# Modeling electrodynamics in the vicinity of metal nanostructures

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsprogramm ProPhys

der Georg-August University School of Science (GAUSS)

vorgelegt von

**Daja Talina Helga Wilhelmine Ruhlandt**

aus Walsrode

Göttingen, 2018

**Betreuungsausschuss:**

Prof. Dr. Jörg Enderlein
Drittes Physikalisches Institut,
Georg-August-Universität Göttingen

Prof. Dr. Claus Ropers
Drittes Physikalisches Institut,
Georg-August-Universität Göttingen

**Mitglieder der Prüfungskommission:**

Referent:       Prof. Dr. Jörg Enderlein
                Drittes Physikalisches Institut,
                Georg-August-Universität Göttingen

Korreferent:    Prof. Dr. Claus Ropers
                Viertes Physikalisches Institut,
                Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

                Dr. Alexander Egner
                Laser-Laboratorium Göttingen e.V.

                Prof. Dr. Andreas Janshoff
                Institut für Physikalische Chemie,
                Georg-August-Universität Göttingen

                Dr. Florian Rehfeldt
                Drittes Physikalisches Institut,
                Georg-August-Universität Göttingen

                Prof. Dr. Tim Salditt
                Institut für Röntgenphysik,
                Georg-August-Universität Göttingen

**Tag der mündlichen Prüfung:**

18. Dezember 2018

# Abstract

In this work, we show how a profound theoretical understanding of a seemingly exotic phenomenon, namely the influence of metal structures on fluorescence lifetimes, has lead to the development of a microscopy technique with outstanding $z$-resolution that is based on a completely different physical concept than other optical super-resolution methods. The first part of this work is dedicated to classical electrodynamics. In particular, the radiation properties of oscillating electric dipoles are derived in both planar and spherical geometries. An interpretation of these results in quantum-mechanical terms is then used to quantitatively describe the change of fluorescence lifetimes in the vicinity of metal nanostructures. The power of this semi-classical approach is demonstrated in the second part of this work, where metal-induced energy transfer (MIET) microscopy is employed to study several biological systems, and the fluorescence quantum yield of dyes in different environments is determined via lifetime measurements in a nanocavity.

# Affidavit

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

# List of related publications

T. Baronsky, **D. Ruhlandt**, B. R. Brückner, J. Schäfer, N. Karedla, S. Isbaner, D. Hähnel, I. Gregor, J. Enderlein, A. Janshoff, and A. I. Chizhik, "Cell-substrate dynamics of the epithelial-to-mesenchymal transition," *Nano Letters*, vol. 17, no. 5, pp. 3320–3326, 2017.

F. Schneider, **D. Ruhlandt**, I. Gregor, J. Enderlein, and A. I. Chizhik, "Quantum yield measurements of fluorophores in lipid bilayers using a plasmonic nanocavity," *Journal of Physical Chemistry Letters*, vol. 8, no. 7, pp. 1472–1475, 2017.

A. M. Chizhik, **D. Ruhlandt**, J. Pfaff, N. Karedla, A. I. Chizhik, I. Gregor, R. H. Kehlenbach, and J. Enderlein, "Three-dimensional reconstruction of nuclear envelope architecture using dual-color metal-induced energy transfer imaging," *ACS Nano*, vol. 11, no. 12, pp. 11839–11846, 2017.

A. M. Chizhik, C. Wollnik, **D. Ruhlandt**, N. Karedla, A. I. Chizhik, L. Hauke, D. Hähnel, I. Gregor, J. Enderlein, and F. Rehfeldt, "Dual-color metal-induced and Förster resonance energy transfer for cell nanoscopy," *Molecular Biology of the Cell*, vol. 29, no. 7, pp. 846–851, 2018.

L. Tarpani, **D. Ruhlandt**, L. Latterini, D. Hähnel, I. Gregor, J. Enderlein, and A. I. Chizhik, "Photoactivation of luminescent centers in single $SiO_2$ nanoparticles," *Nano Letters*, vol. 16, no. 7, pp. 4312–4316, 2016.

S. Isbaner, N. Karedla, **D. Ruhlandt**, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein, "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging," *Optics Express*, vol. 24, no. 9, pp. 9429–9445, 2016.

N. Karedla, **D. Ruhlandt**, A. M. Chizhik, J. Enderlein, and A. I. Chizhik, *Advanced Photon Counting*, ch. Metal-Induced Energy Transfer, pp. 265–281. Springer, 2014.

S. Isbaner, N. Karedla, I. Kaminska, **D. Ruhlandt**, M. Raab, J. Bohlen, A. I. Chizhik, I. Gregor, P. Tinnefeld, J. Enderlein, and R. Tsukanov, "Axial colocalization of single molecules with nanometer accuracy using metal-induced energy transfer," *Nano Letters*, vol. 18, no. 4, pp. 2616–2622, 2018.

N. Karedla, A. M. Chizhik, S. C. Stein, **D. Ruhlandt**, I. Gregor, A. I. Chizhik, and J. Enderlein, "Three-dimensional single-molecule localization with nanometer accuracy using metal-induced energy transfer (MIET) imaging," *The Journal of Chemical Physics*, vol. 148, no. 20, p. 204201, 2018.

# Glossary of acronyms

ADR          -  angular distribution of radiation
dcMIET       -  dual-colour MIET
DOPE         -  1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine
DMSO         -  dimethyl sulfoxide
DOPC         -  1,2-dioleoyl-*sn*-glycero-3-phosphocholine
DPPE         -  1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine
DTF          -  distributed tail-fit
FLIM         -  fluorescence lifetime imaging microcopy
FRET         -  Förster resoncance energy transfer
hMSC         -  human mesenchymal stem cell
INM          -  inner nuclear membrane
IRF          -  instrument response function
MIET         -  metal-induced energy transfer
nMuMG        -  a specific mouse mammary gland epithelial cell line
NPC          -  nuclear pore complex
ONM          -  outer nuclear membrane
PAINT        -  points accumulation for imaging in nanoscale topography
PALM         -  photoactivated localization microscopy
PBS          -  phosphate-buffered saline
PSF          -  point spread function
PW           -  plane wave
QY           -  quantum yield
RI           -  refractive index
SMLM         -  single-molecule localization microscopy
SSR          -  sum of squared residuals
STORM        -  stochastic optical reconstruction microscopy
STED         -  stimulated emission depletion
VA-TIRFM     -  variable-angle total internal reflection fluorescence microscopy
VSH          -  vector spherical harmonic

# Contents

# Contents

# 1 Introduction

Without doubt, light is one of the most fascinating natural phenomena. Sight is our most important sense, and light is what lets us discover the world – from the macroscopic realm of the stars via our everyday world down to the microcosmos with its myriad of stunning inhabitants. It is therefore not surprising that humans have long been striving to understand light and to apply it to a variety of purposes. Over the course of the last two centuries, significant progress has been made in both fields. The description of light as electromagnetic waves, together with a comprehensive theory of these waves, has enabled the modeling of light and light sources in different situations, including the radiation properties of single luminescent molecules. On the experimental side, one of the most important applications is light microscopy. Continuous advances have pushed the size of the smallest observable structures down to a few nanometers, allowing scientists to gain insights into the components that make up living beings. Here, where physics, biology and medicine overlap, numerous exciting questions are waiting to be answered.

This work aims to explain the electrodynamic framework used to model the properties of single-molecule emitters, and to demonstrate a range of applications in the context of optical microscopy. In order to introduce the reader to this vast field, a brief history of optics and electromagnetics is given, demonstrating their joint importance for modern microscopy in general and the techniques presented in this work in particular. This is followed by a summary of state-of-the-art optical super-resolution techniques to provide context for the applications presented later. A short outline of the structure of this work concludes this chapter.

Some of the earliest known theories about the nature of light are more than 2500 years old, dating back to ancient Greece and ancient India. The topic was studied with renewed vigour in the Age of Enlightenment, resulting in the formulation of two competing theories: Christiaan Huygens believed light was a wave [10], while Issac Newton suggested that light consisted of a stream of fast particles [11]. These theories co-existed for roughly one hundred years, until a series of experiments provided strong arguments for the wave nature of light. These experiments, which were conducted and interpreted by Thomas Young, François Arago, Augustin-Jean Fresnel and others, included interference (the ability of waves to locally amplify or cancel each other) and diffraction (the appearance of distinct patterns also in the geometrical shadow when a wave encounters an obstacle). Arguably the most important contribution came from James Clerk Maxwell. In 1862, he was able to link optics to another emerging field of physics: *electrodynamics.*

While both magnetic and electric phenomena had been known for centuries, it was only in 1820 that Hans Christian Oersted discovered that electric currents can move a magnetic needle [12]. This finding signaled the birth of electromagnetics, a field which subsequently attracted much attention, for example by André-Marie Ampère, Michael

Faraday and William Thomson (see e.g. [13] for a historical overview). Inspired by these works, Maxwell formulated a set of equations which provided a relation between static and dynamic electric and magnetic fields, as well as electric charges and currents, and which are – in a slightly modified form – today known as *Maxwell's equations*. From these equations, Maxwell derived transverse waves that propagate with a finite velocity, which he found to be very similar to experimentally obtained values for the speed of light. He concluded that "light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena" [14]. The existence of these electromagnetic waves in air was shown by Heinrich Hertz in 1888, seemingly verifying Maxwell's theory. In the same theoretical framework, Hertz was able to predict the electric field of an oscillating dipole, a model that can explain such diverse systems as television antennas and single light-emitting molecules. This framework is the basis on which all modern calculations of electromagnetic fields are performed, including the work presented here.

However, there were also some experiments which contradicted Maxwell's theory. The most prominent of these is the photoelectric effect, where it was found that the maximum velocity of electrons emitted from an illuminated material depends on the wavelength rather than the intensity of the light. Another mystery that could not be explained classically was *Stoke's law of fluorescence*. A fluorescent substance can absorb light of one wavelength, followed by emission of light of a different wavelength. Stoke's law states that the emitted light always has a longer wavelength than the absorbed light. In 1905, Albert Einstein explained both effects by suggesting that light consists of discrete wave packets, nowadays called *photons*, whose energy $E$ can only taken on discrete values that are determined by the wavelength $\lambda$ of the light, $E = h/\lambda$ with a constant $h$ [15][1]. In 1921, Einstein was awarded the Nobel Prize in Physics for this work. Today, the apparent contradiction of Maxwell's and Einstein's explanations is called the *wave-particle duality*: Neither of the two concepts "wave" or "particle" can explain all effects by itself because light displays characteristics of both.

Parallel to the advances in electrodynamics, the study of optical phenomena was furthered by the improvement of optical instruments, such as optical microscopes. However, as Ernst Abbe proved in 1873, even a perfect microscope cannot resolve illuminated objects that are closer than a certain minimal distance [17], which is called the *diffraction limit*. The same is true for small light sources, such as fluorescent molecules, whose image can never be an infinitesimal point but is always spread out in space. The three-dimensional intensity distribution that is generated by imaging a point-source through an optical microscope is called the instrument's *point spread function* (PSF). For a standard widefield microscope, the diameter of the PSF in the plane perpendicular to the optical axis is close to the wavelength of the light. Typically, the axial resolution is by approximately a factor of two worse than the lateral resolution. Only recently, methods have been invented that can overcome the diffraction limit for some specialized sample types. These include scattering samples (e.g. interferometric scattering microscopy, [18]), reflecting samples (e.g. reflection interference constrast microscopy, [19]) and fluorescent samples. The latter sample type has found the most

---

[1] A concept that had been foreshadowed by Max Planck's description of black-body radiation, where he postulated that the energy of oscillators in a black body is quantized. Nevertheless, Planck himself was a critic of the lightquantum hypothesis [16].

widespread application due to the high specificity with which fluorescent labels can be attached to structures of interest, and because it is possible to image labels fluorescing at different wavelengths within the same sample. The importance of this field was widely acknowledged when Eric Betzig, Stefan W. Hell and W. E. Moerner were awarded the Nobel Prize in Chemistry in 2014 for "the development of super-resolution fluorescence miscoscopy" [20].

The various super-resolution fluorescence microscopy techniques can be classified according to the basic principle that allows them to circumvent the diffraction limit. One class of techniques uses sophisticated methods to define the spatial distribution of possibly excited fluorophores. Members of this class include stimulated emission depletion (STED, [21, 22]), where the excited state of all molecules except those in a very small volume is depleted by stimulated emission, thus determining where detected fluorescence is originating, and variable-angle total internal reflection fluorescence microscopy (VA-TIRFM, e.g. [23, 24]), where axial super-resolution ($10\,\mathrm{nm}$ to $20\,\mathrm{nm}$) is achieved by varying the excitation intensity along the optical axis and comparing the measured fluorescence intensity with theoretical models. A second class consists of interferometric techniques in which the sample is illuminated and/or observed from both sides simultaneously using two opposing lenses (4Pi-microscopy and I$^5$M, see e.g. [25] for a comparison). Finally, a third class of techniques can be summarized as single-molecule localization microscopy (SMLM). SMLM is based on the fact that the center of the PSF – and thus the position of the emitter – can be determined with much higher accuracy than the diameter of the PSF, provided that only one emitter is active within a diffraction limited spot during the acquisition time of an image. Thus, in order to study a densely labeled sample, sparse subsets of the labels have to be activated, imaged and subsequently deactivated. This can be achieved in a variety of ways, resulting in the existence of many SMLM methods, such as photoactivated localization microscopy (PALM, [26]), stochastic optical reconstruction microscopy (STORM, [27]), points accumulation for imaging in nanoscale topography (PAINT, [28]), or direct STORM (dSTORM, [29]), to name just some.

All of these methods have their specific advantages and disadvantages, and are therefore used side by side. Originally, PALM, STORM, dSTORM and PAINT only provided an increased resolution perpendicular to the optical axis, while STED improved the resolution along the optical axis, too, from approximately $500\,\mathrm{nm}$ to $100\,\mathrm{nm}$ [22], and was later improved to yield an almost isotropic resolution of $20\,\mathrm{nm}$ to $30\,\mathrm{nm}$ (isoSTED, [30]). However, already two years after their initial invention, SMLM techniques were extended to the third dimension. This was achieved by designing a PSF that is asymmetric along the optical axis (astigmatic imaging with $50\,\mathrm{nm}$ to $60\,\mathrm{nm}$ axial resolution [31], helical wavefront shaping with $20\,\mathrm{nm}$ axial resolution [32]), by sampling different parts of the PSF simultaneously (biplane imaging, $75\,\mathrm{nm}$ [33]), or by combining SMLM with principles from the other classes of super-resolution methods (interferometric PALM, $10\,\mathrm{nm}$ to $20\,\mathrm{nm}$ [34], 4Pi-STORM, $3\,\mathrm{nm}$ to $7\,\mathrm{nm}$ [35]).

Most of the three-dimensional SMLM techniques mentioned so far obtain estimates of the position of a fluorophore by measuring a two-dimensional intensity distribution and fitting it to a known model. An alternative approach is to use different physical observables for localization in the lateral and axial directions, namely the fluorescence intensity for the former and the *fluorescence lifetime* for the latter. The fluorescence

lifetime, also called excited state lifetime or simply lifetime, is the average time delay between the absorption of a photon and the emission of fluorescence. As Edward Purcell predicted in 1946 [36] and Karl-Heinz Drexhage showed experimentally in 1970 [37], the lifetime of a fluorophore is influenced by the presence of metals. Perhaps surprisingly, this intrinsically quantum mechanical phenomenon can be described in a precise quantitative manner by modeling the fluorophore as a classical oscillating electric dipole and mapping classical to quantum mechanical quantities. Based on this realization, *metal-induced energy transfer* (MIET) microscopy was devised, which allows to obtain an axial resolution in the order of 2 nm to 4 nm [38–40]. Besides providing axial super-resolution, metal-induced lifetime changes can also be used to determine the fluorescence quantum yield of an emitter in a calibration-free measurement that does not require any reference sample. The advancement of these two fruitful applications of a seemingly obscure discovery from the 1970s to the thriving field of super-resolution fluorescence microscopy was the main driving force behind this work.

This work is organized as follows: The next chapter introduces the relevant theory, starting from Maxwell's equations, presenting solutions using different basis functions that are beneficial for different sample geometries, and deriving a number of quantities for an oscillating dipole emitter. These derivations are accompanied by simulated results, for example for the total energy emitted by an oscillating dipole close to various nanostructures. Subsequently, the mechanisms governing fluorescence are explained, followed by an introduction of possible techniques to measure fluorescence lifetimes and a presentation of the theory of MIET. The third chapter focuses on the numerical implementation of the theoretical concepts on a computer system, which entails the approximation of analytical formulas by numerical algorithms as well as possible numerical errors and the convergence behaviour of certain quantities. The fourth chapter consists of two parts treating different applications of the concepts introduced in the first chapters, namely MIET and quantum yield measurements using a nanocavity. In both cases, specific experiments are supplemented by general considerations of factors determining the accuracy of the respective technique. In the fifth chapter, the results of this work are summarized and discussed. Finally, the appendix contains additional figures, calculations, code excerpts, as well as a short overview of further projects which were a part of this work but which would have gone beyond the scope of the main text.

# 2 Theory

Modeling electrodynamics in the vicinity of metal nanostructures would not be possible without a profound knowledge of the behaviour of electromagnetic fields in a variety of situations. In practice, taking into account all the details of a system is often impossible. For example, treating the interaction of an electric field with every single electron of a metallic nanostructure is not feasible. Instead, the properties of the metal are contained in macroscopic, averaged quantities such as the refractive index. Furthermore, we neglect the fact that "plane surfaces" always have some finite surface roughness, that "spherical particles" are never perfect spheres, and that an "infinite halfspace" usually does end somewhere. Within this idealized framework, an astonishing number of effects can be modeled. The good agreement between theory and experiment that is shown in chapter 4 proves that these simplifications allow to explain experimental results to a satisfying degree of accuracy.

This chapter starts by introducing the fundamental equations on which all further derivations are based. Different types of solutions that are advantageous for certain geometries are presented in section 2.1, together with ways to switch between these representations. Section 2.2 introduces the electric dipole emitter, the most important energy source for this work. In the subsequent section 2.3, the influence of nanostructures on electric fields is derived. This also has an impact on the energy transported by the electromagnetic field, as will be shown in section 2.4. Finally, in section 2.5, classical electromagnetics and some quantum mechanical aspects are combined to describe fluorescence, a phenomenon that allows many interesting applications. Well-known facts have been kept short, instead, more attention is given to aspects that are found less often in standard textbooks. While this chapter presents the relevant theory, concrete numerical implementations of some of the more complicated expressions can be found in chapter 3. Since many relations contain infinite sums and integrals, convergence is an important practical aspect. This, too, is discussed in the chapter on numerical implementation.

## 2.1 The wave equation and its solutions in different coordinate systems

Many textbooks treat the fundamentals of optics and their applications, for example [41–43], to name just a few. Most of these books use the International System of Units (SI units), since it is the statutory unit system for commerce, administration, and also measuring instruments, and therefore specifies the units of quantities measured in the lab – be it an intensity in watt per square meter, or a current in ampere. However, this system was only introduced in 1960, therefore all older texts still use the centimetre-gram-second (cgs) unit system. Nowadays, many theoreticians working in the field of optics still prefer the cgs system since the relevant equations are symmetric (electric

and magnetic field have the same units) and the somewhat artifical constants $\varepsilon_0$ and $\mu_0$ (vacuum permittivity and permability, respectively) are not needed anymore. The only non-dimensionless constant appearing in the context of electromagnetics in the cgs system is the speed of light in vacuum, which seems intuitive considering that we want to describe light. Using a different unit system changes the numerical values, but not the general structure of the solutions. Furthermore, for the techniques explained in this work, it is often sufficient to calculate relative changes in electric fields for different situations. Therefore, cgs units are used throughout this work.

### 2.1.1 The wave equation

In a medium with relative permittivity $\varepsilon$ and relative permeability $\mu$, charge density $\rho$ and electric current density $\boldsymbol{j}$, the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ obey *Maxwell's equations*:

$$\nabla \cdot (\varepsilon \boldsymbol{E}) = 4\pi\rho \tag{2.1}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{2.2}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \left(\frac{\boldsymbol{B}}{\mu}\right) = \frac{1}{c}\frac{\partial (\varepsilon \boldsymbol{E})}{\partial t} + \frac{4\pi}{c}\boldsymbol{j} \ . \tag{2.4}$$

Here, $c$ is the speed of light in vacuum. Taking the curl of equation (2.3) leads to

$$\nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} = \nabla \times (\nabla \times \boldsymbol{E}) = -\frac{1}{c}\frac{\partial(\nabla \times \boldsymbol{B})}{\partial t} \ . \tag{2.5}$$

If the medium is isotropic, homogeneous and linear, that is if $\varepsilon$ and $\mu$ are scalars, spatially invariant and independent of the magnitudes of $\boldsymbol{E}$ and $\boldsymbol{B}$, they can be placed before the derivatives. In the absence of free charges or currents ($\rho = 0$, $\boldsymbol{j} = 0$), substituting equations (2.1) and (2.4) on the left and right side, respectively, results in the *wave equation*

$$\nabla^2 \boldsymbol{E} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0 \ . \tag{2.6}$$

By starting with equation (2.4) instead, an identical equation can be derived for $\boldsymbol{B}$:

$$\nabla^2 \boldsymbol{B} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2 \boldsymbol{B}}{\partial t^2} = 0 \ . \tag{2.7}$$

It is convenient to define the *refractive index* $n$ of the medium via

$$n := \sqrt{\varepsilon\mu} \ . \tag{2.8}$$

Then, the speed of light in the medium is given by $c_n := c/n$. In practice, we only work with materials where $\mu$ is very close to one, which is why $\mu$ is omitted in the following. Since both equations (2.6) and (2.7) are linear in $\boldsymbol{E}$ and $\boldsymbol{B}$, any linear combination of

solutions is a solution, too. Therefore, it is useful to find sets of solutions $\{\boldsymbol{f}_i(\boldsymbol{r}, t)\}$ that fulfill the following criteria:

1. The $\{\boldsymbol{f}_i(\boldsymbol{r}, t)\}$ are a complete set of orthogonal functions, and can thus be used to express any possible solution of the wave equation.

