS, it was determined that the mixture ratio was almost 1 : 1. Information on measurement time, burst thresholds, background, and starting parameters can be found in Table 6.3. The sample concentration ranged between 10 pmol/L and 400 pmol/L.

ch.	$T$ [min]	$IPL^{thr}$ [ $\mu$ s]	$\langle IPL^{bg} \rangle$ [ $\mu$ s]	$\langle N \rangle [10^{-3}]$	$\langle \tau_d \rangle$ [ $\mu$ s]	$\langle MB \rangle$ [kHz]
red blue	60	120	1778	$17.95 \pm 0.11$	$1223.1 \pm 5.3$	$34.5 \pm 1.5$
		120	2512	$2.407 \pm 0.030$	$783.6 \pm 6.3$	$28.7 \pm 3.9$
red blue	60	120	1413	$25.78 \pm 0.13$	$1202.3 \pm 4.3$	$34.1 \pm 1.5$
		120	891	$4.507 \pm 0.037$	$634.3 \pm 3.3$	$23.6 \pm 1.2$
red blue	20	100	316	$91.70 \pm 0.46$	$1137.2 \pm 4.0$	$35.5 \pm 2.0$
		100	794	$9.754 \pm 0.093$	$618.3 \pm 3.8$	$28.7 \pm 2.2$
red blue	20	100	200	$122.89 \pm 0.56$	$1169.4 \pm 3.7$	$33.5 \pm 1.5$
		100	708	$12.68 \pm 0.11$	$622.8 \pm 3.3$	$27.6 \pm 1.1$
red blue	20	60	89	$741.2 \pm 2.8$	$1317.1 \pm 2.8$	$37.2 \pm 1.0$
		100	251	$68.34 \pm 0.29$	$745.4 \pm 2.1$	$26.86 \pm 0.64$
red blue	20	60	89	$757.1 \pm 2.7$	$1142.4 \pm 2.3$	$41.9 \pm 1.7$
		100	251	$78.53 \pm 0.32$	$758.3 \pm 2.0$	$27.00 \pm 0.68$

Table 6.3: Measurement time, burst thresholds, background, and starting parameters for measurements of a mixture of red-labeled dsDNA and dual-labeled dsDNA. The measurement rows are ordered with increasing sample concentration.

### 6.2.2 Results and Discussion

To validate the application limits of the chance coincidence correction, first, the optimal brightness threshold for the six measurement rows is determined. Table 6.4 gives an overview of the results for the red channel. Besides the optimal brightness threshold, it includes the molecule number, the relevant dwell time ratio, the uncorrected coincidence fraction, and the corrected coincidence fraction. The position of the optimal brightness thresholds in the application range illustration from the previous chapter can be found in Figure 6.9. The first four measurement rows lie definitively in the application range, while the two measurement rows with the highest concentration are outside it. In general, the measurement results reflect the mixing ratio of 1 : 1 as the corrected coincidence fraction is roughly 50 %. Hence, a reasonable guess for the actual binding fraction is the corrected coincidence fraction for the first measurement. There, the molecule number is small, and it is expected that the chance coincidence correction leads to accurate results. Then, the other measurements can be compared with this reference value. Here, the corrected coincidence fractions spread around the first value, and no systematic trend is observable.

The optimal brightness threshold is different for every measurement row. This complicates the observation of systematical deviations because the coincidence fractions vary additionally.

Hence, another approach is to compare the corrected coincidence fractions for a fixed brightness threshold. Here, the optimal brightness threshold  $n_{br} = 6.40$  of the first measurement row is taken as that value. The results are displayed in Table 6.5. It can be seen that the two measurement rows with the highest molecule number deviate by 3.9 % and 5.1 %, respectively, from the first measurement row. For those measurement rows, a deviation is expected as they lie outside the application range, see Figure 6.9.

$n_{br,opt}$	$\langle N_B \rangle(0) [10^{-3}]$	$\langle \tau_d^R \rangle(n_{br,opt}) / \langle \tau_d^B \rangle(0)$	$f_{RB}(n_{br,opt}) [\%]$	$f_{RB}^{cor}(n_{br,opt}) [\%]$
6.40	$2.529 \pm 0.030$	$7.480 \pm 0.016$	$50.7 \pm 2.3$	$49.6 \pm 2.3$
5.04	$5.057 \pm 0.037$	$8.977 \pm 0.010$	$46.3 \pm 1.4$	$43.5 \pm 1.5$
5.05	$10.538 \pm 0.094$	$8.987 \pm 0.011$	$52.8 \pm 1.3$	$47.6 \pm 1.5$
8.37	$13.68 \pm 0.11$	$12.479 \pm 0.011$	$61.9 \pm 2.2$	$54.2 \pm 2.7$
9.28	$72.52 \pm 0.29$	$15.8994 \pm 0.0056$	$85.2 \pm 1.3$	$49.7 \pm 4.4$
9.63	$83.21 \pm 0.32$	$14.2016 \pm 0.0053$	$86.3 \pm 1.2$	$51.3 \pm 4.4$

Table 6.4: Corrected coincidence fraction  $f_{RB}^{cor}$  of red channel at the optimal brightness threshold  $n_{br,opt}$  for measurements with a mixture of red-labeled dsDNA and dual-labeled dsDNA. For every measurement row, the molecule number, the relevant dwell time ratio, and the uncorrected coincidence fraction are stated. The first measurement is taken as a guess for the true binding fraction. The corrected coincidence fractions spread around the first value, and no systematic trend is observable.

$\langle N_B \rangle(0) [10^{-3}]$	$\langle \tau_d^R \rangle(6.40) / \langle \tau_d^B \rangle(0)$	$f_{RB}(6.40) [\%]$	$f_{RB}^{cor}(6.40) [\%]$
$2.529 \pm 0.030$	$7.480 \pm 0.016$	$50.7 \pm 2.3$	$49.6 \pm 2.3$
$5.057 \pm 0.037$	$10.225 \pm 0.011$	$48.7 \pm 1.9$	$45.7 \pm 2.0$
$10.538 \pm 0.094$	$10.053 \pm 0.012$	$55.2 \pm 1.7$	$49.7 \pm 1.9$
$13.68 \pm 0.11$	$10.711 \pm 0.012$	$59.5 \pm 1.6$	$52.4 \pm 1.9$
$72.52 \pm 0.29$	$12.9324 \pm 0.0052$	$80.24 \pm 0.87$	$45.7 \pm 2.5$
$83.2100 \pm 0.0032$	$11.2796 \pm 0.0048$	$80.00 \pm 0.80$	$44.5 \pm 2.3$

Table 6.5: Corrected coincidence fraction  $f_{RB}^{cor}$  of red channel at  $n_{br} = 6.40$  for measurements with a mixture of red-labeled dsDNA and dual-labeled dsDNA. For every measurement row, the molecule number, the relevant dwell time ratio, and the uncorrected coincidence fraction are stated. The first measurement is taken as a guess for the true binding fraction. The last two corrected coincidence fractions corresponding to the highest molecule numbers deviate by 3.9 % and 5.1 % from the first one.

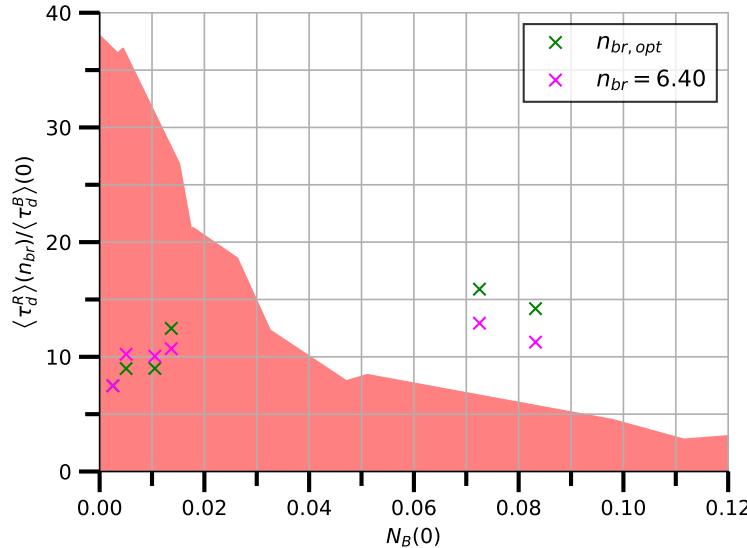


Figure 6.9: Position of the validation experiments for the optimal brightness threshold (see Table 6.4) and for  $n_{br} = 6.40$  (see Table 6.5) in the illustration of the application range of the chance coincidence fraction for the red channel. The first four measurements lie definitively in the application range, while the two measurements with the highest concentration are outside it.

The observed deviation from the reference value is still small. This is caused by the relative small dwell time ratios compared to those for the mixtures of ribosomes and free dye in the previous chapter. For the analysis of the application limits, it would be useful to measure samples with different diffusing properties that lead to larger dwell time ratios. Then, a clearer systematic trend would be expected. However, yet, such measurements are not possible due to an existing shortcoming of BTCCD. It cannot be applied for samples with significantly different initial dwell times if the binding fraction is larger than 0 %.<sup>1</sup> The reason for this is that, if a complex between the molecule types is formed, it has necessarily a larger dwell time than the faster molecule. Then, the application of a brightness threshold will exclude bursts corresponding to the faster single molecule because they are considered as dim compared to the bursts of the complex. Thus, for large  $n_{br}$ , the coincidence fraction falsely approaches 100 %.

A possible solution could be substituting the brightness threshold with another parameter, e.g., the molecular brightness or the IPL. Those parameters allow the exclusion of dim bursts because they take their extreme values for molecules with a central trajectory through the confocal detection volume. However, in contrast to the brightness threshold, they depend only on the dye properties and do not exclude bursts by their absolute dwell time.