2. For certain geometries, the $\{\boldsymbol{f}_i(\boldsymbol{r}, t)\}$ facilitate calculations of interactions of the field with boundaries between different materials.

In the following, we will introduce two such sets of solutions, *plane waves* and *vector spherical harmonics*, and explain in which situations they can be applied profitably.

### 2.1.2 Plane waves

Plane waves are defined as

$$\boldsymbol{A}(\boldsymbol{r}, t) = \boldsymbol{A}_0 \, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t}, \tag{2.9}$$

where the amplitude $\boldsymbol{A}_0$ defines both the *magnitude* $|\boldsymbol{A}_0|$ and the *polarization* $\boldsymbol{A}_0/|\boldsymbol{A}_0|$ of the field, and $\boldsymbol{A}$ can be either the electric field $\boldsymbol{E}$ or the magnetic field $\boldsymbol{B}$. The frequency $f$ of the field's oscillation at a fixed point is related to the *angular frequency* $\omega$ via $\omega = 2\pi f$. Both $\omega$ and $f$ are constant in all media, including vacuum. The vector $\boldsymbol{k}$ is called the *wave vector*, it points in the direction of propagation, while its magnitude is related to the wavelength $\lambda$ via $k = |\boldsymbol{k}| = 2\pi/\lambda$. Finally, the speed of light in the medium connects the other quantities by $c_n = f\lambda = \omega/k$.

Inserting this solution into Maxwell's equations in the absence of free charges and currents leads to

$$\boldsymbol{k} \cdot \boldsymbol{E}_0 = 0 \, , \qquad \boldsymbol{k} \times \boldsymbol{E}_0 = \phantom{-}\frac{\omega}{c}\boldsymbol{B}_0 \, ,$$

$$\boldsymbol{k} \cdot \boldsymbol{B}_0 = 0 \, , \qquad \boldsymbol{k} \times \boldsymbol{B}_0 = -\frac{n^2\omega}{c}\boldsymbol{E}_0 \, . \tag{2.10}$$

Thus, both $\boldsymbol{E}_0$ and $\boldsymbol{B}_0$ are perpendicular to the wave vector $\boldsymbol{k}$ and to each other. Since $k = \omega/c_n = n\omega/c$, the two equalities in the right column imply that the magnitudes of $\boldsymbol{E}_0$ and $\boldsymbol{B}_0$ fulfill $B_0 = nE_0$.

To sum up, plane waves are characterized by infinitely extended planes of constant phase which propagate along the wave vector $\boldsymbol{k}$. At any fixed point, electric and magnetic field oscillate perpendicular to $\boldsymbol{k}$ and to each other. Together, these oscillating fields are called *electromagnetic waves*.

### 2.1.3 Vector spherical harmonics

Samples consisting of nanospheres have a spherical geometry. Plane waves, which, as the name suggests, have a planar symmetry, are not the best choice for this situation. Instead, we are now interested in a set of solutions with spherical symmetries. While plane waves are known to almost everyone in the field of optics, this second set of solutions is less common and will thus be treated in more detail. A number of approaches

to introducing vector spherical harmonics (VSH) can be found in the literature; we base this section on [44], chapter 4. We start by demanding that these functions, like plane waves, give monochromatic solutions, i.e. we assume that the electric field has the form $\boldsymbol{E}(\boldsymbol{r})\exp(-i\omega t)$. Inserting this in the wave equation (2.6) leads to the time-independent vector wave equation

$$\nabla^2\boldsymbol{E} + k^2\boldsymbol{E} = 0 \ , \tag{2.11}$$

which is also known as the Helmholtz equation. As an ansatz, suppose we already know a scalar function $\psi(\boldsymbol{r})$ that is a solution to the scalar wave equation

$$\nabla^2\psi(\boldsymbol{r}) + k^2\psi(\boldsymbol{r}) = 0 \tag{2.12}$$

and use this to define

$$\boldsymbol{M}(\boldsymbol{r}) := \nabla\times(\boldsymbol{r}\psi) = \psi\nabla\times\boldsymbol{r} - \boldsymbol{r}\times\nabla\psi = -\boldsymbol{r}\times\nabla\psi \ . \tag{2.13}$$

It now remains to show that $\boldsymbol{M}$ is a solution to the time-independent vector wave equation (2.11). We find the relation

$$\begin{aligned}
\nabla\times\boldsymbol{M} &= \nabla\times(\nabla\times\boldsymbol{r}\psi) \\
&= \nabla(\nabla\cdot\boldsymbol{r}\psi) - \nabla^2(\boldsymbol{r}\psi) \\
&= \nabla(\nabla\cdot\boldsymbol{r}\psi) - \boldsymbol{r}\nabla^2\psi - 2\nabla\psi \ .
\end{aligned} \tag{2.14}$$

Since the divergence of the curl of a vector is zero, $\nabla\cdot\boldsymbol{M} = 0$. This leads to

$$\begin{aligned}
\nabla^2\boldsymbol{M} &= \nabla(\nabla\cdot\boldsymbol{M}) - \nabla\times(\nabla\times\boldsymbol{M}) \\
&= -\nabla\times(\nabla\times\boldsymbol{M}) \\
&\stackrel{(2.14)}{=} -\nabla\times\left[\nabla(\nabla\cdot\boldsymbol{r}\psi) - \boldsymbol{r}\nabla^2\psi - 2\nabla\psi\right] \\
&= \nabla\times(\boldsymbol{r}\nabla^2\psi),
\end{aligned} \tag{2.15}$$

where the last step used the fact that the curl of the gradient of a differentiable function is zero. Now we exploit that $\psi$ is a solution of the scalar wave equation (2.12) to get

$$\begin{aligned}
\nabla^2\boldsymbol{M} &= -\nabla\times(\boldsymbol{r}k^2\psi) \\
&= -k^2\boldsymbol{M} \ ,
\end{aligned} \tag{2.16}$$

Thus we have found the desired solution to the vector wave equation. We can now define another zero-divergence vector field

$$\boldsymbol{N}(\boldsymbol{r}) := \frac{1}{k}\nabla\times\boldsymbol{M} \ . \tag{2.17}$$

As can be seen from equations (2.15) and (2.16),

$$\nabla\times\boldsymbol{N} = \frac{1}{k}\nabla\times(\nabla\times\boldsymbol{M}) = -\frac{1}{k}\nabla^2\boldsymbol{M} = k\boldsymbol{M} \ . \tag{2.18}$$

## 2.1 The wave equation and its solutions in different coordinate systems

This symmetry will be used many times in this work. For excample, we can use it to show that $\boldsymbol{N}$ is also a solution of the vector wave equation:

$$
\begin{aligned}
\nabla^2 \boldsymbol{N} &= -\nabla \times (\nabla \times \boldsymbol{N}) \\
&= -k \nabla \times \boldsymbol{M} \\
&= -k^2 \boldsymbol{N} \; .
\end{aligned}
\tag{2.19}
$$

However, in order to find concrete formulations for $\boldsymbol{M}$ and $\boldsymbol{N}$, we still have to determine the function $\psi(\boldsymbol{r})$ that solves the scalar wave equation (2.12). As mentioned at the beginning of this section, we are interested in a solution with spherical symmetries. Thus, we express the scalar wave equation (2.12) in spherical coordinates $(r, \theta, \varphi)$:

$$
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2 \psi}{\partial \varphi^2} + k^2 = 0 \; .
\tag{2.20}
$$

We can find a solution of the form

$$
\psi(r, \theta, \varphi) = R(r)\, P(\theta)\, \Phi(\varphi)
\tag{2.21}
$$

by doing the following: First, we demand that

$$
\frac{\partial^2 \Phi}{\partial \varphi^2} + m^2 \Phi = 0 \; ,
\tag{2.22}
$$

which has the solution

$$
\Phi_m(\varphi) = \mathrm{e}^{im\varphi} \; .
\tag{2.23}
$$

In general, $m$ can be any number. However, in any physically meaningful solution of Maxwell's equations, the substitution $\varphi \to \varphi + 2\pi$ must not change the value of $\psi$. Therefore, $m$ has to be an integer (positive or negative). The functions $\Phi_m$ are orthogonal in the sense that

$$
\int_0^{2\pi} \Phi_m(\varphi) \overline{\Phi_{m'}}(\varphi) \mathrm{d}\varphi = \int_0^{2\pi} \mathrm{e}^{i(m-m')\varphi} \mathrm{d}\varphi = 2\pi \delta_{m,m'} \; ,
\tag{2.24}
$$

where the bar means taking the complex conjugate and $\delta_{m,m'}$ is the Kronecker delta. As the next step, we demand that

$$
\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P}{\partial \theta} \right) + \left( \ell(\ell+1) - \frac{m^2}{\sin^2\theta} \right) P = 0 \; .
\tag{2.25}
$$

For given integers $\ell \geq 0$ and $m^2 \leq \ell^2$, this equation is solved by the so-called associated Legendre polynomials $P_\ell^m(\cos\theta)$. For $m \geq 0$, $P_\ell^m(\cos\theta)$ can be defined using Rodrigues' formula for ordinary Legendre polynomials $P_\ell(\cos\theta)$

$$
P_\ell^m(\cos\theta) = (-1)^m \sin^m\theta \frac{\mathrm{d}^m}{\mathrm{d}\cos\theta^m} P_\ell(\cos\theta)
$$

$$= \frac{(-1)^m}{2^\ell \ell!} \sin^m \theta \frac{\mathrm{d}^{\ell+m}}{\mathrm{d}\cos\theta^{\ell+m}}(\cos^2\theta - 1)^\ell \,, \tag{2.26}$$

which immediately shows that $P_\ell^0 = P_\ell$, $P_\ell^1 = \mathrm{d}P_\ell/\mathrm{d}\theta$ and $P_\ell^m = 0$ for $m > \ell$. The polynomial $P_\ell^{-m}$ obtained by substituting $m$ in the above equation by $-m$ obeys[1]

$$P_\ell^{-m}(\cos\theta) = (-1)^m \frac{(\ell-m)!}{(\ell+m)!}P_\ell^m(\cos\theta) \,. \tag{2.27}$$

These functions are orthogonal for fixed $m$ in the sense that ([45], chapter 10, p. 1326)

$$\int_0^\pi P_\ell^m(\cos\theta)P_{\ell'}^m(\cos\theta)\sin(\theta)\mathrm{d}\theta = \frac{2(l+m)!}{(2l+1)(l-m)!}\delta_{\ell,\ell'} \,. \tag{2.28}$$

At this point, we combine both angular functions to get the so-called *spherical harmonics* $Y_{\ell m}(\theta, \varphi)$,

$$\boxed{Y_{\ell m}(\theta, \varphi) := c_{\ell m}P_\ell^m(\cos\theta)\mathrm{e}^{im\varphi} \quad \text{with} \quad c_{\ell m} := \sqrt{\frac{(2\ell+1)(\ell-m)!}{4\pi(\ell+m)!}} \,.} \tag{2.29}$$

The normalization constant $c_{\ell m}$ ensures the orthogonality of the spherical harmonics: Using the orthogonality relations of $\Phi_m$ and $P_\ell^m$ we see that

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta \, \mathrm{d}\theta \, Y_{\ell m}(\theta, \varphi)\overline{Y_{\ell'm'}}(\theta, \varphi) = \delta_{\ell\ell'} \, \delta_{mm'} \,. \tag{2.30}$$

Figure 2.1 illustrates the real part of the first few spherical harmonics. It can be seen that $Y_{\ell m}$ has $\ell$ nodal planes[2], and that for fixed $\varphi$, there are $\ell - |m|$ zero-crossings as $\theta$ is varied from 0 to $\pi$. By inserting our results for $\Phi_m(\varphi)$ and $P_\ell^m(\cos\theta)$ into the ansatz (2.21), equation (2.20) simplifies to

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \left(k^2 - \frac{\ell(\ell+1)}{r^2}\right)R = 0 \,. \tag{2.31}$$

This second order differential equation is solved by two sets of functions, the spherical Bessel and Neumann functions

$$j_\ell(kr) = \sqrt{\frac{\pi}{2kr}}J_{\ell+\frac{1}{2}}(kr) \quad \text{and} \quad n_\ell(kr) = \sqrt{\frac{\pi}{2kr}}N_{\ell+\frac{1}{2}}(kr) \tag{2.32}$$

---

[1] The differential equation satisfied by $P_\ell^m$ is a second order differential equation, therefore the general solution is a linear combination of two linearly independent functions. However, the second class of solutions, the associated Legendre polynomials of the second kind $Q_\ell^m(\cos\theta)$, are divergent at $\theta = \{0, \pi\}$, which our solution to Maxwell's equations should not be. Thus we can assume that $P_\ell^{-m}$ is proportional to $P_\ell^m$. By equating the coefficients of the highest powers of $x$ in $\mathrm{d}^{\ell-m}/\mathrm{d}x^{\ell-m}(x^2-1)^\ell = c(1-x^2)^m\mathrm{d}^{\ell+m}/\mathrm{d}x^{\ell+m}(x^2-1)^l$, one arrives at the formula given above.
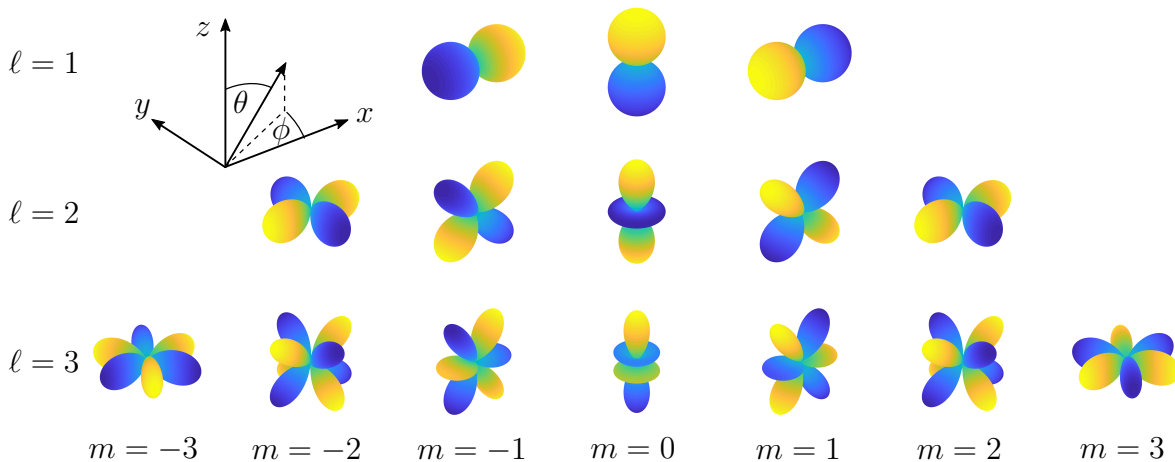[2] Or nodal cones, in the case of $m = 0$.

**Figure 2.1:** Illustration of the real part of the first spherical harmonics $Y_{\ell m}$ for $\ell = 1, 2, 3$ and $m = -\ell, \ldots, \ell$. The distance from the origin in the direction $(\theta, \varphi)$ corresponds to $|Y_{\ell m}(\theta, \varphi)|$, while the colour encodes $Y_{\ell m}(\theta, \varphi)$. All plots have been normalized to $\pm 1$, blue corresponds to negative and yellow to positive values.

with the Bessel and Neumann functions

$$J_\ell(kr) = \sum_{a=0}^{\infty} \frac{(-1)^a}{a! \, \Gamma(a + \ell + 1)} \left(\frac{kr}{2}\right)^{2a+\ell} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(\ell\tau - kr\sin(\tau))} \, d\tau$$

$$\text{and} \quad N_\ell(kr) = \frac{J_\ell(kr)\cos(\ell\pi) - J_{-\ell}(kr)}{\sin(\ell\pi)} \; . \tag{2.33}$$

Any linear combination of these is a solution, too. Two particular linear combinations are the spherical Hankel functions of the first and second kind:

$$h_\ell^{(1)}(kr) = j_\ell(kr) + in_\ell(kr) \quad \text{and} \quad h_\ell^{(2)}(kr) = j_\ell(kr) - in_\ell(kr) \; . \tag{2.34}$$

Later on, we will see that different combinations of these four function sets are needed to fulfill special boundary conditions. A first hint is given by the behaviour of $j_\ell(kr)$ and $y_\ell(kr)$ shown in figure 2.2. For small $kr$, the spherical Bessel function tends to zero, while the spherical Neumann function diverges. Thus, the boundary condition of a finite result at the origin is satisfied by $R(r) = j_\ell(kr)$. On the other hand, for large $kr$, both $j_\ell(kr)$ and $y_\ell(kr)$ are bounded by $\pm(kr)^{-1}$, and oscillate with the cosine and sine of the same argument ([46], §19):

$$j_\ell(kr) \xrightarrow{kr \gg 1} \frac{1}{kr} \cos\left[kr - (\ell + 1)\frac{\pi}{2}\right]$$

$$y_\ell(kr) \xrightarrow{kr \gg 1} \frac{1}{kr} \sin\left[kr - (\ell + 1)\frac{\pi}{2}\right] . \tag{2.35}$$

Therefore, $R(r) = h_\ell^{(1)}(kr)$ has the form of outgoing spherical waves $\propto \exp(ikr)/kr$ for large $kr$. That is why this function is used for the far field of dipole emitters (see section 2.2.2).
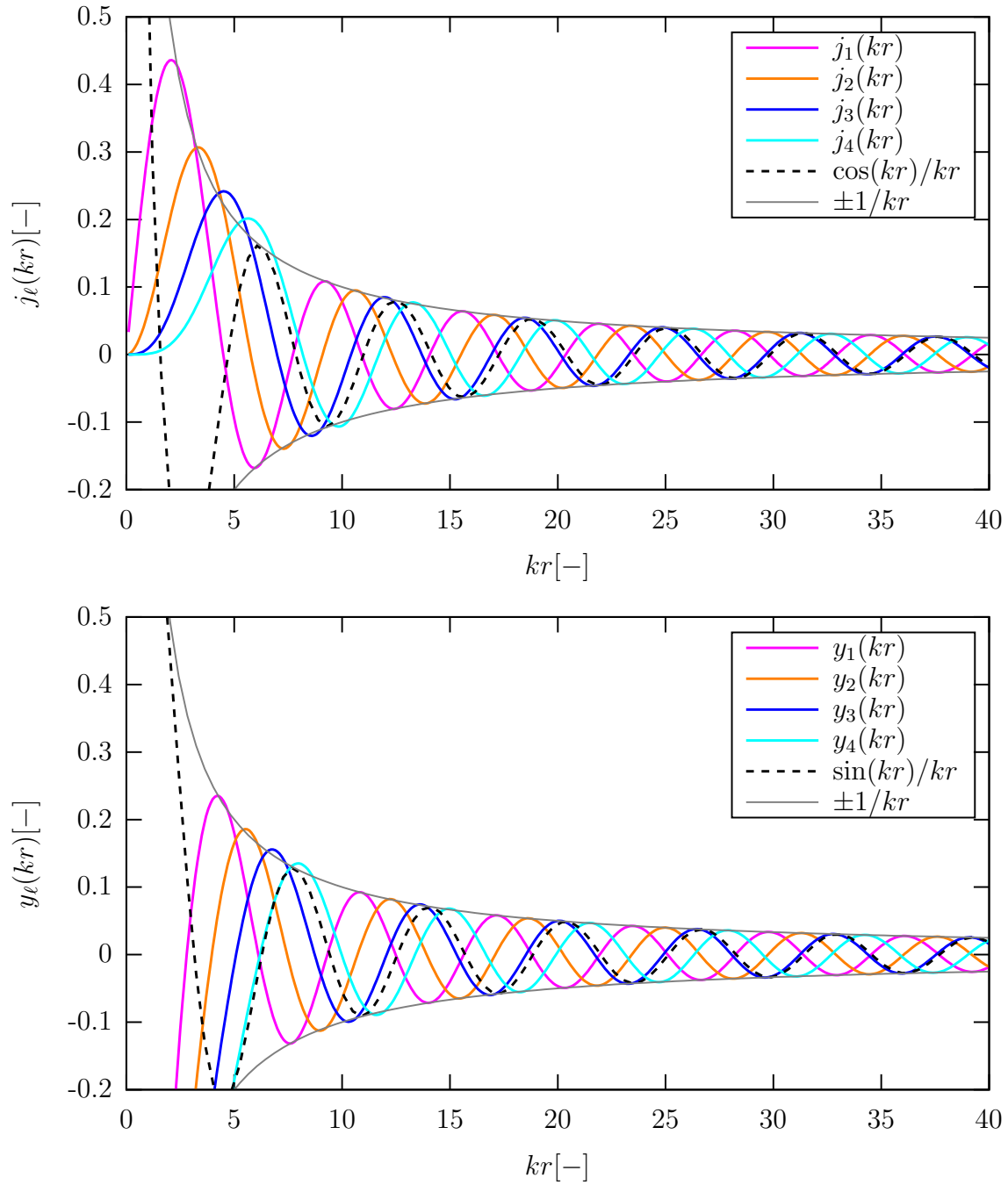
**Figure 2.2:** Spherical Bessel and Neumann functions $j_\ell$ and $y_\ell$ for $\ell = 1 - 4$. For large arguments, the functions converge to $j_\ell(kr) \to \cos[kr - (\ell + 1)\pi/2]/kr$ and $y_\ell(kr) \to \sin(kr - (\ell + 1)\pi/2]/kr$. For small arguments, $j_\ell$ goes to zero, while $y_\ell$ tends to negative infinity.