For the blue channel, the corrected coincidence fraction at the optimal brightness threshold can be found in Table D.1. Table D.2 states the data for a fixed brightness threshold. For both cases, the corrected coincidence fraction is significantly smaller than 100 % as would be expected from the mixing ratio. As explained above, the determination of the coincidence frac-

<sup>1</sup>For the mixtures of ribosomes and free dye in the previous chapter, the dwell times differed significantly. However, since the expected binding fraction was 0 %, BTCCD was applicable.

tion in the blue channel underlies in general larger fluctuations because of the lower number of bursts. This is also the reason for the increasing uncertainties on the corrected coincidence fractions in agreement with the positions in the application range illustration, see Figure D.9.

## 7 Conclusion and Outlook

The central part of this thesis was concerned with the occurrence of chance coincidences during the application of brightness-gated two-color coincidence detection (BTCCD). The investigation of a sample with an expected binding fraction of 0 % verified that the chance coincidence fraction increases with the molecule number and the dwell time ratio of the involved samples.

Chance coincidences lead to a systematic overestimation of the coincidence fraction. In this thesis, the limitation of BTCCD due to chance coincidences was investigated. For this purpose, the applicability of an existing approach to correct for chance coincidences was analyzed. It was shown that the chance coincidence correction only leads to accurate results for a limited range of dwell time ratios that depends on the molecule number. The knowledge of the application range allows choosing a suitable molecule number for given samples, for which the correction is applicable.

The derived application limits of the chance coincidence correction were validated by a trial experiment, where the expected binding fraction was 50 %. The validation revealed an existing shortcoming of BTCCD. It cannot be applied for sample molecules with significantly differing diffusing properties if the actual binding fraction is larger than 0 %. A possible solution could be substituting the brightness threshold with another parameter, e.g., the molecular brightness or the IPL. However, this situation complicated the validation process because only a limited range of dwell time ratios during the application of BTCCD was accessible. Therefore, the corrected coincidence fraction could solely be determined for dwell time ratios slightly outside the application range. However, even those values confirm the derived application limits of the chance coincidence correction.

All in all, the occurrence of chance coincidences limits the precision with which the binding fraction can be determined. Thus, prior to the application of BTCCD, it has to be checked that the expected range of dwell time ratios lies inside the application range of the chance coincidence correction. In practice, the sample concentration has to be adjusted to the specific diffusing properties of the molecules under investigation.

Besides the application limits of chance coincidences, other properties of BTCCD were analyzed.

First of all, it was shown that roughly 10 000 photon bursts need to be recorded if the uncertainty on the coincidence fraction before the correction should be smaller than 5 %.

Moreover, the dependencies of the average values of dwell time, molecule number, and molecular brightness on the brightness threshold were investigated. Based on theoretical assumptions, mathematical expressions for those quantities were derived. The experiment revealed that their development is guided by two processes that take place on the time scale of the dwell time. A fast process is identified as the exclusion of dim bursts due to the application of brightness-gating. Several mechanisms that influence fluorescence characteristics could be the underlying reason for a second, slower process. One of these is the transition of electrons in a triplet state. Another relevant effect could be the charge of the used fluorescent dyes and their Coulomb interaction with the sample, e.g., with the overall positive charge of DNA. The

specific reasons for the second process remain unknown and are object of further research.

Finally, characterizing the starting parameters of a BTCCD analysis, it was noted that the smoothing of the IPL time trace leads to nonphysical bursts that contain zero or one photon. For these bursts, no meaningful dwell time can be defined. Thus, BTCCD could be improved if those bursts were excluded before the application of a brightness threshold.

# Bibliography

- [1] N. R. Zaccai, I. N. Serdyuk, and J. Zaccai. *Methods in Molecular Biophysics*. Second Edition. Cambridge, UK: Cambridge University Press, 2017. ISBN: 978-1-107-05637-4.
- [2] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Third Edition. New York, USA: Springer Science and Business Media LLC, 2006. ISBN: 978-0387-31278-1.
- [3] H. Höfig. “Single-Molecule Characterization of FRET-Based Biosensors and Development of Two-Color Coincidence Detection.” PhD thesis. RWTH Aachen, Germany, 2020.
- [4] I. V. Gopich. “Concentration Effects in Single-Molecule Spectroscopy.” In: *The Journal of Physical Chemistry B* 112.19 (2008), pp. 6214–6220.
- [5] R. Li and E. C. Lim. “Quantitative Study of Luminescence in Aromatic and Heteroaromatic Molecules.” In: *The Journal of Chemical Physics* 57.2 (1972), pp. 605–612.
- [6] J. W. Lichtman and J.-A. Conchello. “Fluorescence Microscopy.” In: *Nature Methods* 2.12 (2005), pp. 910–919.
- [7] Thermo Fisher Scientific Inc. *Fluorescence SpectraViewer*. URL: <https://www.thermofisher.com/de/de/home/life-science/cell-analysis/labeling-chemistry/fluorescence-spectraviewer.html?SID=srch-svtool&UID=21235p72#!/> (visited on 06/16/2020).
- [8] Laser Components GmbH. *Single Photon Counting Module COUNT T-Series*. Product Sheet. Olching, Germany.
- [9]  $\tau$ -SPAD: *Single Photon Counting Module*. Product Sheet. PicoQuant GmbH, Berlin, Germany.
- [10] *MicroTime 200: Fluorescence Lifetime Microscope*. Version 6.1. Manual. PicoQuant GmbH, Berlin, Germany.
- [11] P. Schwille and E. Haustein. “Fluorescence Correlation Spectroscopy - An Introduction to Its Concepts and Applications.” In: *Spectroscopy* 94(22) (2001).
- [12] B. K. Müller et al. “Pulsed Interleaved Excitation.” In: *Biophysical Journal* 89.5 (2005), pp. 3508–3522.
- [13] M. Wahl and S. Orthaus-Müller. *Time Tagged Time-Resolved Fluorescence Data Collection in Life Sciences*. Technical Note. PicoQuant GmbH, Berlin, Germany.
- [14] P. Kapusta. *Absolute Diffusion Coefficients: Compilation of Reference Data For FCS Calibration*. Revision 1. Application Note. PicoQuant GmbH, Berlin, Germany, 2010.
- [15] V. Buschmann et al. *Quantitative FCS: Determination of The Confocal Volume By FCS and Bead Scanning With The MicroTime 200*. Version 1.1. Application Note. PicoQuant GmbH, Berlin, Germany, 2009.
- [16] A. Orte, R. Clarke, and D. Klenerman. “Single-Molecule Two-Colour Coincidence Detection to Probe Biomolecular Associations.” In: *Biochemical Society Transactions* 38.4 (2010), pp. 914–918.

- [17] H. Höfig et al. “Brightness-Gated Two-Color Coincidence Detection Unravels Two Distinct Mechanisms in Bacterial Protein Translation Initiation.” In: *Communications Biology* 2.1 (2019).
- [18] H. Höfig et al. *BTCCD*. 2019. URL: <https://github.com/BTCCD/BTCCD> (visited on 06/04/2020).
- [19] Anaconda. *AnacondaSoftwareDistribution*. 2020. URL: <https://www.anaconda.com/> (visited on 06/28/2020).
- [20] L. C. Zanetti-Domingues et al. “Hydrophobic Fluorescent Probes Introduce Artifacts into Single Molecule Tracking Experiments Due to Non-Specific Binding.” In: *PLoS ONE* 8.9 (2013), e74200.
- [21] K. Bielec et al. “Analysis of Brightness of a Single Fluorophore For Quantitative Characterization of Biochemical Reactions.” In: *The Journal of Physical Chemistry B* 124.10 (2020), pp. 1941–1948.

# A Appendix: Theory and Methods

## A.1 Procedure of BTCCD

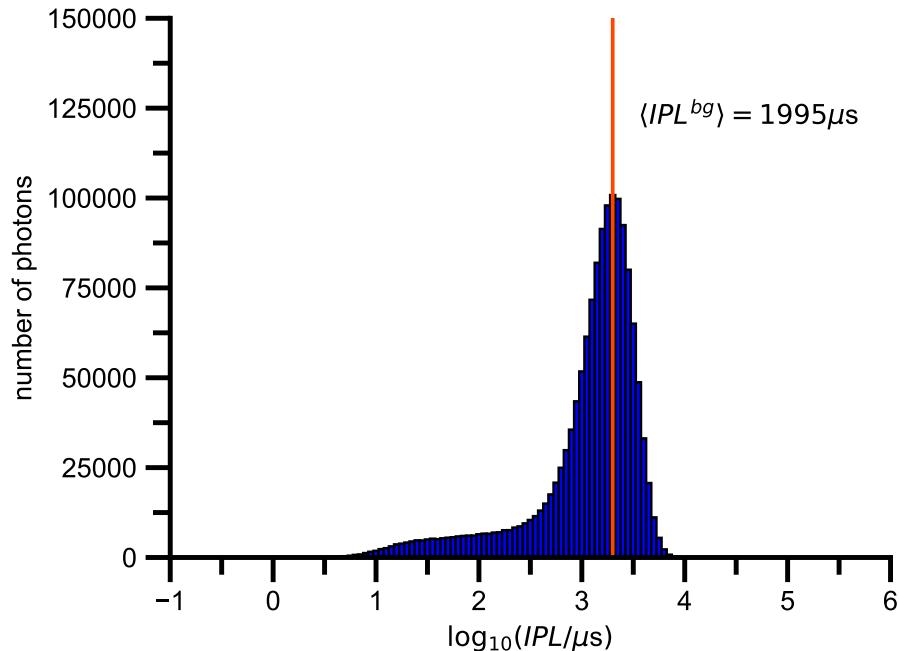


Figure A.1: Histogram of IPL values for the blue channel. The right population reflects the background. In this case, it is characterized by  $\langle IPL^{bg} \rangle = 1995 \mu\text{s}$  as indicated by the vertical orange line. The data is taken from a measurement of dual-labeled dsDNA, see Section 4.1.

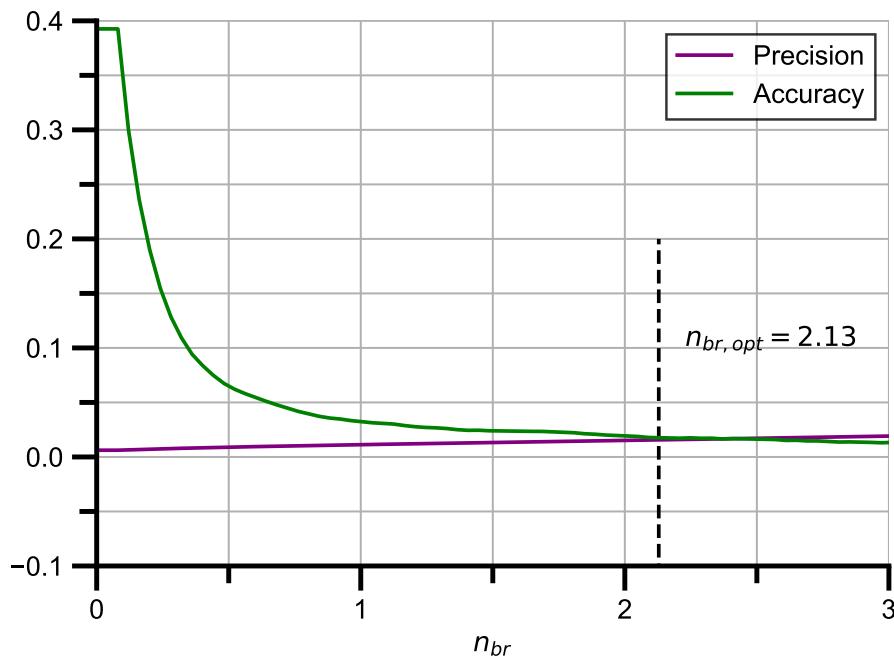


Figure A.2: Search for the optimal brightness threshold  $n_{br,opt}$  for the blue channel.  $n_{br,opt}$  is given by the intersection of precision and accuracy. The data is taken from a measurement of dual-labeled dsDNA, see Section 5.1. In this case,  $n_{br,opt} = 2.13$  leads to a coincidence fraction of  $f_{BR}(n_{br,opt}) = (95.3 \pm 1.7)\%$ .

## A.2 Uncertainty on Dwell Time, Number of Molecules, and Molecular Brightness

The average dwell time  $\langle \tau_d \rangle$  is calculated from the single dwell times  $\tau_{d,i}$  of  $B$  selected bursts according to

$$\langle \tau_d \rangle(n_{br}) = \frac{1}{B(n_{br})} \sum_{i \in B(n_{br})} \tau_{d,i}. \quad (\text{A.1})$$

Here,  $i \in B(n_{br})$  denotes the sum over all selected bursts at a given brightness threshold. The uncertainty on a single dwell time  $\tau_{d,i}$  is given by

$$\sigma_{\tau_{d,i}} = \sqrt{\frac{1}{B(n_{br}) - 1} \sum_{j \in B(n_{br})} (\tau_{d,j} - \langle \tau_d \rangle(n_{br})).} \quad (\text{A.2})$$

Then, the uncertainty on the average  $\langle \tau_d \rangle$  decreases to

$$\sigma_{\langle \tau_d \rangle}(n_{br}) = \frac{\sigma_{\tau_{d,i}}}{\sqrt{B(n_{br})}}. \quad (\text{A.3})$$

The molecular brightness of a single burst is expressed by

$$MB_i = \frac{N_{\gamma,i}}{\tau_{d,i}}, \quad (\text{A.4})$$

where  $N_{\gamma,i}$  denotes the number of photons in the  $i$ -th burst. Since  $N_{\gamma,i}$  is Poisson distributed, its uncertainty is given by  $\sigma_{N_{\gamma,i}} = \sqrt{N_{\gamma,i}}$ . The average molecular brightness is derived by

$$\langle MB \rangle(n_{br}) = \frac{1}{B(n_{br})} \sum_{i \in B(n_{br})} MB_i. \quad (\text{A.5})$$

Thus, the uncertainty of  $N_{\gamma,i}$  and  $\tau_{d,i}$  propagate on  $\langle MB \rangle$  by

$$\sigma_{\langle MB \rangle}(n_{br}) = \sqrt{\sum_{i \in B(n_{br})} \left\{ \left( \frac{\sigma_{N_{\gamma,i}}}{N_{\gamma,i}} \right)^2 + \left( \frac{\sigma_{\tau_{d,i}}}{\tau_{d,i}} \right)^2 + \left( \frac{\sigma_B(n_{br})}{B(n_{br})} \right)^2 \right\}} \quad (\text{A.6})$$

$$= \sqrt{\sum_{i \in B(n_{br})} \left\{ \frac{1}{N_{\gamma,i}} + \left( \frac{\sigma_{\tau_{d,i}}}{\tau_{d,i}} \right)^2 + \frac{1}{B(n_{br})} \right\}}. \quad (\text{A.7})$$

Finally, since  $\langle N \rangle$  is calculated by (2.17), its uncertainty is given by

$$\sigma_{\langle N \rangle}(n_{br}) = \frac{1}{1 - \frac{B(n_{br}) \cdot \langle \tau_d \rangle(n_{br})}{T}} \sqrt{\left( \frac{\langle \tau_d \rangle(n_{br})}{T} \sigma_B(n_{br}) \right)^2 + \left( \frac{B(n_{BR})(n_{br})}{T} \sigma_{\langle \tau_d \rangle}(n_{br}) \right)^2} \quad (\text{A.8})$$

$$= \frac{1}{1 - \frac{B(n_{br}) \cdot \langle \tau_d \rangle(n_{br})}{T}} \sqrt{\left( \frac{\langle \tau_d \rangle(n_{br})}{T} \sqrt{B(n_{br})} \right)^2 + \left( \frac{B(n_{BR})(n_{br})}{T} \sigma_{\langle \tau_d \rangle}(n_{br}) \right)^2} \quad (\text{A.9})$$

### A.3 Uncertainty on Corrected Coincidence Fraction

Formally, the uncertainty on  $f_{RB}^{cor}$  can be expressed by

$$\sigma_{f_{RB}^{cor}}(n_{br}) = f_{RB}^{cor}(n_{br}) \sqrt{\frac{(\sigma_{f_{RB}^{cor}}(n_{br}))^2 + (\sigma_{f_{RB,0}^{chance}}(n_{br}))^2}{(f_{RB}(n_{br}) - f_{RB,0}^{chance}(n_{br}))^2} + \frac{(\sigma_{f_{RB,0}^{chance}}(n_{br}))^2}{(1 - f_{RB,0}^{chance}(n_{br}))^2}}. \quad (\text{A.10})$$

According to Equation (2.15), the uncertainty on the uncorrected coincidence fraction  $f_{RB}$  is given by

$$\sigma_{f_{RB}}(n_{br}) = f_{RB}(n_{br}) \cdot \sqrt{\frac{1}{B_{RB}(n_{br})} + \frac{1}{B_R(n_{br})}},$$

where  $B_R$  is the total number of selected red bursts and  $B_{RB}$  is the number of coincident bursts. The uncertainty on  $f_{RB}^{chance}$  still needs to be determined. It can be expressed by

$$\begin{aligned} \sigma_{f_{RB,0}^{chance}}(n_{br}) &= \exp \left\{ -\langle N_B \rangle(0) \cdot \left( \frac{\langle \tau_d^R \rangle(n_{br})}{\langle \tau_d^B \rangle(0)} + 1 \right) \right\} \times \\ &\quad \left\{ \left( \frac{\langle \tau_d^R \rangle(n_{br})}{\langle \tau_d^B \rangle(0)} + 1 \right)^2 \cdot (\sigma_{\langle N_B \rangle}(0))^2 + \left( \frac{\langle N_B \rangle(0)}{\langle \tau_d^B \rangle(0)} \sigma_{\langle \tau_d^R \rangle}(n_{br}) \right)^2 + \right. \\ &\quad \left. \left( \frac{\langle N_B \rangle(0) \langle \tau_d^R \rangle(n_{br})}{(\langle \tau_d^B \rangle(0))^2} \sigma_{\langle \tau_d^B \rangle}(0) \right)^2 \right\}^{1/2}. \end{aligned} \quad (\text{A.11})$$

The result can also be used to determine the uncertainty  $\sigma_{f_{BR}^{cor}}$  on  $f_{BR}^{cor}$  by substituting  $RB \rightarrow BR$ .