To sum up, we have found a solution to the scalar wave equation in spherical coordinates (2.20) that has the form

$$\psi_{\ell m}(r, \theta, \varphi) = Y_{\ell m}(\theta, \varphi) f_\ell(kr) \quad \text{with } f_\ell = j_\ell, y_\ell, h_\ell^{(1)}, h_\ell^{(2)} \text{ or linear comb.} \tag{2.36}$$

and that is orthogonal in the sense that

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^\pi \sin\theta \; \mathrm{d}\theta \; \psi_{\ell m}(r, \theta, \varphi) \overline{\psi_{\ell' m'}}(r, \theta, \varphi) = \delta_{\ell\ell'} \delta_{mm'} |f_\ell(kr)|^2 \; . \tag{2.37}$$

We can now explicitly write $\boldsymbol{M}$ and $\boldsymbol{N}$ by inserting the function $\psi_{\ell m}$ in equations (2.13) and (2.17),

$$\boldsymbol{M}(\boldsymbol{r}) = -\boldsymbol{r} \times \nabla \psi \tag{2.13}$$

$$\boldsymbol{N}(\boldsymbol{r}) = \frac{1}{k} \nabla \times \boldsymbol{M} \; . \tag{2.17}$$

To do this, we first need the gradient in spherical coordinates:

$$\nabla \psi = \frac{\partial \psi}{\partial r} \hat{\boldsymbol{e}}_r + \frac{1}{r} \frac{\partial \psi}{\partial \theta} \hat{\boldsymbol{e}}_\theta + \frac{1}{r \sin\theta} \frac{\partial \psi}{\partial \phi} \hat{\boldsymbol{e}}_\varphi \; , \tag{2.38}$$

where the position-dependent unit vectors $\hat{\boldsymbol{e}}_r$, $\hat{\boldsymbol{e}}_\theta$, $\hat{\boldsymbol{e}}_\varphi$ at the point $\boldsymbol{r} = (r, \theta, \varphi)$ are defined as

$$\hat{\boldsymbol{e}}_r(\boldsymbol{r}) = \begin{pmatrix} \sin\theta \cos\varphi \\ \sin\theta \sin\varphi \\ \cos\theta \end{pmatrix}, \; \hat{\boldsymbol{e}}_\theta(\boldsymbol{r}) = \begin{pmatrix} \cos\theta \cos\varphi \\ \cos\theta \sin\varphi \\ -\sin\theta \end{pmatrix}, \; \hat{\boldsymbol{e}}_\varphi(\boldsymbol{r}) = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}, \tag{2.39}$$

see also figure 2.3. With $\boldsymbol{r} = r\hat{\boldsymbol{e}}_r$, $\hat{\boldsymbol{e}}_r \times \hat{\boldsymbol{e}}_r = 0$, $\hat{\boldsymbol{e}}_r \times \hat{\boldsymbol{e}}_\theta = \hat{\boldsymbol{e}}_\varphi$ and $\hat{\boldsymbol{e}}_r \times \hat{\boldsymbol{e}}_\varphi = -\hat{\boldsymbol{e}}_\theta$ we get

$$\boldsymbol{M}_{\ell m}^f(r, \theta, \varphi) = -\boldsymbol{r} \times \nabla \psi_{\ell m}^f = c_{\ell m} \left( \frac{im}{\sin\theta} P_\ell^m \hat{\boldsymbol{e}}_\theta - \frac{\partial P_\ell^m}{\partial\theta} \hat{\boldsymbol{e}}_\varphi \right) f_\ell(kr) \; \mathrm{e}^{im\varphi} \; . \tag{2.40}$$

To find $\boldsymbol{N}$, we use the curl in spherical coordinates:

$$\nabla \times \boldsymbol{M} = \frac{1}{r \sin\theta} \left[ \frac{\partial}{\partial\theta} (M_\varphi \sin\theta) - \frac{\partial M_\theta}{\partial\phi} \right] \hat{\boldsymbol{e}}_r + \left[ \frac{1}{r \sin\theta} \frac{\partial M_r}{\partial\phi} - \frac{1}{r} \frac{\partial}{\partial r}(rM_\varphi) \right] \hat{\boldsymbol{e}}_\theta$$
$$+ \frac{1}{r} \left[ \frac{\partial}{\partial r}(rM_\theta) - \frac{\partial M_r}{\partial\theta} \right] \hat{\boldsymbol{e}}_\varphi. \tag{2.41}$$

This leads to

$$\boldsymbol{N}_{\ell m}^f = \frac{1}{kr} \left( -\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left[ \sin\theta \frac{\partial \psi_{\ell m}}{\partial\theta} \right] + \frac{m^2}{\sin^2\theta} \psi_{\ell m} \right) \hat{\boldsymbol{e}}_r$$
$$+ \frac{1}{kr} \frac{\partial}{\partial r} \left( r \frac{\partial \psi_{\ell m}}{\partial\theta} \right) \hat{\boldsymbol{e}}_\theta \quad + \quad \frac{1}{kr} \frac{\partial}{\partial r} \left( r \frac{im}{\sin\theta} \psi_{\ell m} \right) \hat{\boldsymbol{e}}_\varphi$$
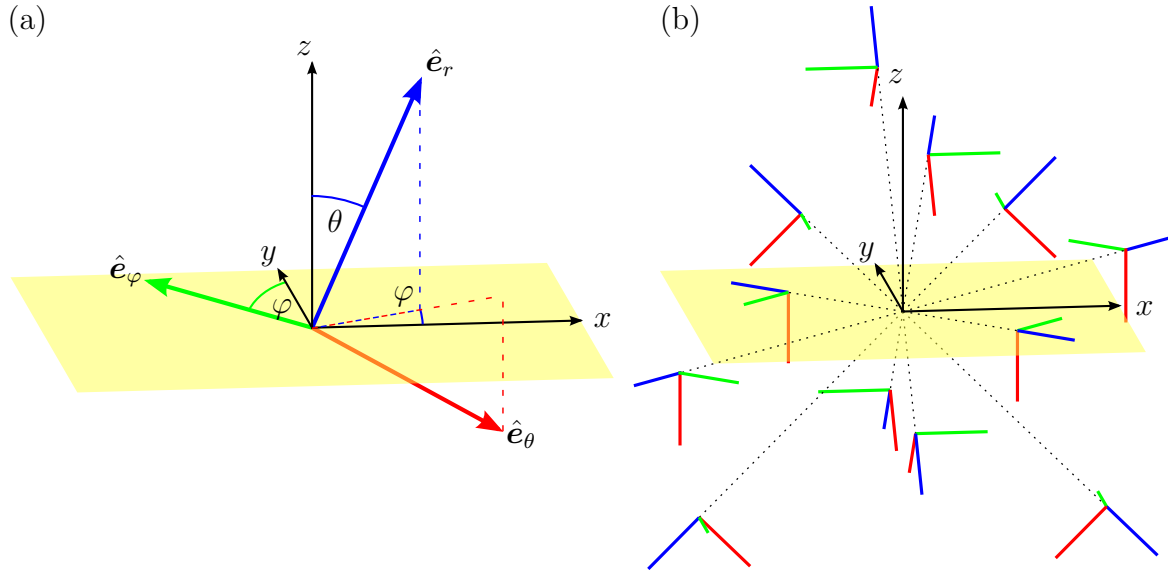
(a)



(b)

Figure 2.3: Unit vectors for spherical coordinates. (a) Definition of $\hat{\boldsymbol{e}}_r$, $\hat{\boldsymbol{e}}_\theta$, $\hat{\boldsymbol{e}}_\varphi$ and the relevant angles $\theta$, $\varphi$. (b) The unit vectors are position-dependent.

$$= \left\{ \ell(\ell+1)\frac{f_\ell}{kr} P_\ell^m \hat{\boldsymbol{e}}_r + \frac{1}{2}\left( f_{\ell-1} + \frac{f_\ell}{kr} - f_{\ell+1} \right) \left[ \frac{\partial P_\ell^m}{\partial \theta} \hat{\boldsymbol{e}}_\theta + \frac{im}{\sin\theta} P_\ell^m \hat{\boldsymbol{e}}_\varphi \right] \right\} c_{\ell m} e^{im\varphi}, \tag{2.42}$$

where we used the fact that $P_\ell^m$ satisfies equation (2.25) and inserted the recurrence relation ([45], chapter 11, p. 1574)

$$\frac{\partial f_\ell(\xi)}{\partial \xi} = \frac{1}{2}\left( f_{\ell-1}(\xi) - \frac{f_\ell(\xi)}{\xi} - f_{\ell+1}(\xi) \right)$$

$$\Rightarrow \frac{\partial}{\partial r}\left[ r f_\ell(kr) \right] = \frac{kr}{2}\left( f_{\ell-1}(kr) + \frac{f_\ell(kr)}{kr} - f_{\ell+1}(kr) \right). \tag{2.43}$$

Since the expressions for $\boldsymbol{M}_{\ell m}^f$ and $\boldsymbol{N}_{\ell m}^f$ are rather long, it seems reasonable to introduce some shorthand notation. As will become obvious in the section on the numerical implementation (see 3.3), it is convenient to define

$$\begin{aligned} \pi_{\ell m}(\theta) &:= c_{\ell m} \frac{m}{\sin\theta} P_\ell^m(\cos\theta) \ , \\ \tau_{\ell m}(\theta) &:= c_{\ell m} \frac{\mathrm{d}P_\ell^m(\cos\theta)}{\mathrm{d}\theta} = -c_{\ell m}\sin\theta \frac{\partial P_\ell^m(\cos\theta)}{\partial\cos\theta} \ , \\ \text{and} \quad \gamma_{\ell m}(\theta) &:= c_{\ell m}\ell(\ell+1)P_\ell^m(\cos\theta) \ . \end{aligned} \tag{2.44}$$

With this, we have

$$
\boldsymbol{M}_{\ell m}^{f}(kr,\theta,\varphi) = \left(i\pi_{\ell m}(\theta)\hat{\boldsymbol{e}}_{\theta} - \tau_{\ell m}(\theta)\hat{\boldsymbol{e}}_{\varphi}\right) f_{\ell}(kr)\mathrm{e}^{im\varphi} \tag{2.45}
$$

$$
\text{and} \quad \boldsymbol{N}_{\ell m}^{f}(kr,\theta,\varphi) = \left[\left\{\tau_{\ell m}(\theta)\hat{\boldsymbol{e}}_{\theta} + i\pi_{\ell m}(\theta)\hat{\boldsymbol{e}}_{\varphi}\right\}\frac{1}{kr}\frac{\mathrm{d}[rf_{\ell}(kr)]}{\mathrm{d}r}\right.
$$

$$
\left. +\gamma_{\ell m}(\theta)\hat{\boldsymbol{e}}_{r}\frac{f_{\ell}(kr)}{kr}\right]\mathrm{e}^{im\varphi}. \tag{2.46}
$$

We started this chapter by claiming that we were going to introduce complete orthogonal function systems for the vector space of solutions of the wave equation. The proof of the orthogonality of $\boldsymbol{M}_{\ell m}$ and $\boldsymbol{N}_{\ell m}$ is quite lengthy, it can be found in section 6.1.1 in the appendix. The main results of the calculations are:

$$
\int_{0}^{2\pi}\mathrm{d}\varphi\int_{0}^{\pi}\mathrm{d}\theta\ \boldsymbol{M}_{\ell m}^{f}\cdot\overline{\boldsymbol{N}_{\ell'm'}^{g}}\sin\theta = 0\ , \tag{2.47}
$$

$$
\int_{0}^{2\pi}\mathrm{d}\varphi\int_{0}^{\pi}\mathrm{d}\theta\ \boldsymbol{M}_{\ell m}^{f}\cdot\overline{\boldsymbol{M}_{\ell'm'}^{g}}\ \sin\theta = \delta_{l,l'}\ \delta_{m,m'}\ \ell(\ell+1)\ f_{\ell}(kr)\overline{g}_{\ell}(kr)\ , \tag{2.48}
$$

$$
\text{and}\ \int_{0}^{2\pi}\mathrm{d}\varphi\int_{0}^{\pi}\mathrm{d}\theta\ \boldsymbol{N}_{\ell m}^{f}\cdot\overline{\boldsymbol{N}_{\ell'm'}^{g}}\sin\theta = \delta_{\ell,\ell'}\ \delta_{m,m'}\ \frac{1}{(kr)^{2}}
$$

$$
\cdot\left\{\ell^{2}(\ell+1)^{2}f_{\ell}\overline{g_{\ell}} + \ell(\ell+1)\frac{\partial}{\partial r}(rf_{\ell})\frac{\partial}{\partial r}\overline{(rg_{\ell})}\right\}\ . \tag{2.49}
$$

Contrarily to what we demanded at the beginning of the chapter, the function set $\boldsymbol{M}_{\ell m}^{f}, \boldsymbol{N}_{\ell m}^{f}$ is actually not complete on the space of all three-dimensional complex-valued functions. It is, however, complete on the subspace of functions with zero divergence [45]. Since we assume homogeneous media without presence of any surface charges, the first of Maxwell's equations is $\nabla\cdot\boldsymbol{E} = 0$ and thus the electrical field is an element of this subspace. Therefore, we can decompose $\boldsymbol{E}$ into VSH.

### 2.1.4 Conversion between PW and VSH

Now that both plane waves and vector spherical harmonics have been introduced, we want to show that they can be converted into each other. This will become important when experimental situations combine both planar and spherical elements. One can show ([47], p. 416) that

$$
\boldsymbol{M}_{\ell m}^{j}(\boldsymbol{r}) = \frac{1}{4\pi i^{\ell-1}}\int_{0}^{\pi}\mathrm{d}\theta'\int_{0}^{2\pi}\mathrm{d}\varphi'\sin\theta'\left[\pi_{\ell m}(\theta')\hat{\boldsymbol{e}}_{p} + i\tau_{\ell m}(\theta')\hat{\boldsymbol{e}}_{s}\right]\mathrm{e}^{im\varphi'+i\boldsymbol{k'}\cdot\boldsymbol{r}}\ , \tag{2.50}
$$

$$
\boldsymbol{N}_{\ell m}^{j}(\boldsymbol{r}) = \frac{1}{4\pi i^{\ell-1}}\int_{0}^{\pi}\mathrm{d}\theta'\int_{0}^{2\pi}\mathrm{d}\varphi'\sin\theta'\left[\tau_{\ell m}(\theta')\hat{\boldsymbol{e}}_{p} + i\pi_{\ell m}(\theta')\hat{\boldsymbol{e}}_{s}\right]\mathrm{e}^{im\varphi'+i\boldsymbol{k'}\cdot\boldsymbol{r}}\ . \tag{2.51}
$$

The unit vectors $\hat{\boldsymbol{e}}_p$ and $\hat{\boldsymbol{e}}_s$ correspond to so-called $p$- and $s$-waves, plane waves with a particular polarization that will be described in more detail in section 2.3.1. For now, it suffices to know the definitions

$$\hat{\boldsymbol{e}}_p := \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} \quad \text{and} \quad \hat{\boldsymbol{e}}_s := \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}. \tag{2.52}$$

The fomulas (2.50) and (2.51) are conveniently evaluated for $\boldsymbol{r}$ in cylindrical coordinates, $\boldsymbol{r} = (\rho\cos\varphi, \rho\sin\varphi, z)$, then $\boldsymbol{k}'\cdot\boldsymbol{r} = k'\sin\theta'\rho\cos(\varphi'-\varphi)+k'\cos\theta'z$. The $\varphi'$-integration can be carried out analytically with the help of the definition of the Bessel function [46],

$$J_\ell(\xi) = \frac{1}{2\pi} \int_a^{a+2\pi} \mathrm{d}\varphi' \mathrm{e}^{i\xi\cos\varphi'} \mathrm{e}^{i\ell(\varphi'-\frac{\pi}{2})} \qquad \text{with arbitrary } a, \tag{2.53}$$

which also implies that:

$$\int_0^{2\pi} \mathrm{d}\varphi' \mathrm{e}^{im\varphi'} \mathrm{e}^{i\xi\cos(\varphi'-\varphi)} = \int_{-\varphi}^{2\pi-\varphi} \mathrm{d}z \mathrm{e}^{im(z+\varphi)} \mathrm{e}^{i\xi\cos(z)}$$
$$= 2\pi J_m(\xi) \mathrm{e}^{im\varphi} i^m . \tag{2.54}$$

With the abbreviation $\xi_m := i^m J_m(k'\sin\theta'\rho)\exp(im\varphi)$, that leads to the relations

$$\boldsymbol{p}_m(\theta',\rho,\varphi) := \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}\varphi' \hat{\boldsymbol{e}}_p \mathrm{e}^{im\varphi'+ik'\sin\theta'\rho\cos(\varphi'-\varphi)} = \begin{pmatrix} \cos\theta'(\xi_{m+1}+\xi_{m-1}) \\ -i\cos\theta'(\xi_{m+1}-\xi_{m-1}) \\ -2\sin\theta'\xi_m \end{pmatrix},$$

$$\boldsymbol{s}_m(\theta',\rho,\varphi) := \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}\varphi' \hat{\boldsymbol{e}}_s \mathrm{e}^{im\varphi'+ik'\sin\theta'\rho\cos(\varphi'-\varphi)} = \begin{pmatrix} i(\xi_{m+1}-\xi_{m-1}) \\ \xi_{m+1}+\xi_{m-1} \\ 0 \end{pmatrix}. \tag{2.55}$$

Equations (2.45) and (2.46) can then be transformed to:

$$\boldsymbol{M}_{\ell m}^j(\rho,\varphi,z) = \frac{1}{4i^{\ell-1}} \int_0^\pi \mathrm{d}\theta' \mathrm{e}^{ik'\cos\theta'z} \left[\pi_{\ell m}(\theta')\boldsymbol{p}_m(\theta',\rho,\varphi) + i\tau_{\ell m}(\theta')\boldsymbol{s}_m(\theta',\rho,\varphi)\right], \tag{2.56}$$

$$\boldsymbol{N}_{\ell m}^j(\rho,\varphi,z) = \frac{1}{4i^{\ell-1}} \int_0^\pi \mathrm{d}\theta' \mathrm{e}^{ik'\cos\theta'z} \left[\tau_{\ell m}(\theta')\boldsymbol{p}_m(\theta',\rho,\varphi) + i\pi_{\ell m}(\theta')\boldsymbol{s}_m(\theta',\rho,\varphi)\right]. \tag{2.57}$$