## B Appendix: Dependencies on Brightness Threshold

### B.1 Dwell Time

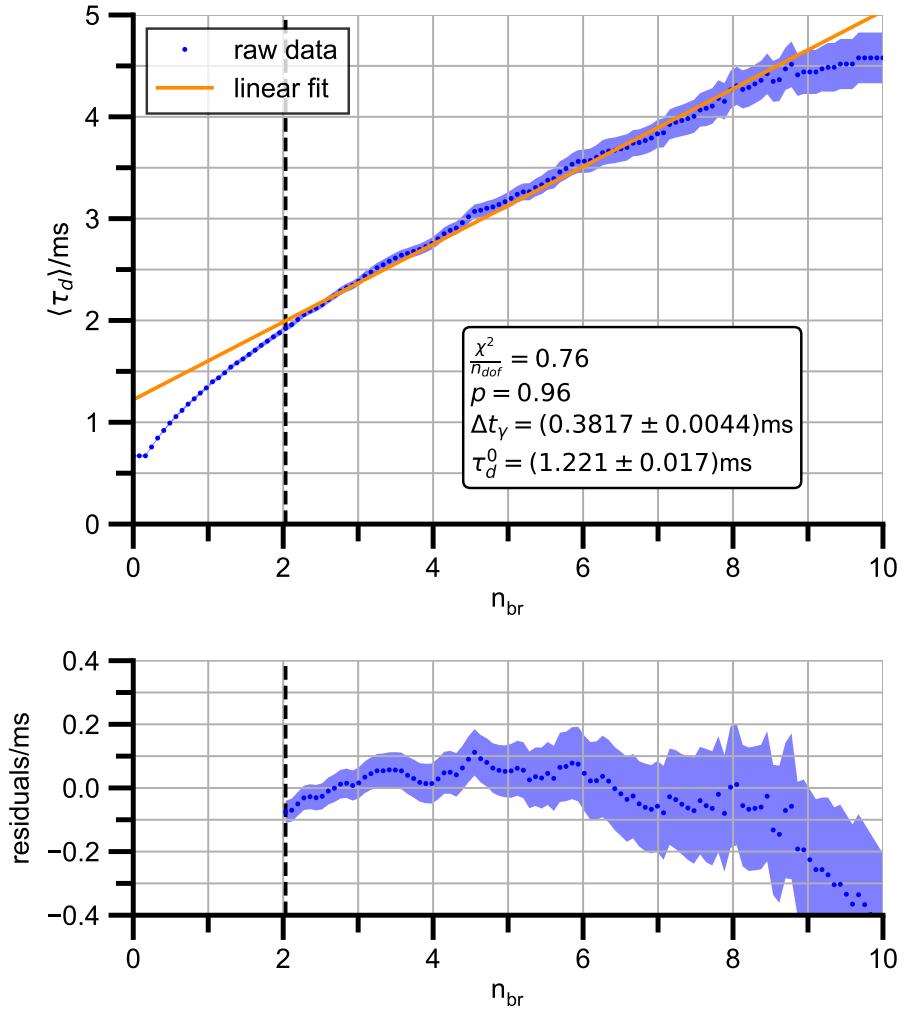


Figure B.1: Linear fit according to  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta t_\gamma \cdot n_{br}$  for the blue channel. Information on the measurement can be found in Section 4.1. The vertical dashed line indicates the range of  $n_{br}$  that was used for fitting. For smaller  $n_{br}$ , the dependency is not linear. For higher  $n_{br}$ , the residual plot reveals a slight deviation of the fitted model from the raw data.

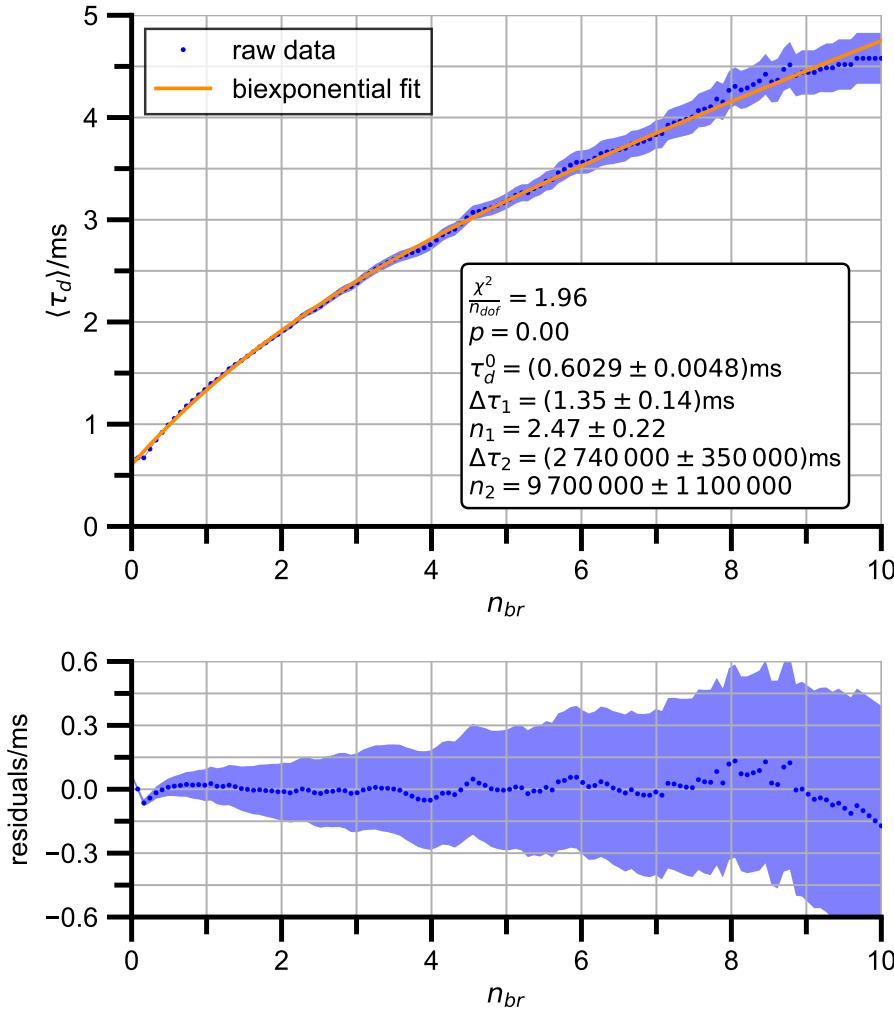


Figure B.2: Biexponential fit according to  $\langle \tau_d \rangle(n_{br}) = \tau_d^0 + \Delta\tau_1(1 - e^{-n_{br}/n_1}) + \Delta\tau_2(1 - e^{-n_{br}/n_2})$  for the blue channel. Information on the measurement can be found in Section 4.1. The biexponential function takes two different saturation processes into account. For the blue channel,  $\tau_2$  and  $n_2$  are notably large. Thus, the second exponential term effectively describes a linear function. The slope is in agreement with the result from the linear fit, see Section 4.2.1.

## B.2 Number of Molecules

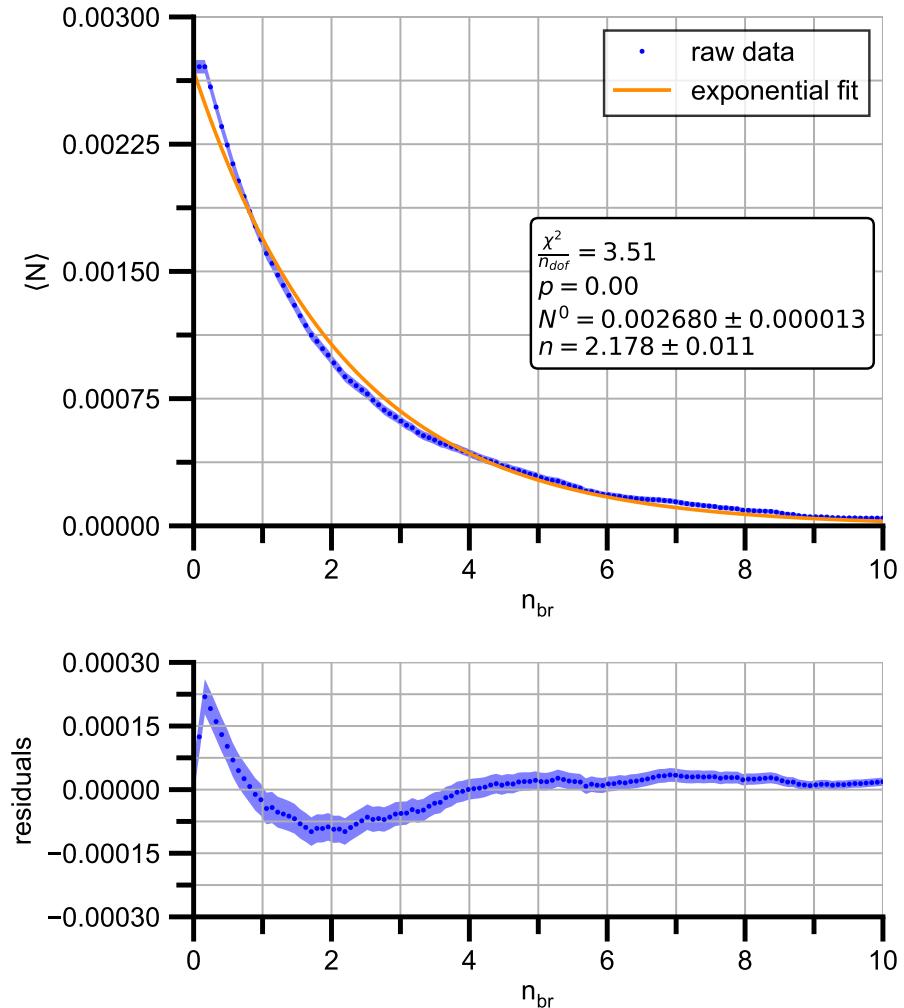


Figure B.3: Exponential fit according to  $\langle N \rangle(n_{br}) = N^0 e^{-n_{br}/n}$  for the blue channel. Information on the measurement can be found in Section 4.1. The fit reflects the general decreasing trend of  $\langle N \rangle$ , but deviates clearly from the specific behavior.

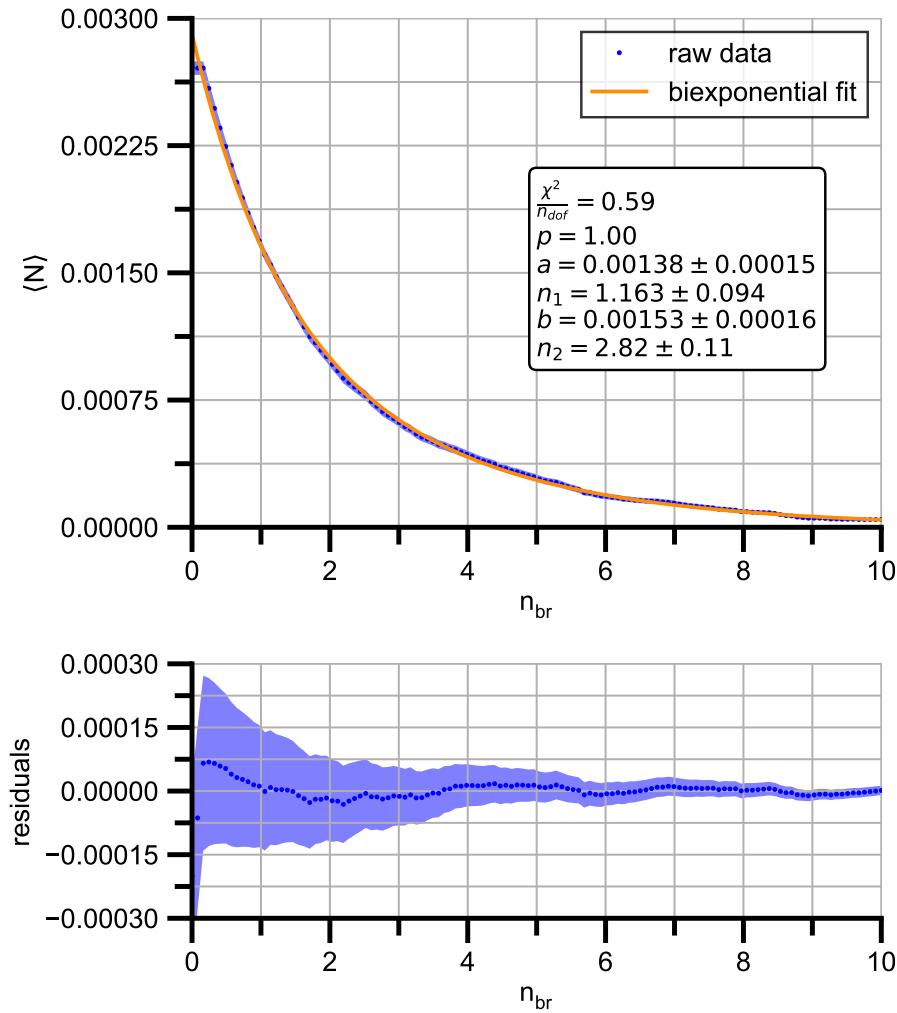


Figure B.4: Biexponential fit according to  $\langle N \rangle(n_{br}) = ae^{-n_{br}/n_1} + be^{-n_{br}/n_2}$  for the blue channel. Information on the measurement can be found in Section 4.1. The two decay processes of the biexponential function modulate the measurement more accurately than the monoexponential model.

### B.3 Molecular Brightness

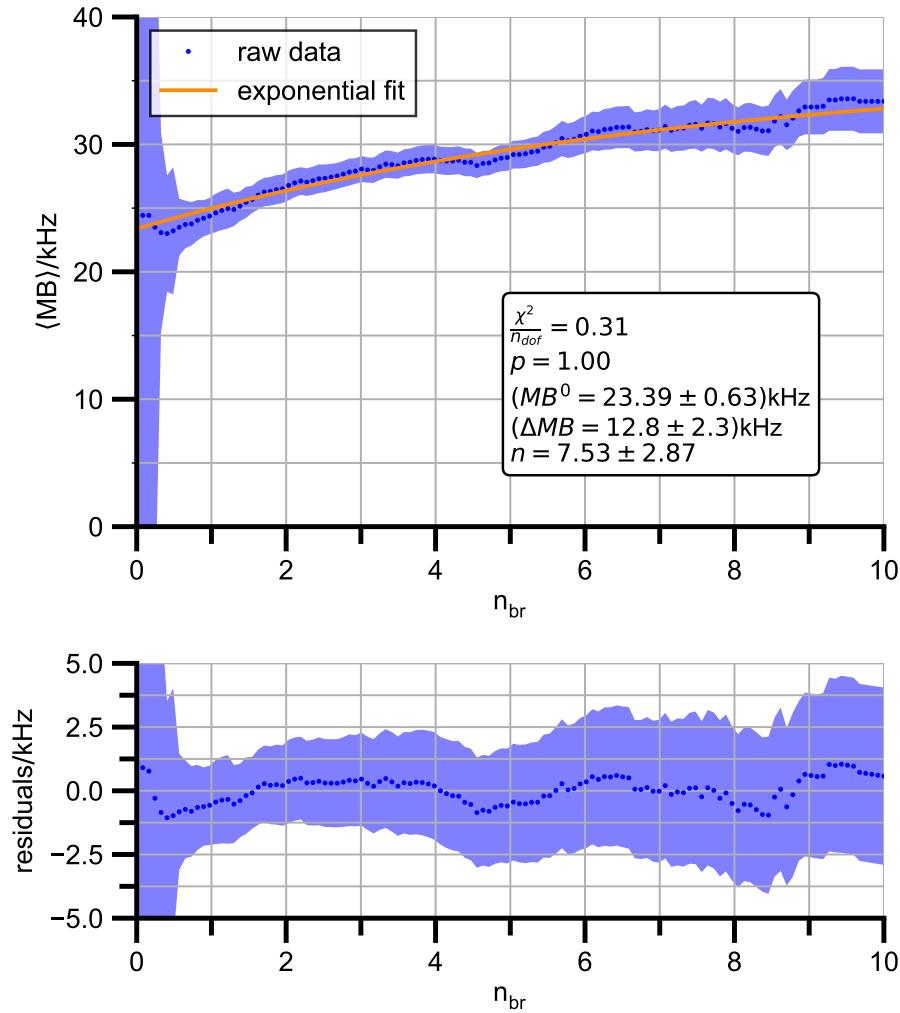


Figure B.5: Exponential fit according to  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB(1 - e^{-n_{br}/n})$  for the blue channel. Information on the measurement can be found in Section 4.1. The model describes the raw data accurately. The comparably large uncertainty on  $\langle MB \rangle$  explains the low value for  $\chi^2/n_{dof}$  for the fits.

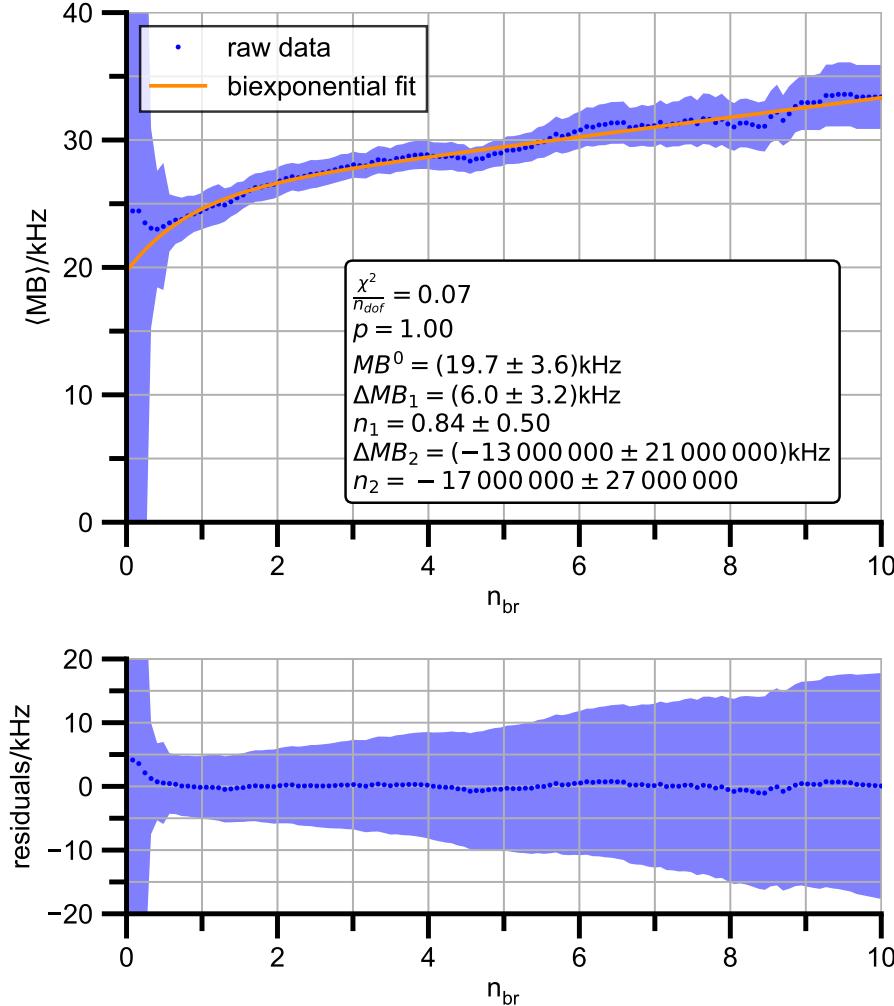


Figure B.6: Fit according to  $\langle MB \rangle(n_{br}) = MB^0 + \Delta MB_1(1 - e^{-n_{br}/n_1}) + \Delta MB_2(1 - e^{-n_{br}/n_2})$  for the blue channel. Information on the measurement can be found in Section 4.1. The parameters of the second exponential term are in agreement with zero. Therefore, for the blue channel, the molecular brightness is effectively described by a monoexponential saturation.