Two exemplary plane-wave decompositions of vector spherical harmonics are shown in figure 2.4, namely $\boldsymbol{M}_{4,1}^j$ and $\boldsymbol{N}_{1,-1}^j$. These particular examples were chosen to illustrate again that the number of symmetry planes rises with $\ell$ and is different for each component. Since the functions are evaluated in the $x$-$z$-plane, $\varphi = 0$ in the right halves of the plots and $\varphi = \pi$ in the left halves. Thus, the exponential functions $\exp(im\varphi)$ are real, which means that $\xi_m$ is real for even $m$ and imaginary for odd $m$. This results in purely real or purely imaginary components of $\boldsymbol{p}_m$ and $\boldsymbol{s}_m$, where a real $p_{m,x}$ is always paired with an imaginary $s_{m,x}$ and vice versa (and the same for $y$ and $z$).
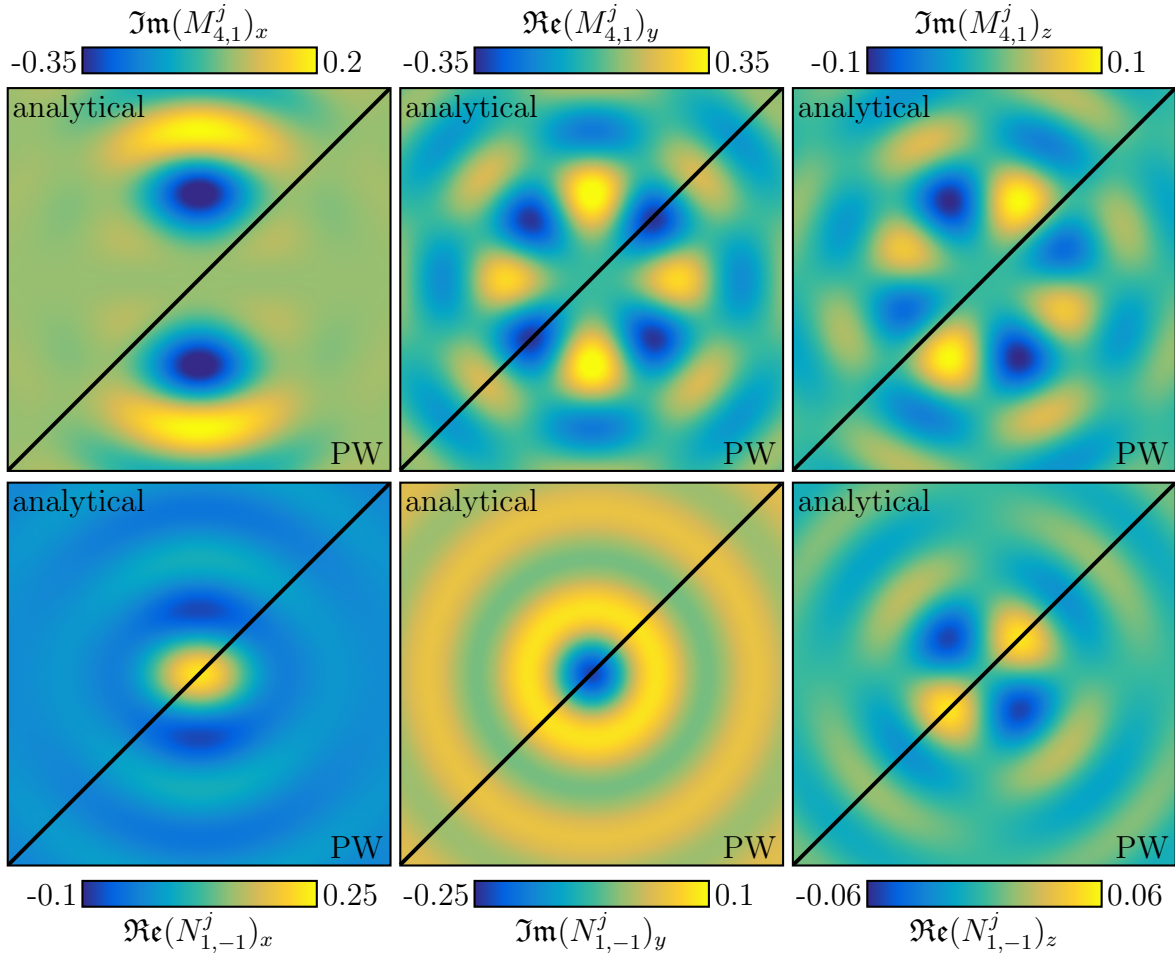
**Figure 2.4:** Plane-wave decompositions of different VSH. The $x$-, $y$- and $z$-component of $\boldsymbol{M}^j_{\ell m}$ and $\boldsymbol{N}^j_{\ell m}$ are either purely real or purely imaginary, shown is always the nontrivial part in the $x$-$z$-plane. The top left corner depicts the analytical result according to equations (2.45,2.46), the bottom right corner the PW decomposition (2.56,2.57) (integral over 250 evenly spaced $\theta'$-values between 0 and $\pi$, analytical $\varphi'$-integration, difference between analytical result and PW decomposition on the order of $10^{-6}$). Note the excellent agreement between analytical result and PW decomposition.

It follows from the formulas (2.56) and (2.57) that the $x$-, $y$- and $z$-components of each VSH are either purely real or purely imaginary. Figure 2.4 only shows these nontrivial components.

If the boundary conditions require the use of spherical Hankel functions instead of spherical Bessel functions, the PW decomposition becomes slightly more complicated. Mathematically, it is possible to change an integral representation of $j_\ell$ to an integral representation of $h_\ell$ by changing the integration path in the complex plane, see e.g. [45], §5.3. Physically, this can be justified as follows: While spherical Bessel functions are finite everywhere, spherical Hankel functions diverge at zero. This necessitates the inclusion of evanescent waves in the decomposition. Furthermore, as mentioned

in the previous section, spherical Hankel functions approach spherical waves for large arguments ([46], §19):

$$h_\ell(kr) \xrightarrow{kr \gg 1} \frac{1}{kr} e^{ikr - i(\ell+1)\pi/2} \,. \tag{2.58}$$

Intuitively, this can be realized by only using plane waves moving away from the origin in the decomposition. In practice, this would mean only using $\boldsymbol{k}$ with $\theta' \in [0, \pi/2]$ for $z > 0$ and $\theta' \in [\pi/2, \pi]$ for $z < 0$. Similarly, evanescent waves with decreasing amplitude for longer distances from the origin are given by $\theta' = \pi/2 + i\theta'', \theta'' \in [-\infty, 0]$ for $z > 0$ and $\theta'' \in [0, \infty]$ for $z < 0$. A comparison with the asymptotic form of the spherical Bessel function ([46], §19)

$$j_\ell(kr) \xrightarrow{kr \gg 1} \frac{1}{kr} \cos\left(kr - (\ell+1)\frac{\pi}{2}\right) = \frac{1}{2kr}\left(e^{ikr - i(\ell+1)\pi/2} + e^{-ikr + i(\ell+1)\pi/2}\right) \tag{2.59}$$

shows that the functions do not only differ in the plane waves that are needed, but also in a prefactor of two. These thoughts result in the PW decompositions

$$\boldsymbol{M}_{\ell m}^h(\rho, \varphi, z) = \frac{1}{2i^{\ell-1}} \int_{\mathcal{C}} d\theta' e^{ik' \cos\theta' z} \left[\pi_{\ell m}(\theta')\boldsymbol{p}_m(\theta') + i\tau_{\ell m}(\theta')\boldsymbol{s}_m(\theta')\right], \tag{2.60}$$

$$\boldsymbol{N}_{\ell m}^h(\rho, \varphi, z) = \frac{1}{2i^{\ell-1}} \int_{\mathcal{C}} d\theta' e^{ik' \cos\theta' z} \left[\tau_{\ell m}(\theta')\boldsymbol{p}_m(\theta') + i\pi_{\ell m}(\theta')\boldsymbol{s}_m(\theta')\right], \tag{2.61}$$

$$\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > 0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < 0 \end{cases}.$$

These formulas, too, were implemented in Matlab. Since it is not possible to integrate to infinity numerically, a cutoff $T$ has to be chosen for the imaginary part of $\theta'$, i.e. $\mathfrak{Im}(\theta') \in [-T, 0]$ (or $[0, T]$ for $z < 0$). The PW decompositions of the same functions $\boldsymbol{M}_{4,1}$ and $\boldsymbol{N}_{1,-1}$ as before, but now with $h$ instead of $j$, are presented for $T = 20$ in figure 2.5. As explained above, for $\boldsymbol{M}_{\ell m}^j$ and $\boldsymbol{N}_{\ell m}^j$, the $x$-, $y$- and $z$-components are each either purely real or purely imaginary in the $x$-$z$-plane. Since $h_\ell = j_\ell + iy_\ell$, changing from spherical Bessel to Hankel functions only changes the previously zero real (or imaginary) part. Thus, only this new result is shown in figure 2.5. The most striking feature is the appearance of very large, rapidly oscillating values close to $z = 0$. We call them ringing-artefacts, because this effect is similar to the Gibbs phenomenon that causes ringing artefacts in signal processing (e.g. [48] §5.7). It is caused by the evanescent waves, which have large values at $z = 0$ and then drop off quickly. To get the correct result everywhere, one would have to include evanescent waves up to very large imaginary part of $\theta'$, and with a dense sampling. This is further investigated in the chapter on the numerical implementation, 3.1.

The next step is to decompose a plane wave in vector spherical harmonics:

$$(E_p \hat{\boldsymbol{e}}_p + E_s \hat{\boldsymbol{e}}_s) e^{i\boldsymbol{k}' \cdot \boldsymbol{r}} = 4\pi \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \frac{i^{\ell-1}}{\ell(\ell+1)} e^{-im\varphi'} \,.$$

$$\left([E_p \pi_{\ell m}(\theta') - iE_s \tau_{\ell m}(\theta')] \boldsymbol{M}_{\ell m}^j(\boldsymbol{r}) + [E_p \tau_{\ell m}(\theta') - iE_s \pi_{\ell m}(\theta')] \boldsymbol{N}_{\ell m}^j(\boldsymbol{r})\right). \tag{2.62}$$

The derivation of the formula can be found e.g. in [49]. As before, dashed quantities $(\theta', \varphi')$ describe the direction of the wavevector $\boldsymbol{k}'$. Details on the convergence for a finite approximation of the infinite sum are presented in section 3.1.
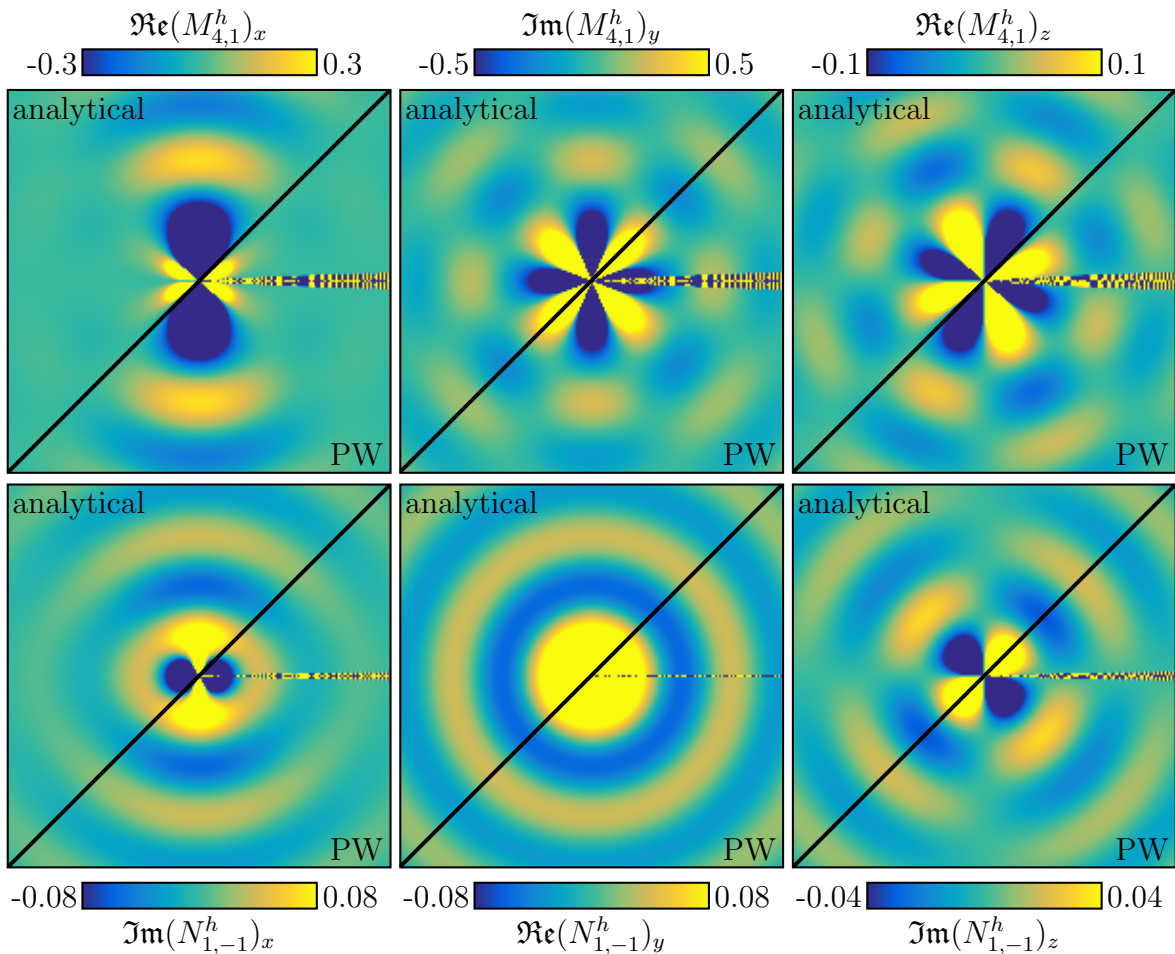


**Figure 2.5:** Plane-wave decompositions of different VSH. Either the real or the imaginary part of the $x$-, $y$- and $z$-component of $\boldsymbol{M}_{\ell m}^h$ ($\boldsymbol{N}_{\ell m}^h$) is diverging at the origin. The other part is finite everywhere and identical to the same part of the same component of $\boldsymbol{M}_{\ell m}^j$ ($\boldsymbol{N}_{\ell m}^j$), and therefore not shown here. The color range was restricted in order to make the patterns better visible. The top left corner depicts the analytical result in the $x$-$z$-plane according to equations (2.45,2.46), the bottom right corner the PW decomposition (2.60,2.61) (integral over 250 evenly spaced $\theta'$-values between 0 and $\pi/2$ and 200 evenly spaced $\theta'$-values between $\pi/2$ and $\pi/2 + 20i$ for $z > 0$; $\pi$ minus these angles for $z < 0$; analytical $\varphi'$-integration). Note the appearance of atefacts near $z = 0$.

## 2.2 Field of a dipole emitter

Remarkably, one can show [50] that most fluorescent dyes of practical interest can be modeled as ideal electric dipole emitters[3]. Therefore, most of this work investigates the behaviour of electric dipole emitters in different situations. In this section, we start by deriving the electrical field of such an emitter. Imagine a point-dipole with dipole moment $\boldsymbol{p}$ that is situated at the origin in a homogeneous material with refractive index $n$ and that oscillates with the angular frequency $\omega$. It has zero net charge, but is described by the current density $\boldsymbol{j} = -i\omega\boldsymbol{p}\delta(\boldsymbol{r})\exp(-i\omega t)$. As in the previous sections, we assume that the resulting electric and magnetic fields have the same time-dependence, such that the physical fields $\boldsymbol{E}_{\mathrm{phys}}$, $\boldsymbol{B}_{\mathrm{phys}}$ can be represented by their complex-valued temporal Fourier coefficients $\boldsymbol{E}$ and $\boldsymbol{B}$ via

$$\boldsymbol{E}_{\mathrm{phys}} = \mathfrak{Re}\left\{\boldsymbol{E}\mathrm{e}^{-i\omega t}\right\} \quad \text{and} \quad \boldsymbol{B}_{\mathrm{phys}} = \mathfrak{Re}\left\{\boldsymbol{B}\mathrm{e}^{-i\omega t}\right\} \ . \tag{2.63}$$

With these prerequisites, the third and fourth of Maxwell's equations read

$$\nabla \times \boldsymbol{E} = \ i\frac{\omega}{c}\boldsymbol{B} \ ,$$
$$\nabla \times \boldsymbol{B} = -i\frac{\omega n^2}{c}\boldsymbol{E} - \frac{4\pi}{c}i\omega\boldsymbol{p}\delta(\boldsymbol{r}). \tag{2.64}$$

Combined, this results in:

$$\nabla \times \nabla \times \boldsymbol{E} = \frac{n^2\omega^2}{c^2}\boldsymbol{E} + \frac{4\pi\omega^2}{c^2}\boldsymbol{p}\delta(\boldsymbol{r}). \tag{2.65}$$

In the following, we will present three different representations of $\boldsymbol{E}$, namely a closed-form solution, an expansion in vector spherical harmonics, and a decomposition of the field in plane waves. In later chapters, it will become obvious how these different forms can be adapted to describe the field of a dipole close to various kinds of nanostructures.

### 2.2.1 Closed-form solution

To find a closed-form solution, we solve equation (2.65) in Fourier space and then perform the inverse Fourier transform. In order to avoid confusing the Fourier space coordinate (arbitrary magnitude possible) with the wavevector used in the previous sections (magnitude defined by the wavelength), we will use the Fourier space coordinate $\boldsymbol{u}$ with magnitude $u = |\boldsymbol{u}|$. Then, equation (2.65) transforms to

$$-\boldsymbol{u} \times \boldsymbol{u} \times \tilde{\boldsymbol{E}} = \frac{n^2\omega^2}{c^2}\tilde{\boldsymbol{E}} + \frac{4\pi\omega^2}{c^2}\boldsymbol{p}, \tag{2.66}$$

---

[3] Or at least as an ensemble of such dipoles with different oscillation frequency and different dipole strength but identical orientation if one wants to take into account the fluorescence spectrum.

where the tilde represents a Fourier transform. With the definition $k_v := \omega/c$ and some basic vector algebra, this can be rearranged to

$$u^2\tilde{\boldsymbol{E}} - \boldsymbol{u}(\boldsymbol{u} \cdot \tilde{\boldsymbol{E}}) = k_v^2 \left(n^2\tilde{\boldsymbol{E}} + 4\pi\boldsymbol{p}\right). \tag{2.67}$$

By multiplying both sides with $\boldsymbol{u}$, we find the expression

$$0 = n^2\boldsymbol{u} \cdot \tilde{\boldsymbol{E}} + 4\pi\boldsymbol{u} \cdot \boldsymbol{p}. \tag{2.68}$$

When inserting this back into equation (2.67), we arrive at the solution in Fourier space:

$$u^2\tilde{\boldsymbol{E}} + \frac{4\pi}{n^2}\boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{p}) = k_v^2 \left(n^2\tilde{\boldsymbol{E}} + 4\pi\boldsymbol{p}\right) \tag{2.69}$$

$$\Rightarrow \quad \tilde{\boldsymbol{E}} = \frac{4\pi}{n^2(u^2 - n^2k_v^2)}\left[n^2k_v^2\boldsymbol{p} - \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{p})\right]. \tag{2.70}$$

As the next step, we have to perform the inverse Fourier transform to find the solution in direct space:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{4\pi}{n^2}\int \frac{\mathrm{d}^3u}{(2\pi)^3}\frac{\exp(i\boldsymbol{u} \cdot \boldsymbol{r})}{u^2 - n^2k_v^2}\left[n^2k_v^2\boldsymbol{p} - \boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{p})\right]. \tag{2.71}$$

By recalling that $\nabla \cdot \boldsymbol{p}\exp(i\boldsymbol{u} \cdot \boldsymbol{r}) = i\boldsymbol{u} \cdot \boldsymbol{p}\exp(i\boldsymbol{u} \cdot \boldsymbol{r})$ and $\nabla[\nabla \cdot \boldsymbol{p}\exp(i\boldsymbol{u} \cdot \boldsymbol{r})] = -\boldsymbol{u}[\boldsymbol{u} \cdot \boldsymbol{p}\exp(i\boldsymbol{u} \cdot \boldsymbol{r})]$ since the derivatives only act on $\boldsymbol{r}$, not $\boldsymbol{u}$, we can rewrite this as

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{4\pi}{n^2}\left[n^2k_v^2 + \nabla(\nabla\cdot)\right]\boldsymbol{p}\int \frac{\mathrm{d}^3u}{(2\pi)^3}\frac{\exp(i\boldsymbol{u} \cdot \boldsymbol{r})}{u^2 - n^2k_v^2}. \tag{2.72}$$

We will now solve this integral in spherical coordinates $(u, \theta, \varphi)$. The coordinate system can always be chosen such that $\boldsymbol{u} \cdot \boldsymbol{r} = ur\cos\theta$. Then, the integrand does not depend on $\varphi$ and the $\varphi$-integration simply gives a factor $2\pi$. The expression left to evaluate is then

$$\frac{1}{(2\pi)^2}\int_0^\infty \mathrm{d}u \int_0^\pi \mathrm{d}\theta\ \sin\theta\ u^2\frac{\exp(iur\cos\theta)}{u^2 - n^2k_v^2}$$

$$= \frac{1}{(2\pi)^2}\int_0^\infty \mathrm{d}u \int_{-1}^1 \mathrm{d}(\cos\theta)\ u^2\frac{\exp(iur\cos\theta)}{u^2 - n^2k_v^2}$$

$$= \frac{1}{ir(2\pi)^2}\int_0^\infty \mathrm{d}u\ u\frac{\exp(iur) - \exp(-iur)}{u^2 - n^2k_v^2}$$

$$= \frac{1}{ir(2\pi)^2}\int_{-\infty}^\infty \mathrm{d}u\ u\frac{\exp(iur)}{u^2 - n^2k_v^2}. \tag{2.73}$$

This integral can be closed in the upper halfspace of the complex $u$-plane without changing its value because $\exp(iur) \to 0$ for $\mathfrak{Im}\{u\} \to \infty$. Then, the integrand has two poles, namely $u = \pm nk_v$. Taking into account the physical situation – a radiating

dipole – only outgoing waves with positive $u$ are reasonable. Thus, the integration path $\mathcal{C}$ is deformed to exclude the pole $u = -nk_v$. Cauchy's residue theorem then yields

$$\frac{1}{ir(2\pi)^2} \oint_{\mathcal{C}} \mathrm{d}u \frac{u \exp(iur)}{(u - nk_v)(u + nk_v)} = \frac{\pi}{r(2\pi)^2} \mathrm{e}^{ink_v r}. \tag{2.74}$$

Thus, we find for the electrical field:

$$\boxed{\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{n^2} \left[ n^2 k_v^2 + \nabla(\nabla \cdot) \right] \boldsymbol{p} \frac{\exp(ink_v r)}{r}.} \tag{2.75}$$

Recalling the considerations from section 2.1.1, we see that $k_v = \omega/c$ is the wavevector in vacuum, and that $nk_v$ is the wavevector in the homogeneous medium. Thus, we denote $k = nk_v$. We will now carry out the partial derivatives explicitly to find a more direct representation of the field. We start with

$$\nabla \cdot \left[ \boldsymbol{p} \frac{\exp(ikr)}{r} \right] = \left( \frac{ik}{r^2} - \frac{1}{r^3} \right) (\boldsymbol{r} \cdot \boldsymbol{p}) \exp(ikr). \tag{2.76}$$

For the gradient, consider the derivative with respect to any of the three components of $\boldsymbol{r}$:

$$\frac{\partial}{\partial r_j} \left( \frac{ik}{r^2} - \frac{1}{r^3} \right) (\boldsymbol{r} \cdot \boldsymbol{p}) \exp(ikr) = \left[ r_j \left( -\frac{2ik}{r^4} + \frac{3}{r^5} \right) (\boldsymbol{r} \cdot \boldsymbol{p}) + \left( \frac{ik}{r^2} - \frac{1}{r^3} \right) p_j \right.$$
$$\left. + \left( \frac{ik}{r^2} - \frac{1}{r^3} \right) (\boldsymbol{r} \cdot \boldsymbol{p}) ik \frac{r_j}{r} \right] \exp(ikr) \tag{2.77}$$

$$\Rightarrow \nabla(\nabla \cdot) \left[ \boldsymbol{p} \frac{\exp(ikr)}{r} \right] = \left[ \left( \frac{ik}{r^2} - \frac{1}{r^3} \right) \boldsymbol{p} + \left( -\frac{k^2}{r^3} - \frac{3ik}{r^4} + \frac{3}{r^5} \right) \boldsymbol{r}(\boldsymbol{r} \cdot \boldsymbol{p}) \right] \mathrm{e}^{ikr} \tag{2.78}$$

This leads to the final result:

$$\boldsymbol{E}_D(\boldsymbol{r}) = nk_v^3 \left[ \left( 1 + \frac{i}{kr_d} - \frac{1}{(kr_d)^2} \right) \boldsymbol{p} + \left( -1 - \frac{3i}{kr_d} + \frac{3}{(kr_d)^2} \right) \hat{\boldsymbol{r}}_d (\hat{\boldsymbol{r}}_d \cdot \boldsymbol{p}) \right] \frac{\exp(ikr_d)}{kr_d}, \tag{2.79}$$

where we now generalised for an arbitrary dipole position $\boldsymbol{r}_0$ by defining the vector from the position of the dipole to the position where the field is evaluated, $\boldsymbol{r}_d = \boldsymbol{r} - \boldsymbol{r}_0$, with magnitude $r_d$ and direction $\hat{\boldsymbol{r}}_d$. Some exemplary plots of the resulting field are shown in figure 2.6 for dipoles oriented parallel to the $x$-, $y$- or $z$-axis.

## 2.2.2 VSH-decomposition

It is often advisable to decompose the dipole field into vector spherical harmonics in order to satisfy boundary conditions imposed by the experimental setup. The derivation
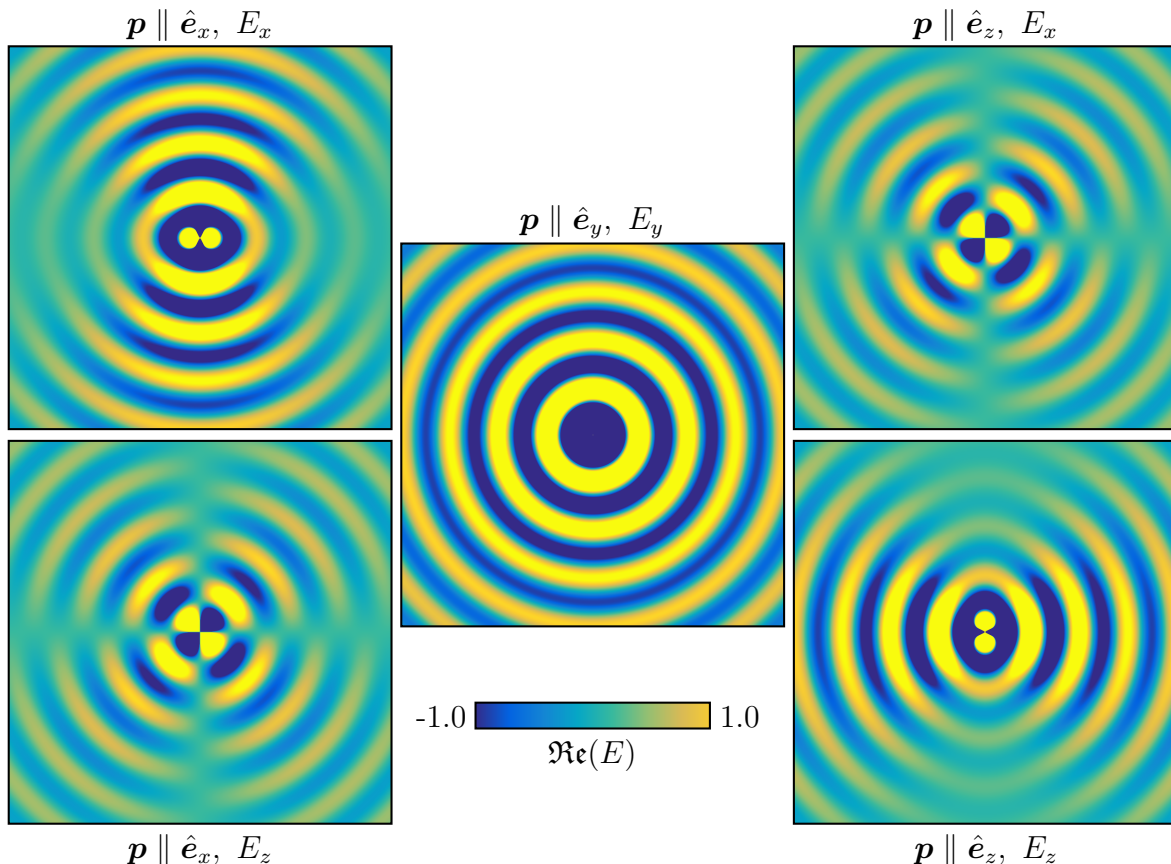
**Figure 2.6:** Real part of the electric field of an $x$-, $y$- or $z$-dipole situated at the origin, with arbitrary amplitude. Shown are the VSH-versions from equation (2.82) but the analytical version from equation (2.79) is identical within numerical accuracy. The field components that are not shown are zero, the evaluation is done in the $x$-$z$-plane in the interval $[-4\lambda, 4\lambda]$.

is rather complicated, it can be found for example in Morse and Feshbach's *Methods of Theoretical Physics* [45], §13.3[4]. The result is:

$$\boldsymbol{E}_{\mathrm{VSH}}(\boldsymbol{r}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^D \boldsymbol{M}_{\ell m}^f(k\boldsymbol{r}) + b_{\ell m}^D \boldsymbol{N}_{\ell m}^f(k\boldsymbol{r}) \tag{2.80}$$

with the expansion coefficients

$$a_{\ell m}^D = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{M}_{\ell m}^g(k\boldsymbol{r}_0)} \quad \text{and} \quad b_{\ell m}^D = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{N}_{\ell m}^g(k\boldsymbol{r}_0)}, \tag{2.81}$$

where the bar denotes complex conjugation. This expansion is actually a piecewise-defined function because the choice of $f$ and $g$ depends on $\boldsymbol{r}$ and $\boldsymbol{r}_0$: If $r < r_0$, the VSH need to be finite at the origin, and thus $f = j_\ell$. Then, $g = h_\ell^2$, which corresponds to

---

[4] Note the different definition of VSH used by Morse and Feshbach: they omit the factor $c_{\ell m}$.
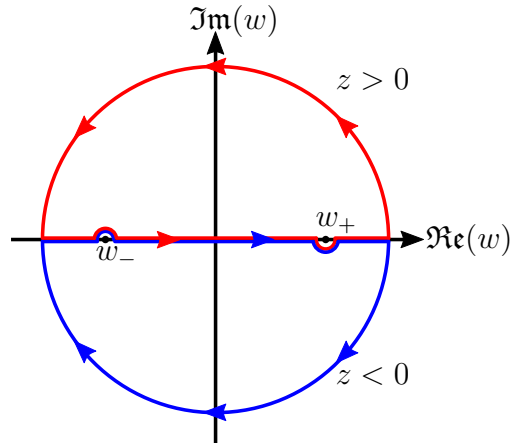
**Figure 2.7:** Path integrals used to integrate over $w$ in equation (2.83). Note that the blue path for $z < 0$ is taken clockwise and thus adds a negative sign to the result according to Cauchy's residue theorem.

incoming spherical waves[5]. On the other hand, if $r > r_0$, the VSH ultimately have to approximate outgoing spherical waves at very large $r$. Therefore, $f = h_\ell^1$ and $g = j_\ell$. By taking the limit $r_0 \to 0$, one finds very simple expansions for an $x$-, $y$- or $z$-dipole:

$$\boldsymbol{E}_{x\text{-dipole}} = -ink_v^3 p\sqrt{\frac{4\pi}{3}}\frac{1}{\sqrt{2}}\left(\boldsymbol{N}_{1,1}^{h^1} - \boldsymbol{N}_{1,-1}^{h^1}\right),$$

$$\boldsymbol{E}_{y\text{-dipole}} = -\ nk_v^3 p\sqrt{\frac{4\pi}{3}}\frac{1}{\sqrt{2}}\left(\boldsymbol{N}_{1,1}^{h^1} + \boldsymbol{N}_{1,-1}^{h^1}\right),$$

$$\boldsymbol{E}_{z\text{-dipole}} = \ \ ink_v^3 p\sqrt{\frac{4\pi}{3}}\ \ \boldsymbol{N}_{1,0}^{h^1}. \tag{2.82}$$

In other words, only $\ell = 1$ is needed if the dipole is situated in the origin – which can be a useful test case for different types of simulations. Several different orientations are explored in figure 2.6. More useful for the evaluation of experimental data is certainly the case that the dipole is not situated in the origin. Then, generally all VSH are needed for replicating the analytical solution. Since it is not possible to sum infinitely many terms on a computer, a truncation $\ell \leq L$ has to be chosen for the numerical evaluation. A study of the convergence behaviour can be found in section 3.1.

### 2.2.3 PW-decomposition

Last but not least, it is sometimes also necessary to express the electric field of a dipole emitter by the other set of basis functions that we introduced earlier, namely plane waves. For this, we return to the derivation of the analytical solution, but choose a different evaluation method starting from equation (2.71):

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{4\pi}{n^2}\int\frac{\mathrm{d}^3\boldsymbol{u}}{(2\pi)^3}\frac{\exp(i\boldsymbol{u}\cdot\boldsymbol{r}_d)}{u^2 - n^2 k_v^2}\left[n^2 k_v^2\boldsymbol{p} - \boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{p})\right]. \tag{2.83}$$

---

[5] Everywhere else in this text, $h_\ell$ means $h_\ell^1 = j_\ell + iy_\ell$, while $h_\ell^2 = j_\ell - iy_\ell$ with the spherical Neumann function $y_\ell$.

The term $\exp(i\boldsymbol{u} \cdot \boldsymbol{r}_d)$ looks like a plane wave with wavevector $\boldsymbol{u}$, so one might think that this is already the desired PW decomposition of the dipole field. However, the integral extends over the whole $\boldsymbol{u}$-space, which in particular includes all possible magnitudes of $|\boldsymbol{u}|$ – but we know that the magnitude of the physical wavevector is given by $2\pi/\lambda$. Thus, the wavevectors of the plane waves in the decomposition can only differ in direction, not magnitude. Therefore, some further conversions are necessary. Instead of transforming this equation using the gradient and divergence as in section 2.2.1, we switch to cylindrical coordinates $\boldsymbol{u} = \boldsymbol{q} + w\hat{\boldsymbol{e}}_z$, where $\boldsymbol{q} = q\cos(\varphi)\hat{\boldsymbol{e}}_x + q\sin(\varphi)\hat{\boldsymbol{e}}_y$ is in the $x$-$y$-plane, and carry out the integration from $-\infty$ to $\infty$ along the $w$-axis. As in section 2.2.1, we want to close the integration path in order to use Cauchy's residue theorem. If $\boldsymbol{r}_d$ has a positive $z$-component, the exponential function is negligible for $\mathfrak{Im}\{w\} \to \infty$. On the other hand, if $\boldsymbol{r}_d$ has a negative component, the integral can only be closed in the *negative* halfspace of the complex $w$-space without changing its value. To find the poles of the integrand, we rewrite

$$u^2 - n^2 k_v^2 = (q^2 + w^2) - n^2 k_v^2 = w^2 - (n^2 k_v^2 - q^2). \tag{2.84}$$

This shows that two poles exist at $w_\pm = \pm\sqrt{n^2 k_v^2 - q^2}$. As before, we demand outgoing waves, and thus the positive pole is included for $z_d > 0$ and the negative pole is included for $z_d < 0$. The final paths are shown in figure 2.7. With Cauchy's theorem, expressing $\boldsymbol{r}_d$ in cylindrical coordinates ($\boldsymbol{r}_d = (\boldsymbol{\rho}_d, z_d)$) and defining $\boldsymbol{k}^\pm = (\boldsymbol{q}, w_\pm)$ we then arrive at:

$$\boldsymbol{E}(x, y, z > z_0) = \frac{i}{n^2\pi} \int_0^\infty \mathrm{d}q \int_0^{2\pi} \mathrm{d}\varphi \, q \frac{\exp(i\boldsymbol{q} \cdot \boldsymbol{\rho}_d + iw_+ z_d)}{2w_+} \left[ n^2 k_v^2 \boldsymbol{p} - \boldsymbol{k}^+ (\boldsymbol{k}^+ \cdot \boldsymbol{p}) \right]$$

$$\boldsymbol{E}(x, y, z < z_0) = \frac{-i}{n^2\pi} \int_0^\infty \mathrm{d}q \int_0^{2\pi} \mathrm{d}\varphi \, q \frac{\exp(i\boldsymbol{q} \cdot \boldsymbol{\rho}_d + iw_- z_d)}{2w_-} \left[ n^2 k_v^2 \boldsymbol{p} - \boldsymbol{k}^- (\boldsymbol{k}^- \cdot \boldsymbol{p}) \right].$$

$$\tag{2.85}$$

This expression is called the *Weyl representation* of the electric field of an oscillating dipole, after Hermann Weyl, who was the first to include evanescent waves[6] in the plane-wave decomposition of spherical waves [51]. What is the physical meaning of this expression? We can interpret the sign change of $w$ for the two domains as integrals over plane waves with different angles $\theta$ relative to the $z$-axis, such that $w = k\cos\theta$. An upwards-traveling wave ($w > 0$) is given for $\theta \in [0, \pi/2]$, while a downwards-traveling wave ($w < 0$) is obtained for $\theta \in [\pi/2, \pi]$. In order to obtain the correct sign for the evanescent waves where $q > k$, we have to choose $\theta = \pi/2 + i\theta''$ with $\theta'' \in [-\infty, 0]$ for $\mathfrak{Im}\{w\} > 0$ (i.e. $z > z_0$) and $\theta'' \in [0, \infty]$ for $\mathfrak{Im}\{w\} < 0$ (i.e. $z < z_0$)[7]. This choice reflects the physical situation: Plane waves are moving away from the oscillating dipole, which – as we will see in section 2.4– corresponds to an energy flow away from the dipole. But this just means that the dipole is an energy source. Furthermore, this choice of plane waves is completely analogous to the integration paths we encountered in the plane-wave decomposition of vector spherical harmonics with Hankel functions

---

[6] These are waves with $q > k$ and imaginary $w$, i.e. waves whose amplitude is exponentially in- or decreasing with $z$. They will be introduced more thoroughly in section 2.3.1.

[7] If $\theta = \pi/2 + i\theta''$, then $\cos(\theta) = [\exp(i\theta) + \exp(-i\theta)]/2 = [i\exp(-\theta'') - i\exp(\theta'')]/2$.
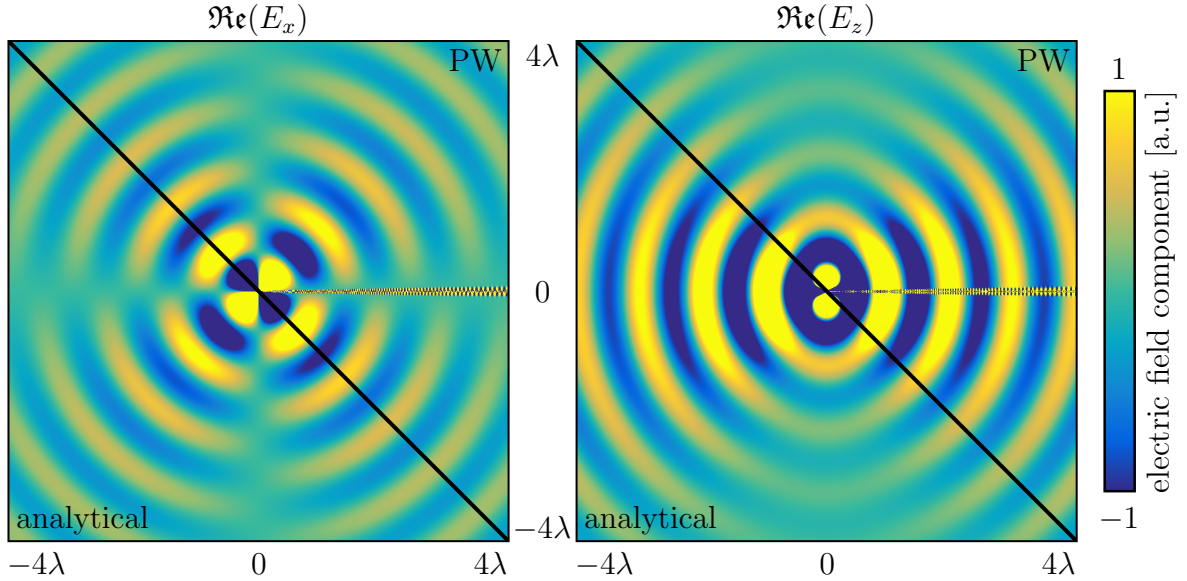
**Figure 2.8:** Comparison of the PW decomposition of a dipole field with the analytical solution. Shown are the real parts of $E_x$ and $E_z$ in the $x$-$z$-plane for a $z$-dipole situated in the origin, the $y$-component is zero in this plane. Calculated according to equation (2.86) with 250 evenly spaced $\theta$-values between 0 and $\pi/2$ and 1000 evenly spaced $\theta$-values between $\pi/2$ and $\pi/2 + 20i$ for $z > 0$; $\pi$ minus these angles for $z < 0$; analytical $\varphi$-integration. Note the artefacts at $z = 0$, which will be adressed in section 3.1.

in equations (2.60,2.61). Since we saw that the far field of a dipole can be expressed in exactly this type of VSH, this shows the consistency of the different representations. To sum up these thoughts, we can express the Weyl representation as an integral over the angles $\theta$ and $\varphi$ as

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{ink_v}{2n^2\pi} \int_{\mathcal{C}} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi \ \sin\theta e^{ik\rho_d \sin\theta \cos(\varphi-\varphi_d)+ik\cos\theta z_d} \left[ n^2 k_v^2 \boldsymbol{p} - \boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{p}) \right]$$

$$\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > z_0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < z_0 \end{cases}, \tag{2.86}$$

where now simply $\boldsymbol{k} = k(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ and $\varphi_d$ is the angle between $\boldsymbol{\rho}_d$ and the $x$-axis[8]. Which of the equations (2.85) or (2.86) is used is partly a matter of taste, and partly depends on which coordinate system fits the given situation best. Figure 2.8 shows the numerical result of this formula for a $z$-dipole situated in the origin, where the $\varphi$-integration was carried out analytically using Bessel functions as in section 2.1.4, and the cutoff on the imaginary axis is $\pm 20i$. Note the appearance of artefacts near $z = 0$, which will be discussed in section 3.1.