## C Appendix: Optimal Number of Bursts

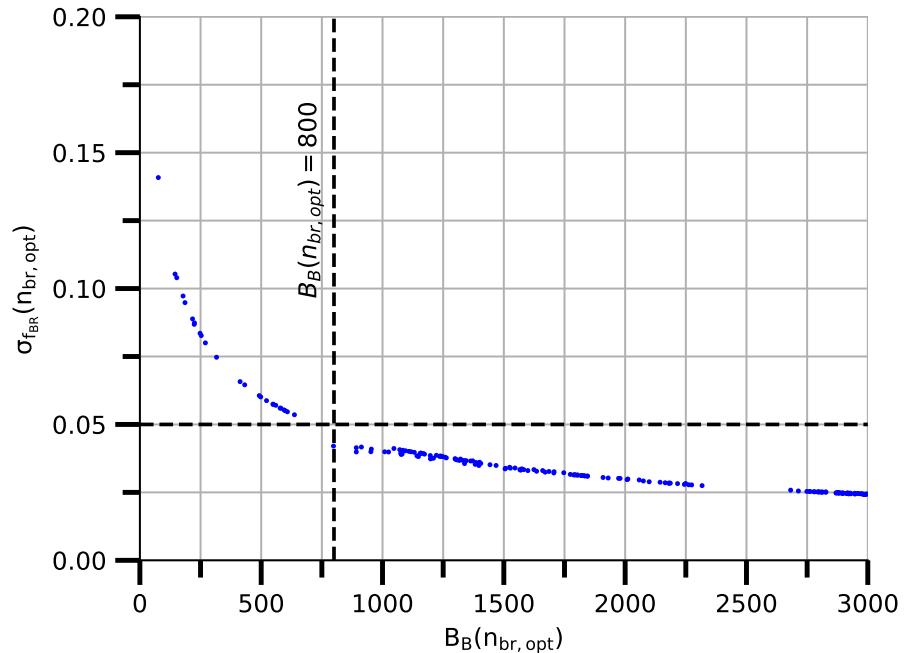


Figure C.1: Uncertainty on the coincidence fraction  $f_{BR}$  at the optimal brightness threshold  $n_{br,opt}$  as a function of the selected number of bursts  $B_B(n_{br,opt})$  for the blue channel. Information on the measurement can be found in Section 5.1. The optimal number of selected bursts, for which the uncertainty falls below 5 %, is approximately  $B_B(n_{br,opt}) = 800$ .

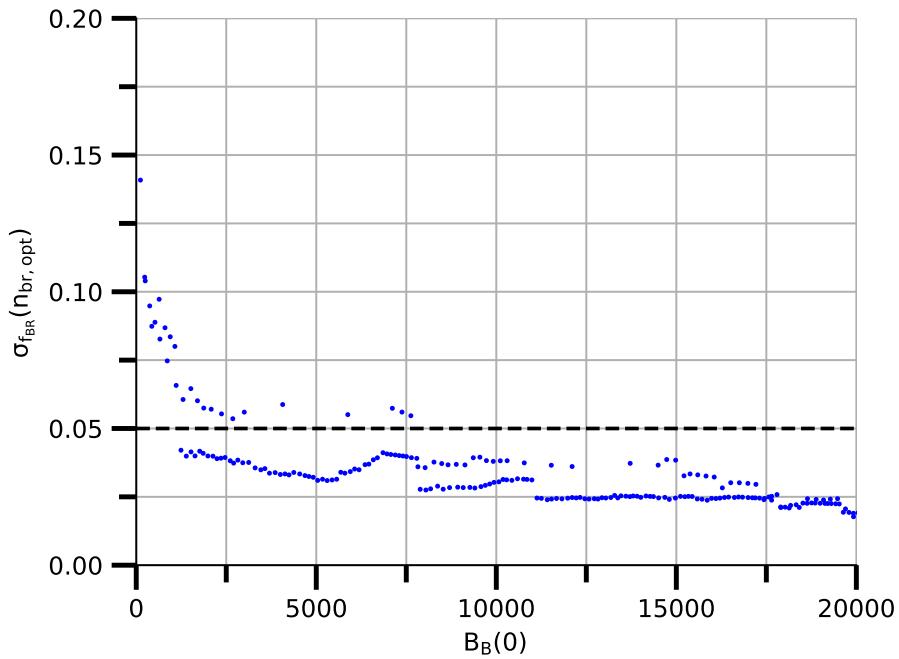


Figure C.2: Uncertainty on the coincidence fraction  $f_{BR}$  at the optimal brightness threshold  $n_{br,opt}$  as a function of the initial number of bursts  $B_B(0)$  for the blue channel. Information on the measurement can be found in Section 5.1. The graph does not reveal a completely deterministic behavior. However, for a number of about 7500 initial bursts, the uncertainty is typically lower than 5 %.

## D Appendix: Chance Coincidence Limitation

### D.1 Application Limits of Chance Coincidence Correction

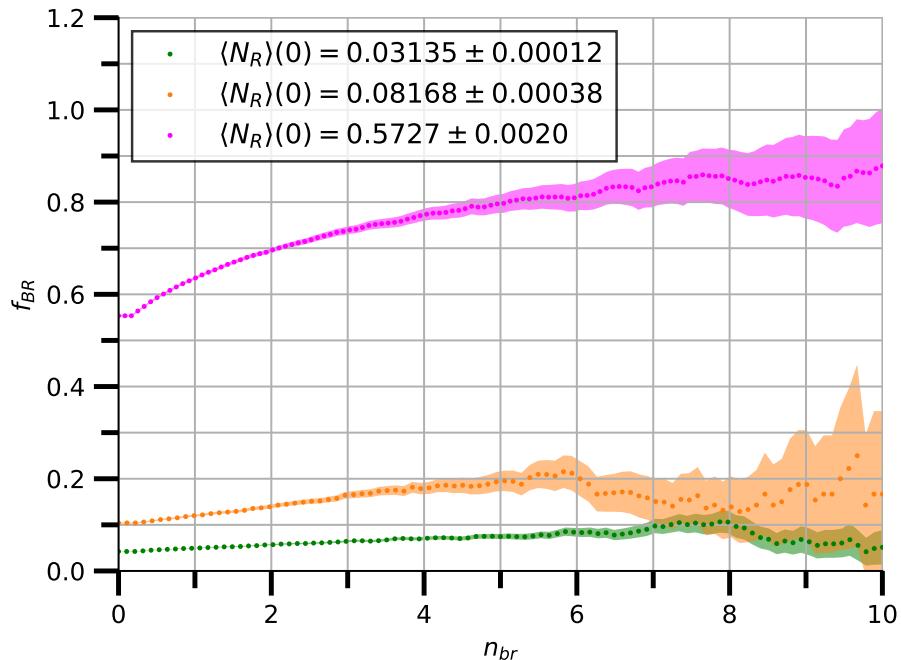


Figure D.1: Coincidence fraction of blue channel  $f_{BR}$  for the mixture of red-labeled dsDNA and free blue dye for a small, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. The observed coincidence fraction is primarily caused by chance coincidences and increases with  $\langle N_R \rangle(0)$  and  $n_{br}$ .

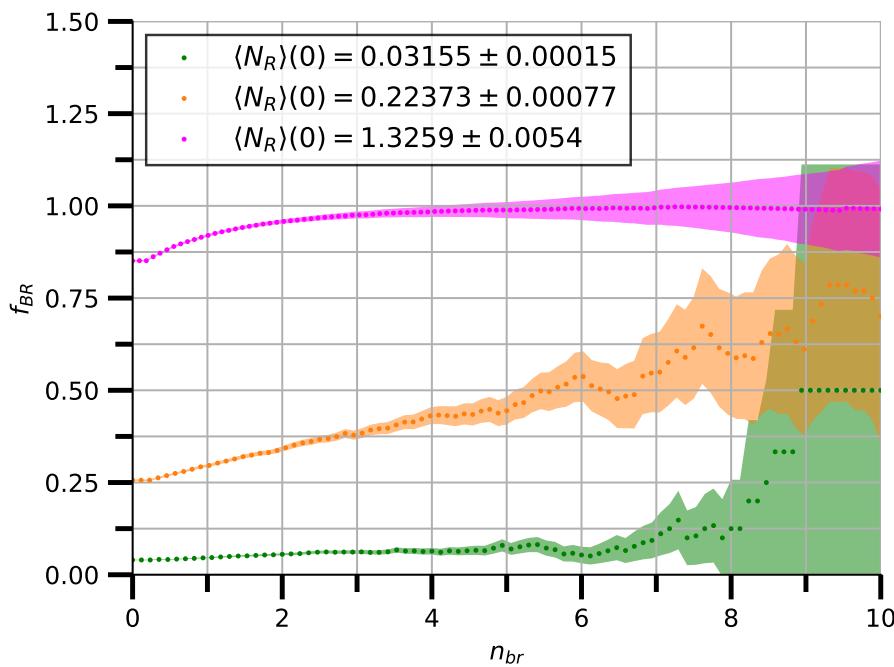


Figure D.2: Coincidence fraction of blue channel  $f_{BR}$  for the mixture of red-labeled ribosomes and free blue dye for a small, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. The observed coincidence fraction is primarily caused by chance coincidences and increases with  $\langle N_R \rangle(0)$  and  $n_{br}$ .

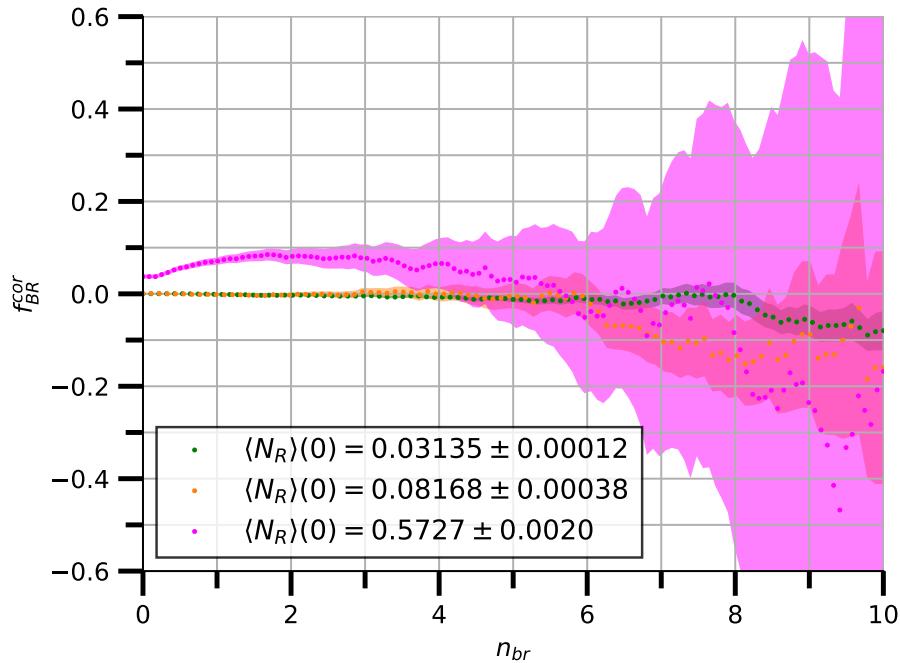


Figure D.3: Corrected coincidence fraction of blue channel  $f_{BR}^{cor}$  for the mixture of red-labeled dsDNA and free blue dye for a low, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. For the highest molecule number, the corrected coincidence fraction deviates slightly from the expected value of 0 %. However, more pronounced is the uncertainty that - compared to the red channel - becomes extremely large for high brightness thresholds. It is caused by a smaller amount of bursts in the blue channel.

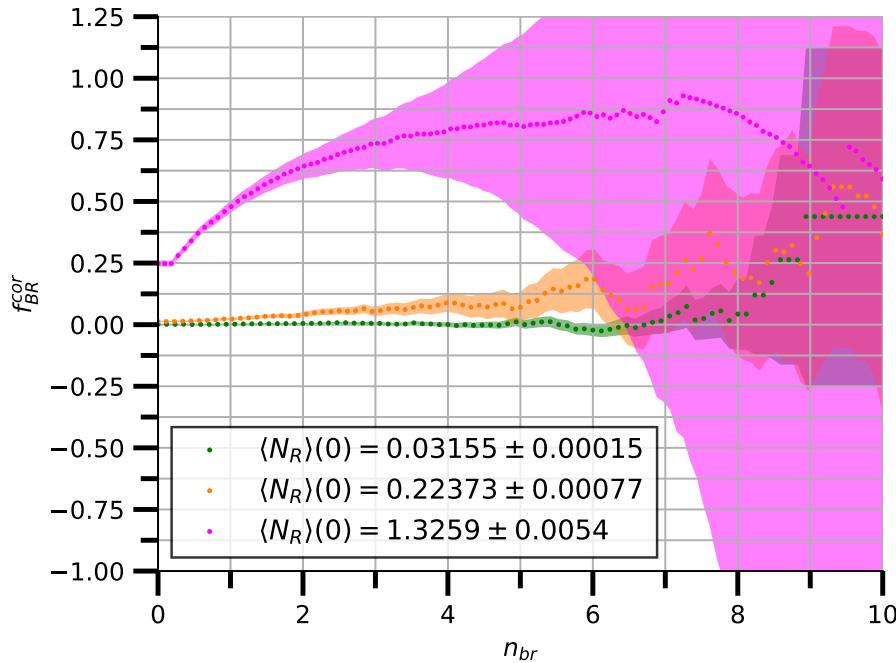


Figure D.4: Corrected coincidence fraction of blue channel  $f_{BR}^{cor}$  for the mixture of red-labeled ribosomes and free blue dye for a low, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. Only for the highest molecule number, the corrected coincidence fraction deviates significantly from the expected value of 0 %. However, more pronounced is the uncertainty that - compared to the red channel - becomes extremely large for high brightness thresholds. It is caused by a smaller amount of bursts in the blue channel.

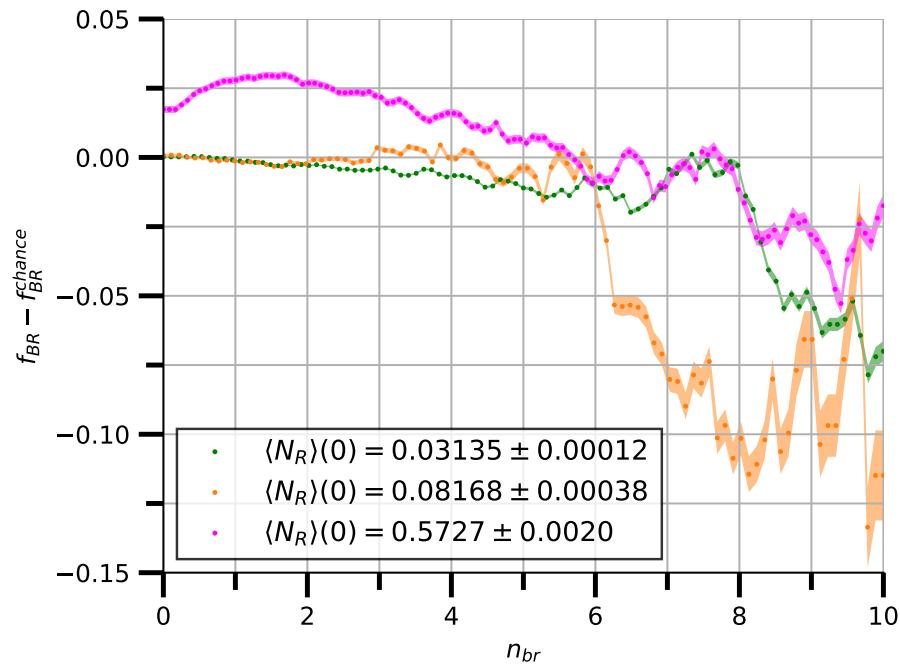


Figure D.5: First term of corrected coincidence fraction of blue channel  $f_{BR} - f_{BR,0}^{chance}$  for the mixture of red-labeled dsDNA and free blue dye for a small, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. For large  $n_{br}$ , the first term deviates clearly from the expected value of 0 %. There, the number of bursts is low and explains the deviation.

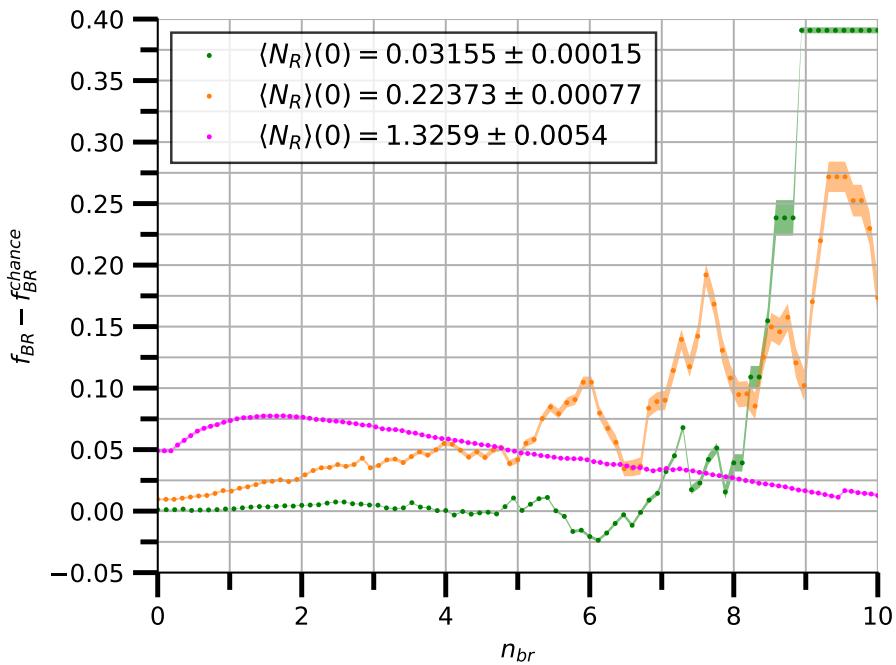


Figure D.6: First term of corrected coincidence fraction of blue channel  $f_{BR} - f_{BR,0}^{chance}$  for the mixture of red-labeled ribosomes and free blue dye for a small, intermediate, and high molecule number. Information on the measurement can be found in Section 6.1.1. For large  $n_{br}$ , the first term deviates clearly from the expected value of 0 %. There, the number of bursts is low and explains the deviation.