---

[8] Note that with $q = k\sin\theta$ we have $\mathrm{d}q = k\cos\theta\mathrm{d}\theta = w\mathrm{d}\theta$. Therefore, the term $q\mathrm{d}q/w$ becomes $k\sin\theta\mathrm{d}\theta$.

# 2.3 Interactions of electromagnetic waves with nanostructures
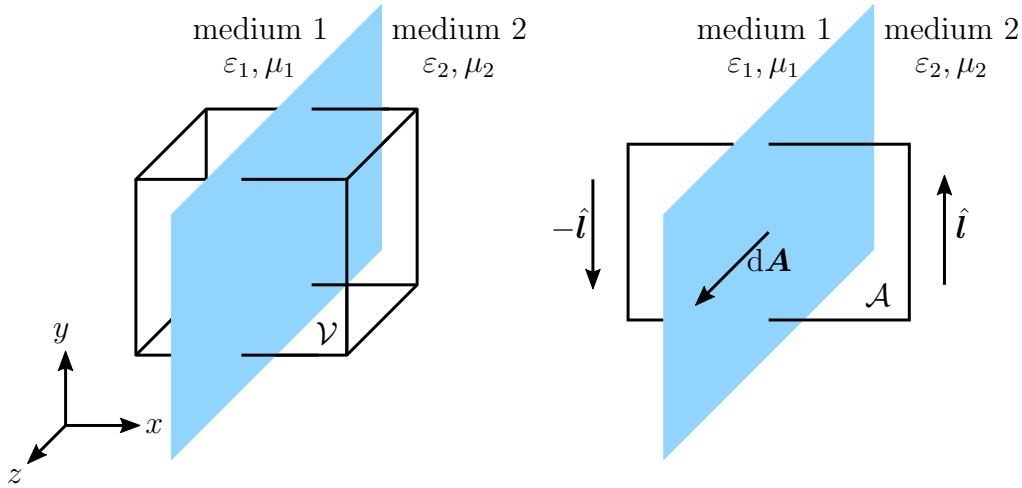


**Figure 2.9:** Geometries used to deduce boundary conditions for $\boldsymbol{E}$ and $\boldsymbol{B}$ from Maxwell's equations at the interface (shown in blue) between two materials 1 and 2. On the left, a small cuboid $\mathcal{V}$ intersects the boundary, on the right, there is a small rectangle $\mathcal{A}$ perpendicular to the interface.

In a real biophysical experiment, it is very rare to observe a dipole emitter in "free space". The closest we come to this ideal situation is a solution of dye molecules (or fluorescent proteins) in a solvent, imaged far away from any surfaces. However, most of this work is concerned with imaging fluorescently labeled structures in cells. The latter usually adhere to the cover slip (or Petri dish), so there is at least one interface present close to the emitters. This boundary has a strong influence on the energy emission of dipole emitters, changing for example the direction in which energy is emitted and the total amount of energy emitted per time. In order to quantitatively determine the change of the dipole field close to an interface, it is beneficial to decompose the dipole field into basis functions whose interaction with the interface is already known. Therefore, in this section, we study the interaction of increasingly complex fields with different types of interfaces. Starting with plane waves at a planar interface, which allows us to derive the field of a dipole emitter close to a planar interface, we then move on to vector spherical harmonics at spherical interfaces, which leads to the field of a dipole close to a spherical interface.

As a basis for these calculations, we need to know the boundary conditions that have to be fulfilled by electric and magnetic fields at the interfaces of materials with different refractive indices, which we will derive now. Consider the small cuboid $\mathcal{V}$ at the boundary between two media with different refractive indices $n_1$ and $n_2$ in figure (2.9). If we assume that there are no free charges in the media or at the interface, then according to equation (2.1) the electric field has to fulfill the relation

$$\nabla \cdot (\varepsilon_i \boldsymbol{E_i}) = 0 \ , i = 1, 2 \ , \tag{2.87}$$

at every point in medium 1 or 2, respectively. Using Gauss's theorem, we find

$$0 = \int_{\mathcal{V}} \nabla \cdot [\varepsilon(\boldsymbol{r})\boldsymbol{E}]\mathrm{d}V = \int_{\partial\mathcal{V}} \varepsilon(\boldsymbol{r})\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{A} = \varepsilon_1 E_{1,x} - \varepsilon_2 E_{2,x} \ , \qquad (2.88)$$

where we made the cuboid infinitely thin in the last step and therefore ignored the contributions from all sides except those perpendicular to the $x$-axis. For the magnetic field, one finds $B_{1,x} = B_{2,x}$ using the same reasoning with equation (2.2). Now consider the small rectangle $\mathcal{A}$ in figure (2.9). Using Stokes' theorem on the third of Maxwell's equations (2.3), we find

$$\int_{\partial\mathcal{A}} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = \int_{\mathcal{A}} (\nabla \times \boldsymbol{E}) \cdot \mathrm{d}\boldsymbol{A} = -\frac{1}{c}\int_{\mathcal{A}} \frac{\partial\boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{A} \approx \frac{i\omega}{c}\frac{\mathcal{A}}{2}(B_{1,z} + B_{2,z}) \ . \qquad (2.89)$$

By letting the sides perpendicular to the interface shrink to infinitesimal length, $\mathcal{A}$ approaches zero while all other terms stay finite. Thus we arrive at

$$0 = (\boldsymbol{E}_2 \cdot \hat{\boldsymbol{l}})l - (\boldsymbol{E}_1 \cdot \hat{\boldsymbol{l}})l \qquad (2.90)$$

with the tangent unit vector $\hat{\boldsymbol{l}}$. Since this is true for any orientation of the small rectangle $\mathcal{A}$, we can conclude that the tangential component of the electric field is the same in both media. For the magnetic field, equation (2.4) leads to the finding that the tangential component of $\boldsymbol{B}/\mu$ is the same in both media. Importantly, since we let both the cuboid and the rectangle shrink in the derivation of the boundary conditions, these are true both for planar and for curved interfaces – as long as one compares field components locally.

### 2.3.1 Plane waves at a planar interface

The four boundary conditions derived above can now be used to predict the behaviour of plane waves at a planar interface. The most general case is shown in figure 2.10, with one incoming and one outgoing wave in each medium. The "+"- or "−"-sign denotes if the wave is travelling in the positive or negative $z$-direction. We decompose each $\boldsymbol{k}$-vector into one component $\boldsymbol{q}$ parallel to the interface and one component $w$ perpendicular to it.

The fact that the periodicity of the waves has to be the same on both sides of the medium directly implies that $q$ is identical for all four waves. Since the wavelength and thus $k_i = 2\pi/\lambda_i$ are constant within one medium, it follows that $w_1^+ = -w_1^-$ and $w_2^+ = -w_2^-$. This is just another form of the well-known law of reflection, which states that the angle between the incident light and the normal to the surface is the same as the angle between the reflected light and the normal. This angle can be calculated directly from the components of the wave vector via

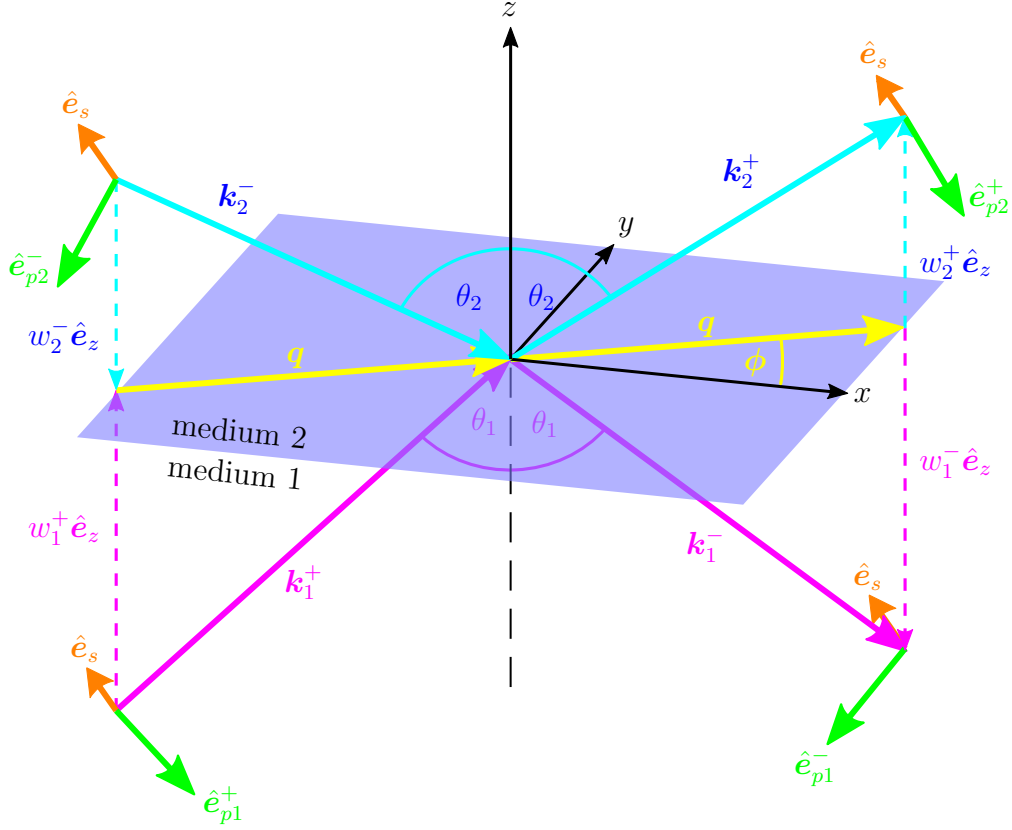$$\sin\theta_i = \frac{q}{k_i} = \frac{q}{n_i k_v} \ , \qquad (2.91)$$

**Figure 2.10:** Geometry of plane waves at an interface, definition of $p$- and $s$-waves and the relevant notation. The subscript 1 or 2 discriminates quantities in the two different media. The superscript $\pm$ shows if the quantitiy belongs to a wave traveling in the positive or negative $z$-direction. The vector $\hat{\boldsymbol{e}}_s$ is identical for all four waves.

which leads to the famous relation known as Snell's law,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \ . \tag{2.92}$$

We will now consider the amplitudes of the different fields. As stated in section 2.1.1, a linear combination of solutions of the wave equation is a solution itself. Furthermore, we have seen that both $\boldsymbol{E}_0$ and $\boldsymbol{B}_0$ are perpendicular to $\boldsymbol{k}$. Thus, if we find a basis $\{\boldsymbol{b}_i\}$ for the vector space $\{\boldsymbol{r}|\boldsymbol{r} \cdot \boldsymbol{k} = 0\}$, any plane wave of the form (2.9) can be expressed as a linear combination of

$$\boldsymbol{b}_i \ \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega t} \ . \tag{2.93}$$

Since the vector space $\{\boldsymbol{r}|\boldsymbol{r} \cdot \boldsymbol{k} = 0\}$ is two-dimensional, just two basis vectors are needed. The boundary conditions take on a simple form if one uses the basis vectors

$$\hat{\boldsymbol{e}}_p = \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{pmatrix} \quad \text{and} \quad \hat{\boldsymbol{e}}_s = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}, \tag{2.94}$$

where $\hat{e}_p$ lies in the *plane of incidence* spanned by $\boldsymbol{k}$ and the normal to the interface, and $\hat{e}_s$ is perpendicular to this plane. The corresponding waves are called transversal electric or $p$-wave and transversal magnetic or $s$-wave. In the following, we will study the behaviour of $p$- and $s$-waves at the interface, and use this to derive the behaviour of arbitrary plane waves. But first, we introduce a shorthand to facilitate calculations: The angle $\theta$ that appears in the definition of $\hat{e}_p$ is usually defined between the positive $z$-axis and $\boldsymbol{k}$. As figure 2.10 shows, this is true for $\theta_1$ and $\theta_2$ with $\boldsymbol{k}_1^+$ and $\boldsymbol{k}_2^+$. However, for $\boldsymbol{k}_1^-$ and $\boldsymbol{k}_2^-$, the angle to the positive $z$-axis is given by $\pi - \theta_1$ and $\pi - \theta_2$, respectively. Since $\sin(\pi - \theta) = \sin\theta$ and $\cos(\pi - \theta) = -\cos\theta$, we can introduce a new unit vector $\hat{e}_p^-$ for these two $p$-waves that directly uses $\theta_{1,2}$ instead of $\pi - \theta_{1,2}$:

$$
\hat{e}_{pj}^- = \begin{pmatrix} -\cos\theta_j \cos\phi \\ -\cos\theta_j \sin\phi \\ -\sin\theta_j \end{pmatrix}, \quad \hat{e}_{pj}^+ = \hat{e}_{pj} = \begin{pmatrix} \cos\theta_j \cos\phi \\ \cos\theta_j \sin\phi \\ -\sin\theta_j \end{pmatrix}, \quad j = 1, 2. \tag{2.95}
$$

The boundary conditions derived above place restrictions of the tangential and perpendicular components of the electric and magnetic field. We start by considering four $p$-waves with $\boldsymbol{E}_i^\pm = E_{pi}^\pm \hat{e}_{pi}^\pm$ and $\boldsymbol{B}_i^\pm = B_{pi}^\pm \hat{e}_s$. Simple geometry (see figure 2.10) leads to

$$
E_{pi,\parallel}^\pm = \pm \frac{w_i}{k_i} E_{pi}^\pm
$$
$$
E_{pi,\perp}^\pm = \frac{q}{k_i} E_{pi}^\pm
$$
$$
B_{pi,\parallel}^\pm = B_{pi}
$$
$$
B_{pi,\perp}^\pm = 0 . \tag{2.96}
$$

Under the additional assumption of non-magnetic materials, i.e. $\mu_1 = 1 = \mu_2$ and thus $n_i = \sqrt{\varepsilon_i}$, the boundary conditions for the electric field can now be compactly written as

$$
\begin{pmatrix} w_1/k_1 & -w_1/k_1 \\ qn_1^2/k_1 & qn_1^2/k_1 \end{pmatrix} \begin{pmatrix} E_{p1}^+ \\ E_{p1}^- \end{pmatrix} = \begin{pmatrix} w_2/k_2 & -w_2/k_2 \\ qn_2^2/k_2 & qn_2^2/k_2 \end{pmatrix} \begin{pmatrix} E_{p2}^+ \\ E_{p2}^- \end{pmatrix} . \tag{2.97}
$$

Both $q$ and $2\pi/\lambda$ (from $k_i = n_i 2\pi/\lambda$) can be canceled out. Simplifying the equation and multiplying with the inverse of the left matrix finally results in

$$
\begin{pmatrix} E_{p1}^+ \\ E_{p1}^- \end{pmatrix} = \frac{1}{2w_1} \begin{pmatrix} n_1 & w_1/n_1 \\ -n_1 & w_1/n_1 \end{pmatrix} \begin{pmatrix} w_2/n_2 & -w_2/n_2 \\ n_2 & n_2 \end{pmatrix} \begin{pmatrix} E_{p2}^+ \\ E_{p2}^- \end{pmatrix}
$$
$$
= \underline{\underline{M}}_{21}^p \begin{pmatrix} E_{p2}^+ \\ E_{p2}^- \end{pmatrix} \quad \text{with} \quad \underline{\underline{M}}_{21}^p = \frac{1}{2} \begin{pmatrix} w/n+n & -w/n+n \\ -w/n+n & w/n+n \end{pmatrix} , \tag{2.98}
$$

where the abbreviations $n = n_2/n_1$ and $w = w_2/w_1$ were introduced. Since both the relations between $\boldsymbol{E}$ and $\boldsymbol{B}$ described in the equations (2.10) and the boundary conditions are based on Maxwell's equations, one can now either construct a similar matrix equation for $\boldsymbol{B}$, or derive the magnetic fields of the four waves from the

corresponding electric fields. In practice it is usually easier to do the latter. Following the same reasoning, we also find matrix equations for $s$-waves. With

$$
\begin{aligned}
E_{si,\parallel}^{\pm} &= E_{si}^{\pm} \\
E_{si,\perp}^{\pm} &= 0 \\
B_{si,\parallel}^{\pm} &= \pm\frac{w_i}{k_i}B_{si}^{\pm} \\
B_{si,\perp}^{\pm} &= \frac{q}{k_i}B_{si}^{\pm}
\end{aligned}
\tag{2.99}
$$

and $B_i = n_i E_i$ (see discussion after equation (2.10)), the boundary conditions for the electric field have the form

$$
\begin{pmatrix} 1 & 1 \\ w_1 & -w_1 \end{pmatrix}\begin{pmatrix} E_{s1}^+ \\ E_{s1}^- \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ w_2 & -w_2 \end{pmatrix}\begin{pmatrix} E_{s2}^+ \\ E_{s2}^- \end{pmatrix}.
\tag{2.100}
$$

This can then be rearranged to yield

$$
\begin{pmatrix} E_{s1}^+ \\ E_{s1}^- \end{pmatrix} = \underline{\underline{M}}_{21}^s \begin{pmatrix} E_{s2}^+ \\ E_{s2}^- \end{pmatrix} \quad \text{with} \quad \underline{\underline{M}}_{21}^s = \frac{1}{2}\begin{pmatrix} 1+w & 1-w \\ 1-w & 1+w \end{pmatrix}.
\tag{2.101}
$$

Equations (2.98) and (2.101) describe the relation between the complex electric field amplitudes of the incoming and outgoing waves in both media for $p$- and $s$-waves. They have several useful applications. We will start by deriving the reflection and transmission coefficients

$$
r_{p,s} = \frac{E_{(p,s)1}^-}{E_{(p,s)1}^+} \quad \text{and} \quad t_{p,s} = \frac{E_{(p,s)2}^+}{E_{(p,s)1}^+}
\tag{2.102}
$$

for the special case of a plane wave incident from medium 1. Then, $E_{(p,s)2}^- = 0$ and we find:

$$
\begin{aligned}
r_p &= \frac{n^2 - w}{n^2 + w} & t_p &= \frac{2n}{n^2 + w} \\
r_s &= \frac{1 - w}{1 + w} & t_s &= \frac{2}{1 + w}
\end{aligned}
\tag{2.103}
$$

These so-called *Fresnel coefficients* describe the relationship between the amplitudes of the incoming, reflected and transmitted plane waves *at the interface*. This is important to keep in mind when discussing special cases such as total internal reflection or medias that absorb radiation, which will be discussed in more detail later. When $n_2 < n_1$,

$$
\theta_2 = \arcsin\left(\frac{n_1}{n_2}\sin\theta_1\right) > \theta_1 ,
\tag{2.104}
$$

i.e. the wave is refracted away from the surface normal. Thus there has to be an incident angle $\theta_c$ where $\theta_2$ equals ninety degrees. This angle is called the *critical angle*, it is given by $\theta_c = \arcsin(n_2/n_1)$. At even higher incident angles, Snell's law can no longer be

fulfilled and there is no propagating wave in the second medium. However, a different type of wave emerges: For $\theta_1 > \theta_c$,

$$w_2 = \sqrt{k_2^2 - q^2} = \sqrt{k_0^2 n_2^2 - k_0^2 n_1^2 \sin^2 \theta_1} = i \cdot k_0 \cdot \sqrt{n_1^2 \sin^2 \theta_1 - n_2^2} \qquad (2.105)$$

is purely imaginary. Thus the plane wave

$$\boldsymbol{E}_2^+ \, e^{iq_x x + iq_y y + iw_2 z} = \boldsymbol{E}_2^+ e^{-i|w_2|z} e^{iq_x x + iq_y y} \qquad (2.106)$$

has an amplitude that decays exponentially with $z$. This type of wave is called an *evanescent wave*. In this case, both $r_p$ and $r_s$ have an amplitude of 1, which means that the intensity of the incident and reflected wave are identical. No energy is transported into the second, optically thinner medium, as we will prove in section 2.4.2.