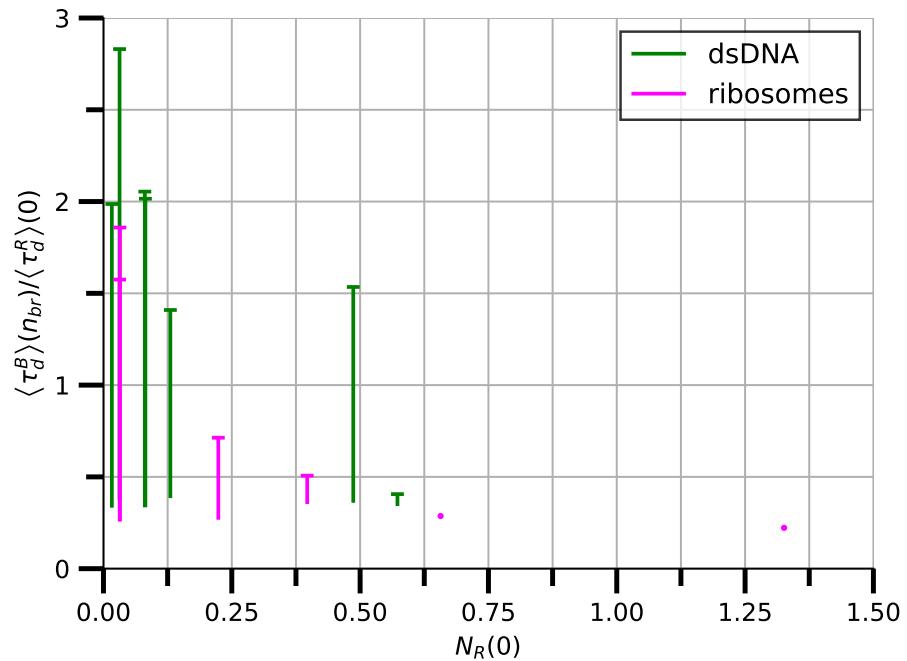


Figure D.7: Application range in terms of dwell time ratio as a function of the molecule number for the blue channel. Information on the measurements with dsDNA and ribosomes can be found in Section 6.1.1. Every vertical line represents one measurement. A short horizontal line at the highest dwell time ratio indicates that the deviation of the corrected coincidence fraction or its statistical uncertainty becomes more than 5 % for higher dwell time ratios. A simple dot denotes that the chance coincidence correction is unsuitable for the whole measurement.

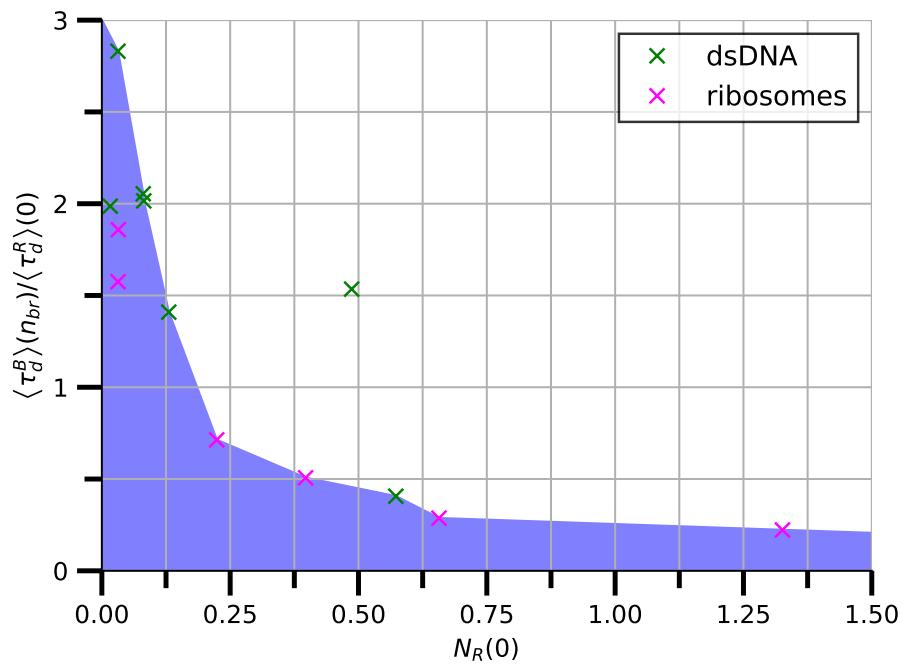


Figure D.8: Interpolated application range in terms of dwell time ratio as a function of the molecule number for the blue channel. Information on the measurements with ds-DNA and ribosomes can be found in Section 6.1.1. The interpolation serves only for illustration and is not derived from theoretical knowledge. It allows to determine the application range for an arbitrary molecule number.

## D.2 Validation of Application Limits

$n_{br,opt}$	$\langle N_R \rangle(0) [10^{-3}]$	$\langle \tau_d^B \rangle(n_{br,opt}) / \langle \tau_d^R \rangle(0)$	$f_{BR}(n_{br,opt}) [\%]$	$f_{BR}^{cor}(n_{br,opt}) [\%]$
1.30	$18.47 \pm 0.11$	$1.398 \pm 0.011$	$84.2 \pm 2.1$	$83.5 \pm 2.2$
3.52	$26.54 \pm 0.13$	$1.807 \pm 0.011$	$85.6 \pm 2.9$	$84.5 \pm 3.2$
1.95	$94.56 \pm 0.56$	$1.4415 \pm 0.0091$	$90.6 \pm 2.2$	$88.2 \pm 2.7$
2.00	$126.83 \pm 0.56$	$1.4124 \pm 0.0083$	$91.5 \pm 2.0$	$88.4 \pm 2.7$
0.87	$621.7 \pm 2.0$	$1.3982 \pm 0.0043$	$97.44 \pm 0.64$	$88.6 \pm 2.9$
3.72	$777.8 \pm 2.8$	$2.6371 \pm 0.0057$	$99.6 \pm 1.7$	$93 \pm 28$

Table D.1: Corrected coincidence fraction  $f_{BR}^{cor}$  of blue channel at the optimal brightness threshold  $n_{br,opt}$  for measurements with a mixture of red-labeled dsDNA and dual-labeled dsDNA. For every measurement row, the molecule number, the relevant dwell time ratio, and the uncorrected coincidence fraction is stated. Information on the measurements can be found in Section 6.2.1. The first measurement row is taken as a guess for the true binding fraction. The corrected coincidence fractions increases systematically with the molecule number.

$\langle N_R \rangle(0) [10^{-3}]$	$\langle \tau_d^B \rangle(3.52) / \langle \tau_d^R \rangle(0)$	$f_{BR}(3.52) [\%]$	$f_{BR}^{cor}(3.52) [\%]$
$18.47 \pm 0.11$	$2.313 \pm 0.015$	$84.5 \pm 4.8$	$83.5 \pm 5.2$
$26.54 \pm 0.13$	$1.807 \pm 0.011$	$85.6 \pm 2.9$	$84.5 \pm 3.2$
$94.56 \pm 0.56$	$1.944 \pm 0.011$	$93.3 \pm 3.7$	$91.2 \pm 4.9$
$126.83 \pm 0.56$	$1.966 \pm 0.011$	$93.6 \pm 3.5$	$90.6 \pm 5.1$
$621.7 \pm 2.0$	$2.7692 \pm 0.0058$	$99.2 \pm 1.6$	$92 \pm 17$
$777.8 \pm 2.8$	$2.5580 \pm 0.0056$	$99.5 \pm 1.6$	$92 \pm 25$

Table D.2: Corrected coincidence fraction  $f_{BR}^{cor}$  of blue channel at  $n_{br} = 3.53$  for measurements with a mixture of red-labeled dsDNA and dual-labeled dsDNA. For every measurement row, the molecule number, the relevant dwell time ratio, and the uncorrected coincidence fraction is stated. Information on the measurements can be found in Section 6.2.1. The first measurement row is taken as a guess for the true binding fraction. The corrected coincidence fractions increases systematically with the molecule number, and for the last two measurement rows the uncertainty becomes extremely large.

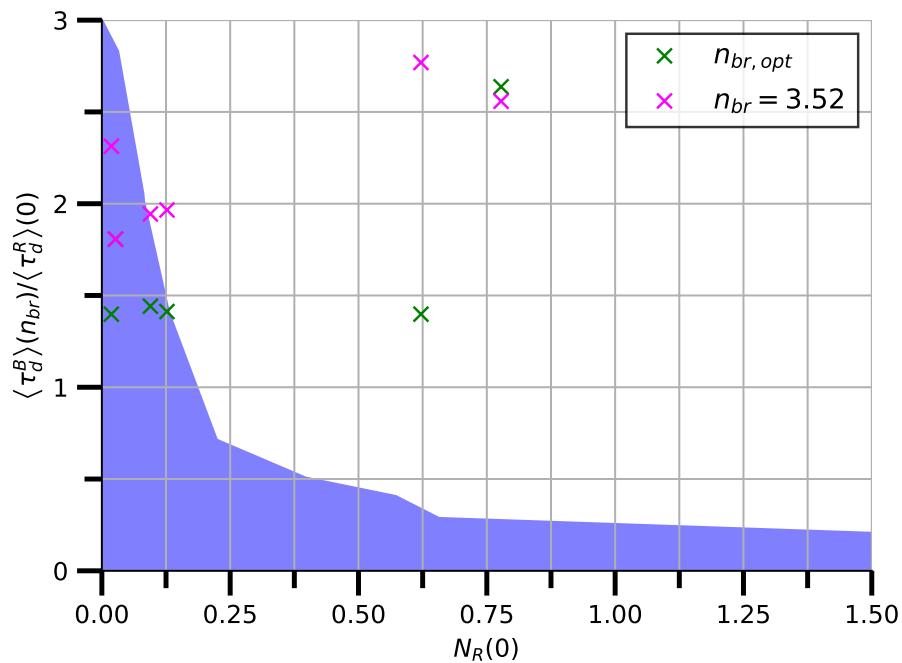


Figure D.9: Position of the validation experiments for the optimal brightness threshold (see Table D.1) and for  $n_{br} = 3.52$  (see Table D.2) in the illustration of the application range of the chance coincidence fraction for the blue channel. Only the first two measurements lie inside the application range. For the other measurement rows, the corrected coincidence fraction and its uncertainty increases systematically.