The simple situation of a single interface between two media described by equations (2.98) and (2.101) can easily be extended to a multi-layered sample. To this end, we need to describe the effect that traversing a layer has on the phase and amplitude of a plane wave. Remember that

$$\boldsymbol{A}_0 \, e^{i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t} = \boldsymbol{A}_0 \, e^{i(q_x x + q_y y + wz) - i\omega t}, \qquad (2.107)$$

where $q_x$ and $q_y$ are the same in all media. Thus, when traversing layer $j$ with thickness $d_j$, the wave accumulates the phase $iw_j d_j$ when traveling in positive $z$-direction and $-iw_j d_j$ when traveling in negative $z$-direction. If we denote the amplitudes of the plane waves at the top and bottom of layer $j$ by $E_{j,t}^{\pm}$ and $E_{j,b}^{\pm}$, we have

$$\begin{pmatrix} E_{j,b}^+ \\ E_{j,b}^- \end{pmatrix} = \underline{\underline{T}}_j \begin{pmatrix} E_{j,t}^+ \\ E_{j,t}^- \end{pmatrix} \quad \text{with} \quad \underline{\underline{T}}_j = \begin{pmatrix} \exp(-iw_j d_j) & 0 \\ 0 & \exp(iw_j d_j) \end{pmatrix}. \qquad (2.108)$$

Thus, when combining three different media as in figure 2.11, we have

$$\begin{pmatrix} E_{p1,t}^+ \\ E_{p1,t}^- \end{pmatrix} = \underline{\underline{M}}_{12}^p \begin{pmatrix} E_{p2,b}^+ \\ E_{p2,b}^- \end{pmatrix}$$

$$= \underline{\underline{M}}_{12}^p \underline{\underline{T}}_2 \begin{pmatrix} E_{p2,t}^+ \\ E_{p2,t}^- \end{pmatrix}$$

$$= \underline{\underline{M}}_{12}^p \underline{\underline{T}}_2 \underline{\underline{M}}_{23}^p \begin{pmatrix} E_{p3,b}^+ \\ E_{p3,b}^- \end{pmatrix} \qquad (2.109)$$

or the equivalent for $s$-waves. By multiplying the three matrices, effective reflection and transmission coefficients

$$r_{p,s} = \frac{E_{(p,s)1,t}^-}{E_{(p,s)1,t}^+} \quad \text{and} \quad t_{p,s} = \frac{E_{(p,s)3,b}^+}{E_{(p,s)1,t}^+} \qquad (2.110)$$

can immediately be deduced. This procedure scales well to even more layers, especially when automating the calculations on a computer. Some numerical difficulties can arise when any of the $w_j$ become imaginary, these are treated in section 3.2.
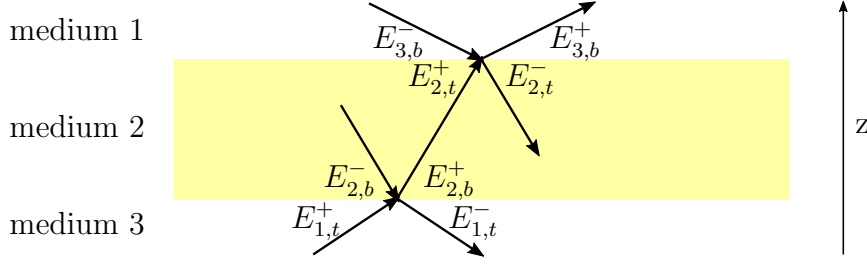
**Figure 2.11:** Reflection and transmission of electromagnetic waves in a stratified sample consisting of three planar layers. Naming convention for the amplitudes of the waves at the top or bottom (subscript $t/b$) of the three layers and moving in the positive or negative $z$-direction (superskript $\pm$).

## 2.3.2 Dipole emitter above a planar interface

The results from the previous section can immediately be combined with the plane-wave representation of a dipole field from section 2.2.3 to calculate the electrical field of a dipole emitter that is situated close to an interface, such as a dye molecule deposited on a cover slip and surrounded by air. We start by examining the amplitudes of the single plane waves in the Weyl representation (2.85)

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{ink_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \ \sin\theta \mathrm{e}^{ik\rho_d \sin\theta \cos(\phi-\phi_d)+ik\cos\theta z_d} \left[ \boldsymbol{p} - \hat{\boldsymbol{k}}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p}) \right]$$

$$\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > z_0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < z_0 \end{cases},$$

where again $\boldsymbol{k} = k(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $\boldsymbol{\rho}_d = (x - x_0, y - y_0)$, $z_d = z - z_0$ and $\phi_d = \sphericalangle(\boldsymbol{\rho_d}, \hat{\boldsymbol{e}}_x)$. The expression $\hat{\boldsymbol{k}}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p})$ is the projection of the dipole moment $\boldsymbol{p}$ on the unit vector in the direction of the wave vector $\boldsymbol{k}$. When subtracting this from $\boldsymbol{p}$, what remains is the component of $\boldsymbol{p}$ perpendicular to $\boldsymbol{k}$. But as we saw in the previous section, such a vector can always be expressed as a linear combination of the two basis vectors $\hat{\boldsymbol{e}}_p$ and $\hat{\boldsymbol{e}}_s$. The weights of this linear combination are simply

$$\hat{\boldsymbol{e}}_p \cdot \left[ \boldsymbol{p} - \hat{\boldsymbol{k}}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p}) \right] = \hat{\boldsymbol{e}}_p \cdot \boldsymbol{p}$$

$$\text{and} \quad \hat{\boldsymbol{e}}_s \cdot \left[ \boldsymbol{p} - \hat{\boldsymbol{k}}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p}) \right] = \hat{\boldsymbol{e}}_s \cdot \boldsymbol{p} \tag{2.111}$$

since $\hat{\boldsymbol{e}}_p$ and $\hat{\boldsymbol{e}}_s$ are, per definition, orthogonal to $\boldsymbol{k}$. Thus, we arrive at a new form of the Weyl representation of the field of a dipole emitter in a homogeneous medium:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{ink_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \ \sin\theta \mathrm{e}^{ik\rho_d \sin\theta \cos(\phi-\phi_d)+ik\cos\theta z_d} \left[ \hat{\boldsymbol{e}}_p(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p}) + \hat{\boldsymbol{e}}_s(\hat{\boldsymbol{e}}_s \cdot \boldsymbol{p}) \right]$$

$$\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > z_0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < z_0 \end{cases}. \tag{2.112}$$

**Theory**

Let us now imagine that there is a planar interface at $z = 0$ between two media with refractive indices $n_1$ ($z > 0$) and $n_2$ ($z < 0$). If the dipole is positioned at $\boldsymbol{r}_0 = (x_0, y_0, z_0)$ above the interface, the field directly at the interface is given by $\boldsymbol{E}(x, y, 0)$ with the integration path along $\pi/2 + i\infty \to \pi/2 \to \pi$ and with $n = n_1$. The reflected field $\boldsymbol{E}_R(x, y, z > 0)$ in the upper halfspace and the transmitted field $\boldsymbol{E}_T(x, y, z < 0)$ in the lower halfspace can then be found by treating each plane wave of the decomposition separately. The procedure is as follows:

1. Calculate the reflection and transmission coefficients $r_{(p,s)}(\theta)$ and $t_{(p,s)}(\theta)$ according to equation (2.103).

2. Multiply the amplitude of the wave, i.e. $(in_1 k_v^3)/(2\pi) \exp(-ik_1 \cos\theta_1 z_0)(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p})$ for a $p$-wave or $(in_1 k_v^3)/(2\pi) \exp(-ik_1 \cos\theta_1 z_0)(\hat{\boldsymbol{e}}_s \cdot \boldsymbol{p})$ for an $s$-wave, with the corresponding coefficient to get the amplitude of the reflected or transmitted wave. Note that the phase that the wave has accumulated by traveling from $z = z_0$ to $z = 0$ is taken into account here.

3. Find the new unit vectors of polarization: For the transmitted wave, $\hat{\boldsymbol{e}}_{p2}$ and $\hat{\boldsymbol{e}}_{s2}$ are obtained by using $\theta_2 = \arcsin(\sin\theta_1 n_1/n_2)$. For the reflected wave, $\theta_1$ has to be replaced by $\theta_1^* = \pi - \theta_1$, which does not change $\hat{\boldsymbol{e}}_{s1}$ but flips the sign of the $x$- and $y$-components of $\hat{\boldsymbol{e}}_{p1}^*$ compared to $\hat{\boldsymbol{e}}_{p1}$.

4. Construct the final wave by multiplying the amplitude with the new unit vector of polarization and the new exponential function. Since $n_i k_v \sin\theta_i =: q$ is constant everywhere, the only change is in the $z$-component of the exponent. This is given by $-ik_1 \cos\theta_1 z$ for the reflected wave and $ik_2 \cos\theta_2 z$ for the transmitted wave.

Thus the final fields are:

$$\boldsymbol{E}_T(x, y, z < 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \, \sin\theta_1 \left[ T_p \hat{\boldsymbol{e}}_{p2}(\hat{\boldsymbol{e}}_{p1} \cdot \boldsymbol{p}) + T_s \hat{\boldsymbol{e}}_{s2}(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p}) \right]$$
$$\cdot \exp\left(iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0 + ik_2 \cos\theta_2 z\right)$$

$$\boldsymbol{E}_R(x, y, z > 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \, \sin\theta_1 \left[ R_p \hat{\boldsymbol{e}}_{p1}^*(\hat{\boldsymbol{e}}_{p1} \cdot \boldsymbol{p}) + R_s \hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p}) \right]$$
$$\cdot \exp\left(iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0 + ik_1 \cos\theta_1^* z\right)$$

with the path $\quad \mathcal{C} = \pi/2 + i\infty \to \pi/2 \to \pi \quad$ and with $\theta_1^* := \pi - \theta_1$. $\qquad$ (2.113)

Note that the reflection of a wave causes the $z$-component of the wavevector to change its sign, which is equivalent to the change from propagation direction $\theta$ to propagation direction $\theta^* := \pi - \theta$. As mentioned in step 3, this also has an effect on the polarization of the wave, since the vector $\hat{\boldsymbol{e}}_p$ changes the sign of its $x$- and $y$-component when $\theta$ is replaced by $\theta^*$ (denoted by $\hat{\boldsymbol{e}}_p^*$). However, this change does not affect the coefficient $(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p})$ that is part of the amplitude of the $p$-wave. Also, due to the different definitions of the angle $\theta$ in this section and the previous one, we have to use $\theta^*$ when calculating all reflection and transmission coefficients.

As an example, figure 2.12 shows the electric field of a vertical dipol emitter in a water halfspace close to an interface with an air, glass or silver halfspace. The first

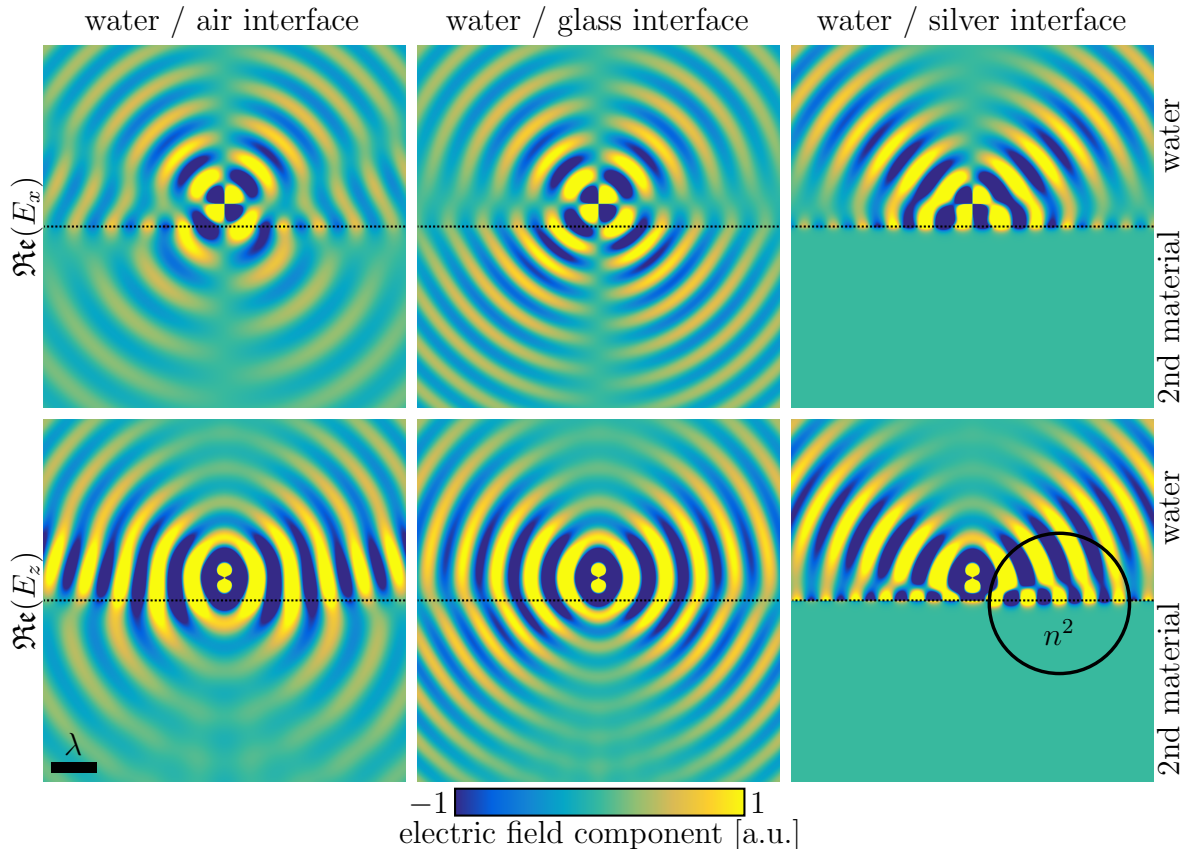**Figure 2.12:** Interaction of a dipole field with a planar interface. In all three cases, a vertical dipole ($\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_z$) is placed at height $z = \lambda/2$ above the interface between water and a second material – air ($n = 1$), glass ($n = 1.52$) or silver ($n = 0.05 + 2.80i$ at $\lambda = 500\,\text{nm}$). Shown are the real parts of $E_x$ and $E_z$ in the $x$-$z$-plane in the region $[-4\lambda, 4\lambda]$. $E_x$ is continuous across the interface, $E_z$ is discontinuous. However, the small circle in the bottom right panel shows that $n^2 E_z$ is continuous.

obvious effect of the different refractive index is the changed wavelength in the second medium, it is longer for air and shorter for glass compared to water. Since silver absorbs radiation, the field quickly decays to zero and no wavelength can be seen[9]. Secondly, the field in the water halfspace changes due to the added reflected waves. Thirdly, a wave that is confined to the interface itself appears in the presence of silver. This phenomenon, called *surface plasmon-polaritons*, will be explained in section 2.4.3. And finally, the figure nicely illustrates the boundary conditions from the previous section: $E_x$, i.e. the component parallel to the interface, is continuous, while $E_z$ is discontinuous. However, as the small circle in the bottom right panel demonstrates, $n^2 E_z = \varepsilon E_z$ is again continuous.

---

[9] The behaviour of electromagnetic fields in metals will be described in more detail in section 2.4.2.

### 2.3.3 Vector spherical harmonics at a spherical interface

After having found a description for planar interfaces, we now turn our attention to spherical interfaces. We encounter these e.g. when measuring dye molecules embedded in polymer beads, fluorescent proteins in solution close to metal nanospheres, or when scattering a laser beam at a glass bead. Historically, the term *reflection* is used when light interacts with a planar boundary, and when the previously found law "angle of incidence equals angle of reflected light" is valid. On the other hand, if light changes its propagation direction due to encountering spherical particles, the term *scattering* is used[10]. The most obvious difference between these situations is that upon reflection of an incident beam of light, all light travels in the same direction, while scattering generally leads to a broad distribution of propagation directions. However, the physical mechanism behind both reflection and scattering is the same, namely the interaction of electromagnetic fields with the interface between two media with different refractive indices. The general boundary conditions derived at the beginning of this chapter are still valid: The tangential component of the electric field, as well as the product of $n^2$ and the perpendicular component of the electric field have to be constant across the interface. In a spherical geometry where the origin of the coordinate system coincides with the center of the sphere, the tangential field components are given by $E_\theta$ and $E_\phi$, while the perpendicular component is $E_r$. Conveniently, these are exactly the same components which are used in the definition of vector spherical harmonics. Thus, in the following, the origin of the coordinate system is always chosen to coincide with the center of the spherical interface. We will now set out to derive how an incident field made up of VSH creates a scattered and a transmitted field that are also decomposed into VSH.

Before doing any calculations, we have to choose which VSH to include in the model. As stated in section 2.1.3, VSH with different radial functions exist. Since the differential equation (2.31) describing the $r$-dependence[11] is of second order, a linear combination of two linearly independent solutions is needed to fulfill general boundary conditions. It is convenient to use $j_\ell(kr)$ and $h_\ell^1(kr)$: The former is finite at the origin, while the latter has the form of outgoing spherical waves in the far field. Furthermore, both are needed in the VSH decomposition of the field of an oscillating dipole, as seen in section 2.2.2. Since we ultimately want to describe the interaction of a dipole emitter with spherical particles, it is crucial that we can use a dipole field as the incident field. Thus, the general structure of the solution can be summarised as one linear combination of $\{\boldsymbol{M}_{\ell m}^j, \boldsymbol{N}_{\ell m}^j, \boldsymbol{M}_{\ell m}^h, \boldsymbol{N}_{\ell m}^h \mid \ell \in \mathbb{N}, m \in \mathbb{Z}, |m| \leq \ell\}$ on the inside of the spherical interface, and another such linear combination on the outside of the interface. However, not all types of VSH are always needed to fulfill the boundary conditions. This can be illustrated on the example of a sphere embedded in a medium and illuminated by different types of light sources:

---

[10]It is assumed here that the particle is smaller than the diameter of the beam of light, such that the surface cannot be approximated as one plane across the whole cross section of the beam. A counterexample where the term reflection is more appropriate is the interaction of the beam of light of a laser pointer hitting a metal sphere the size of a watermelon. There, if the surface is sufficiently smooth, all reflected light will travel in roughly the same direction.

[11]$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial R}{\partial r} \right) + \left( k^2 - \frac{\ell(\ell+1)}{r^2} \right) R = 0$ (2.31).

| light source | VSH inside sphere | | VSH outside of sphere | |
| --- | --- | --- | --- | --- |
| | $\boldsymbol{M}_{\ell m}^{j}, \boldsymbol{N}_{\ell m}^{j}$ | $\boldsymbol{M}_{\ell m}^{h}, \boldsymbol{N}_{\ell m}^{h}$ | $\boldsymbol{M}_{\ell m}^{j}, \boldsymbol{N}_{\ell m}^{j}$ | $\boldsymbol{M}_{\ell m}^{h}, \boldsymbol{N}_{\ell m}^{h}$ |
| plane wave from outside | transmitted field | - | incident field | scattered field |
| dipole emitter within sphere | scattered field | incident field | - | transmitted field |
| dipole emitter outside of sphere | transmitted field | - | incident field | scattered field |

The list of needed VSH always refers to the field directly at the interface[12]. It becomes apparent that out of the four possible types of function sets, only three are needed in a particular situation. For example, if the light source is situated outside the sphere, there is no reason why the field should become infinitely large at the center of the sphere. Thus, VSH with spherical Hankel functions are not needed inside the sphere. Contrarily, if the light source is situated inside the sphere, then the electric field far away from the sphere has to have the form of outgoing spherical waves. Thus, VSH with spherical Bessel functions are not needed outside the sphere. This observation is similar to the reflection of plane waves at a planar boundary, where generally four types of waves are defined: In each medium, there could be waves traveling towards or away from the interface. However, if we assume an incoming wave in one of the media, then outgoing waves in both media are generated (the reflected and transmitted waves), but no incoming wave in the second medium occurs. Thus, for plane waves at a planar interface, always one of the waves traveling towards the interface is missing, while the other one is the incident field. Analogously, in the VSH decomposition of the field in a spherical geometry, either $\boldsymbol{M}_{\ell m}^{h}, \boldsymbol{N}_{\ell m}^{h}$ inside the sphere or $\boldsymbol{M}_{\ell m}^{j}, \boldsymbol{N}_{\ell m}^{j}$ outside the sphere are unneeded, while the other set defines the incident field. We use the former case to now quantify the relation between incident, scattered and transmitted fields.

Imagine a sphere with radius $R$ and refractive index $n_{\mathrm{sph}}$ embedded in a homogeneous medium with refractive index $n_{\mathrm{med}}$. Let us assume that the incident field in the environment is given by $\boldsymbol{N}_{\ell m}^{j}$. Then, the conditions for the transmitted field $\boldsymbol{E}_{tr}$ inside the sphere and the scattered field $\boldsymbol{E}_{sc}$ in the environment are:

$$
\begin{aligned}
\boldsymbol{E}_{tr} \cdot \hat{\boldsymbol{e}}_{\theta} &= \left(\boldsymbol{N}_{\ell m}^{j} + \boldsymbol{E}_{sc}\right) \cdot \hat{\boldsymbol{e}}_{\theta}, \\
\boldsymbol{E}_{tr} \cdot \hat{\boldsymbol{e}}_{\phi} &= \left(\boldsymbol{N}_{\ell m}^{j} + \boldsymbol{E}_{sc}\right) \cdot \hat{\boldsymbol{e}}_{\phi}, \\
n_{\mathrm{sph}}^{2} \boldsymbol{E}_{tr} \cdot \hat{\boldsymbol{e}}_{r} &= n_{\mathrm{med}}^{2} \left(\boldsymbol{N}_{\ell m}^{j} + \boldsymbol{E}_{sc}\right) \cdot \hat{\boldsymbol{e}}_{r}.
\end{aligned} \tag{2.114}
$$

According to our previous considerations, we now define

$$
\boldsymbol{E}_{tr} = \sum_{\ell'=1}^{\infty} \sum_{m'=-\ell'}^{\ell'} a_{\ell'm'}^{tr} \boldsymbol{M}_{\ell'm'}^{j} + b_{\ell'm'}^{tr} \boldsymbol{N}_{\ell'm'}^{j}, \tag{2.115}
$$

---

[12]Recall the piecewise definition of the VSH decomposition of a dipole field: When there is a dipole inside the sphere at a distance $r_0$ from the origin, $\boldsymbol{M}_{\ell m}^{j}$ and $\boldsymbol{N}_{\ell m}^{j}$ are needed for the field at $r < r_0$. Conversely, if the dipole is outside the sphere at distance $r_0$ from the origin, the scattered field and the original dipole field are both made up of $\boldsymbol{M}_{\ell m}^{h}$ and $\boldsymbol{N}_{\ell m}^{h}$ at $r > r_0$.

$$\boldsymbol{E}_{sc} = \sum_{\ell'=1}^{\infty} \sum_{m'=-\ell'}^{\ell'} a_{\ell'm'}^{sc} \boldsymbol{M}_{\ell'm'}^{h} + b_{\ell'm'}^{sc} \boldsymbol{N}_{\ell'm'}^{h}. \tag{2.116}$$

Let us now look at the first condition of equation (2.114), using the shorthand $f_\ell'(kR) := 1/(kR) \, \mathrm{d}[rf_\ell(kr)]/\mathrm{d}r|_{r=R}$:

$$\sum_{\ell',m'} c_{\ell'm'} \left( a_{\ell'm'}^{tr} i\pi_{\ell'm'}(\theta) j_{\ell'}(k_{\mathrm{sph}}R) + b_{\ell'm'}^{tr} \tau_{\ell'm'}(\theta) j_{\ell'}'(k_{\mathrm{sph}}R) \right) \mathrm{e}^{im'\phi} = c_{\ell m} \tau_{\ell m}(\theta) j_\ell'(k_{\mathrm{med}}R) \mathrm{e}^{im\phi}$$

$$+ \sum_{\ell''m''} c_{\ell''m''} \left( a_{\ell''m''}^{sc} i\pi_{\ell''m''}(\theta) h_{\ell''}(k_{\mathrm{med}}R) + b_{\ell''m''}^{sc} \tau_{\ell''m''}(\theta) h_{\ell'}'(k_{\mathrm{med}}R) \right) \mathrm{e}^{im''\phi}. \tag{2.117}$$

This seems rather daunting. However, we know that the equation has to be fulfilled at *every set* $(\theta, \phi)$ – which means that we can make use of the orthogonality and thus the linear independence of the angle-dependent functions. For example, since

$$\int_0^{2\pi} \mathrm{e}^{im\phi} \mathrm{e}^{-im'\phi} = 2\pi\delta_{m,m'} \ , \tag{2.118}$$

the function $\mathrm{e}^{im\phi}$ cannot be expressed as a sum over exponential functions with all other $m' \neq m$,

$$\mathrm{e}^{im\phi} \neq \sum_{m'\neq m} c_{m'} \mathrm{e}^{im'\phi}. \tag{2.119}$$

Therefore, equality at all $\phi \in [0, 2\pi]$ can only be reached if the decomposition only uses terms with $m' = m = m''$. For the $\theta$-dependent functions, recall that we showed orthogonality of $(\boldsymbol{M}_{\ell,m}, \boldsymbol{N}_{\ell,m})$, $(\boldsymbol{M}_{\ell m}, \boldsymbol{M}_{\ell'm'})$ and $(\boldsymbol{N}_{\ell m}, \boldsymbol{N}_{\ell'm'})$ in section 2.1.3 (eq. 2.47-2.49). Thus, the complicated equation (2.117) reduces to:

$$b_{\ell m}^{tr} j_\ell'(k_{\mathrm{sph}}R) = j_\ell'(k_{\mathrm{med}}R) + b_{\ell m}^{sc} h_\ell'(k_{\mathrm{med}}R). \tag{2.120}$$

We now have one equation with two unknowns, so a second equation is needed to determine the expansion coefficients. For that, we turn to the radial component of the electric field and find:

$$n_{\mathrm{sph}}^2 b_{\ell m}^{tr} \frac{j_\ell(k_{\mathrm{sph}}R)}{k_{\mathrm{sph}}R} = n_{\mathrm{med}}^2 \frac{j_\ell(k_{\mathrm{med}}R)}{k_{\mathrm{med}}R} + n_{\mathrm{med}}^2 b_{\ell m}^{sc} \frac{h_\ell(k_{\mathrm{med}}R)}{k_{\mathrm{med}}R}$$

$$\Rightarrow \quad n_{\mathrm{sph}} b_{\ell m}^{tr} j_\ell(k_{\mathrm{sph}}R) = n_{\mathrm{med}} j_\ell(k_{\mathrm{med}}R) + n_{\mathrm{med}} b_{\ell m}^{sc} h_\ell(k_{\mathrm{med}}R). \tag{2.121}$$

Combined, this leads to a matrix equation similar to eq. (2.97) and (2.100):

$$\begin{pmatrix} j_\ell'(k_{\mathrm{sph}}R) & -h_\ell'(k_{\mathrm{med}}R) \\ n_{\mathrm{sph}} j_\ell(k_{\mathrm{sph}}R) & -n_{\mathrm{med}} h_\ell(k_{\mathrm{med}}R) \end{pmatrix} \begin{pmatrix} b_{\ell m}^{tr} \\ b_{\ell m}^{sc} \end{pmatrix} = \begin{pmatrix} j_\ell'(k_{\mathrm{med}}R) & -h_\ell'(k_{\mathrm{sph}}R) \\ n_{\mathrm{med}} j_\ell(k_{\mathrm{med}}R) & -n_{\mathrm{sph}} h_\ell(k_{\mathrm{sph}}R) \end{pmatrix} \begin{pmatrix} b_{\ell m}^{i,\mathrm{med}} \\ b_{\ell m}^{i,\mathrm{sph}} \end{pmatrix}, \tag{2.122}$$

where we now allowed an arbitrary prefactor $b_{\ell m}^{i,\mathrm{med}}$ for the initial wave in the environment, and also added an initial wave $b_{\ell m}^{i,\mathrm{sph}} \boldsymbol{N}_{\ell m}^h$ inside the sphere to account for the other

possible illumination scenarios mentioned above. The same procedure leads to an equation for VSH of type $M_{\ell m}^{j,h}$, with the one difference that there the condition $B_\theta^{\text{sph}} = B_\theta^{\text{med}}$, $B_\phi^{\text{sph}} = B_\phi^{\text{med}}$ (with $\boldsymbol{B} = \nabla \times \boldsymbol{E}/(ik_0)$ and $\nabla \times \boldsymbol{M}_{\ell m}^f = k\boldsymbol{N}_{\ell m}^f$) has to be used as second equation because $\boldsymbol{M}_{\ell m}^f$ does not have an $r$-component:

$$\begin{pmatrix} j_\ell(k_{\text{sph}}R) & -h_\ell(k_{\text{med}}R) \\ n_{\text{sph}}j_\ell'(k_{\text{sph}}R) & -n_{\text{med}}h_\ell'(k_{\text{med}}R) \end{pmatrix}\begin{pmatrix} a_{\ell m}^{tr} \\ a_{\ell m}^{sc} \end{pmatrix} = \begin{pmatrix} j_\ell(k_{\text{med}}R) & -h_\ell(k_{\text{sph}}R) \\ n_{\text{med}}j_\ell'(k_{\text{med}}R) & -n_{\text{sph}}h_\ell'(k_{\text{sph}}R) \end{pmatrix}\begin{pmatrix} a_{\ell m}^{i,\text{med}} \\ a_{\ell m}^{i,\text{sph}} \end{pmatrix}. \tag{2.123}$$

Thus, the expansion coefficients for the scattered and transmitted fields can be found by simple matrix operations. Since these equations are independent for different sets $(\ell, m)$ and for VSH type $\boldsymbol{M}$ and $\boldsymbol{N}$, one can calculate a "lookup table" for given refractive indices $n_{\text{sph}}$, $n_{\text{med}}$ and given sphere radius $R$. Then, the interaction of any initial field that has been decomposed into VSH with the sphere can easily be determined.

As a side note, we want to mention that this method scales readily to the case of several concentric shells with different refractive indices $n_1, n_2, n_3, \ldots$ and radii $r_1, r_2, r_3, \ldots$ and the environment $n_{\text{med}}$. Then, the conditions treated above have to be valid at each single interface, and one simply gets larger matrices.

## 2.3.4 Dipole emitter above a spherical interface

We will now apply the formulas derived above to the case of a dipole emitter close to a spherical interface. It was stated in section 2.2.2 that the electric field of a dipole emitter situated at $\boldsymbol{r}_0$ is given by

$$\boldsymbol{E}_{\text{VSH}}(\boldsymbol{r}) = \sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}^D \boldsymbol{M}_{\ell m}^f(k\boldsymbol{r}) + b_{\ell m}^D \boldsymbol{N}_{\ell m}^f(k\boldsymbol{r}) \tag{2.124}$$

with the expansion coefficients

$$a_{\ell m}^D = \frac{4\pi i n k_v^3}{\ell(\ell+1)}\boldsymbol{p}\cdot\overline{\boldsymbol{M}_{\ell m}^g(k\boldsymbol{r}_0)} \quad\text{and}\quad b_{\ell m}^D = \frac{4\pi i n k_v^3}{\ell(\ell+1)}\boldsymbol{p}\cdot\overline{\boldsymbol{N}_{\ell m}^g(k\boldsymbol{r}_0)}, \tag{2.125}$$

and with $f = h^1, g = j$ for $r > r_0$ and $f = j, g = h^2$ for $r < r_0$. Assuming that the sphere with radius $R$ is situated at the origin and the dipole is located outside the sphere ($r_0 > R$), the field incident at the surface of the sphere is decomposed into VSH with radial dependence $j$. Thus, to compare with the notation of equations (2.122) and (2.123), we have

$$\begin{aligned} a_{\ell m}^{i,\text{med}} &= a_{\ell m}^D, & a_{\ell m}^{i,\text{sph}} &= 0, \\ b_{\ell m}^{i,\text{med}} &= b_{\ell m}^D, & b_{\ell m}^{i,\text{sph}} &= 0. \end{aligned} \tag{2.126}$$

We can then directly solve for the coefficients $a_{\ell m}^{tr}, b_{\ell m}^{tr}$ of the field inside the sphere and $a_{\ell m}^{sc}, b_{\ell m}^{sc}$ of the scattered field. As an example, see figure 2.13, where a dipole emitter oriented parallel to the $z$-axis resides half a wavelength above the surface of a sphere with radius $0.4\,\lambda$. It becomes obvious that the resulting field depends strongly on the
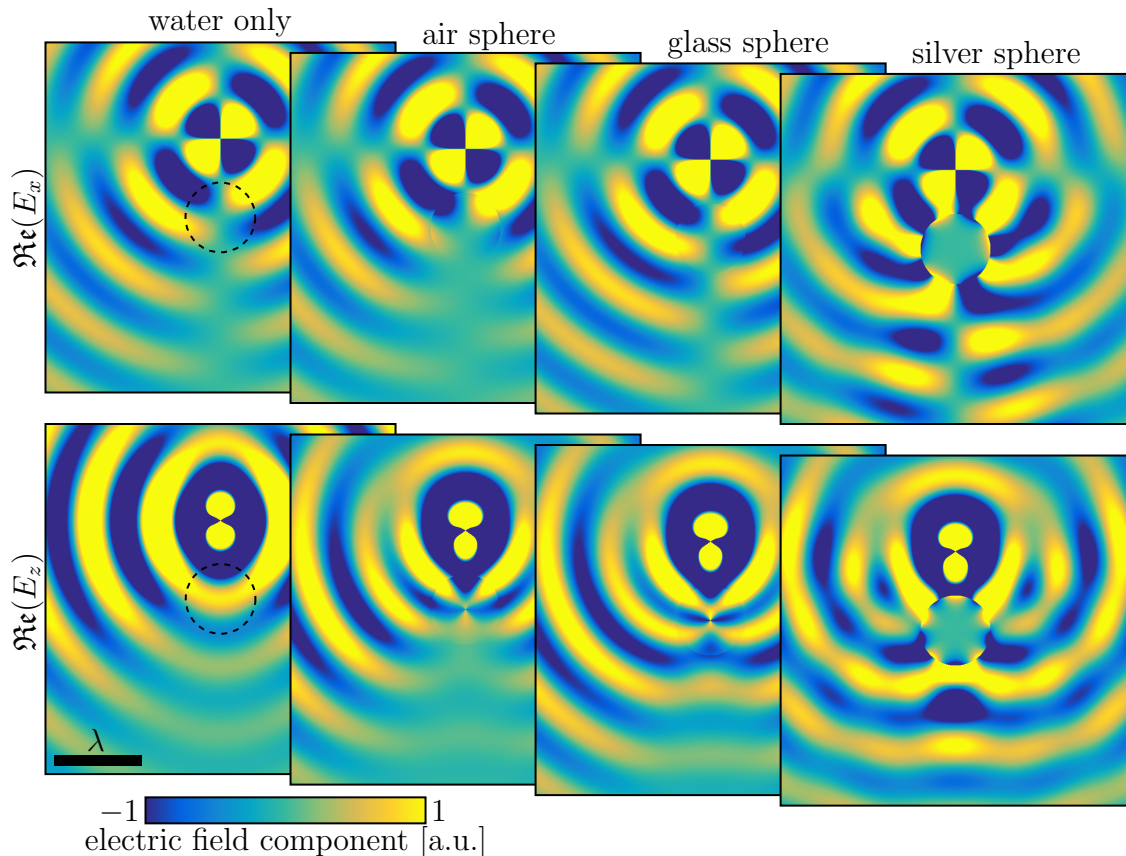
**Figure 2.13:** Electric field of a dipole emitter close to a sphere of radius $R = 200\,\text{nm}$ made of water, air, glass or silver and embedded in water. The dipole is oriented parallel to the $z$-axis and placed at $\boldsymbol{r}_0 = (0, 0, 450)\,\text{nm}$, while the sphere is centered at the origin. The scalebar corresponds to one wavelength, i.e. to $\lambda = 500\,\text{nm}$.

material of the sphere – air, glass or silver. The refractive index of water ($n = 1.33$) is closest to that of glass ($n = 1.52$), which is why the distortion of the field is smallest in this case. For glass and air ($n = 1.0$), the field penetrates the sphere and is focused at its center. On the other hand, silver ($n = 0.05 + 2.80i$ at $\lambda = 500\,\text{nm}$) absorbs light, which is why the field only penetrates some tens of nanometers into the sphere. Finally, the water only sample shows that $E_x$ is almost zero close to the origin, while $E_z$ is large in this region. Accordingly, the $x$-component of the field is less strongly influenced by the presence of the sphere than the $z$-component. For the latter, the highest amplitudes shift from $\theta = \pi/2$ relative to the dipole (water only) to larger $\theta$-values, "bending" the emission around the sphere. This will be discussed in more detail in section 2.4.6, which investigates a dipole's energy flux.

## 2.4 Energy flux of an electromagnetic field

A fluorescent molecule emits light and thus energy – a process that will be explained in more detail in section 2.5.1. Both the spatial distribution of the energy flux and the total rate of energy emission are of practical interest. The former determines the detection efficiency of the emitted light (see section 4.3.1) and can be used to predict the patterns that arise when taking a defocused image of a dipole emitter [52]. The latter is connected to the excited state lifetime of the emitter (see chapter 2.5), which was measured and used to infer properties such as cell-substrate distances in the course of this work.

In the most general form, the energy flux of an electromagnetic field, i.e. the rate of energy flow through a reference surface of unit area, is given by the Poynting vector $\tilde{\boldsymbol{S}}(\boldsymbol{r}, t) = (c/4\pi)\boldsymbol{E}_{\mathrm{phys}}(\boldsymbol{r}, t) \times \boldsymbol{B}_{\mathrm{phys}}(\boldsymbol{r}, t)$, where $\boldsymbol{E}_{\mathrm{phys}}$ and $\boldsymbol{B}_{\mathrm{phys}}$ are the physical (real!) fields. When using the complex-valued temporal Fourier amplitudes $\boldsymbol{E}$ and $\boldsymbol{B}$ (as we have during this whole text), the Poynting vector averaged over one oscillation period is given by

$$
\begin{aligned}
\boldsymbol{S}(\boldsymbol{r}) &= \frac{c}{8\pi}\mathfrak{Re}\left(\boldsymbol{E} \times \overline{\boldsymbol{B}}\right) \\
&= \frac{c}{8\pi}\mathfrak{Re}\left(\boldsymbol{E} \times \overline{\left[\frac{1}{ik_v}\nabla \times \boldsymbol{E}\right]}\right) \\
&= -\frac{c}{8\pi k_v}\mathfrak{Im}\left(\boldsymbol{E} \times \overline{\left[\nabla \times \boldsymbol{E}\right]}\right),
\end{aligned}
\tag{2.127}
$$

where we used the fact that $\boldsymbol{B} = \nabla \times \boldsymbol{E}/ik_v$ (third of Maxwell's equations). This simple relation can now be filled with life by considering the energy flux in several different situations.

### 2.4.1 Energy flux of a plane wave

We start by calculating the energy flux of a single plane wave,

$$
\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}_0 \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}.
\tag{2.128}
$$

Then, $\nabla \times \boldsymbol{E} = i\boldsymbol{k} \times \boldsymbol{E}$, and with $\boldsymbol{k} = k\hat{\boldsymbol{e}}_k$,

$$
\boldsymbol{E} \times \left[\overline{\nabla \times \boldsymbol{E}}\right] = -i\overline{k}\left[\hat{\boldsymbol{e}}_k|\boldsymbol{E}|^2 - \boldsymbol{E}(\hat{\boldsymbol{e}}_k \cdot \boldsymbol{E})\right].
\tag{2.129}
$$

In ideal dielectric media, the refractive index $n$ is real and thus $k$ is also real. Since we saw in section 2.1.2 that $\boldsymbol{k} \cdot \boldsymbol{E} = 0$, the time-averaged Poynting vector of a plane wave in a dielectric medium (such as air or water) is simply given by

$$
\boldsymbol{S}_{\mathrm{PW}}(\boldsymbol{r}) = \frac{cn}{8\pi}|\boldsymbol{E}_0|^2\hat{\boldsymbol{e}}_k.
\tag{2.130}
$$

From this, we can conclude that:

1. The time-averaged energy flux of an infinitely extended plane wave is identical everywhere in space.

2. The energy flows in the direction of $\hat{\boldsymbol{e}}_k$. This seems reasonable, since $\boldsymbol{k}$ points in the direction of propagation of the wave fronts.

It is also quite enlightening to compare the result for $\boldsymbol{S}_{\mathrm{PW}}$ with the energy density of a plane wave. The energy stored in an electromagnetic field per unit volume is given by

$$\mathcal{U} = \frac{1}{8\pi} \left( \varepsilon \boldsymbol{E}_{\mathrm{phys}} \cdot \boldsymbol{E}_{\mathrm{phys}} + \boldsymbol{B}_{\mathrm{phys}} \cdot \boldsymbol{B}_{\mathrm{phys}} \right), \tag{2.131}$$

or, when again using the complex-valued Fourier coefficients and averaging over one oscillation period,

$$\langle \mathcal{U} \rangle = \frac{1}{16\pi} \left( \varepsilon \boldsymbol{E} \cdot \overline{\boldsymbol{E}} + \boldsymbol{B} \cdot \overline{\boldsymbol{B}} \right). \tag{2.132}$$

For a plane wave in a lossless dielectric ($n = \sqrt{\varepsilon} \in \mathbb{R}$), this is simply

$$\langle \mathcal{U}_{\mathrm{PW}} \rangle = \frac{1}{16\pi} \left( n^2 |\boldsymbol{E_0}|^2 + \frac{1}{k_v^2} |\nabla \times \boldsymbol{E}|^2 \right) = \frac{n^2}{8\pi} |\boldsymbol{E_0}|^2 \tag{2.133}$$

A comparison with equation (2.130) shows that

$$\boldsymbol{S}_{\mathrm{PW}} = \frac{c}{n} \langle \mathcal{U}_{\mathrm{PW}} \rangle \, \hat{\boldsymbol{e}}_k. \tag{2.134}$$

But $c/n$ is simply the speed of light in the medium with refractive index $n$. Thus, for a plane wave, the energy flux (energy per time and area) equals the product of the energy density (energy per volume) and the speed of propagation.

## 2.4.2 Energy flux of evanescent and lossy waves

We have introduced evanescent waves in the context of total internal reflection in section 2.3.1, where we already stated that these waves do not transport any energy perpendicular to the surface. We will prove this claim now. Assume that the medium is lossless, i.e. $n \in \mathbb{R}$, and that the interface where total internal reflection occurred is at $z = 0$. Then,

$$\boldsymbol{E} = (E_p \hat{\boldsymbol{e}}_p + E_s \hat{\boldsymbol{e}}_s) \exp(iq_x x + iq_y y - w'' z),$$

$$\boldsymbol{B} = \frac{1}{ik_v} \nabla \times \boldsymbol{E} = n(E_p \hat{\boldsymbol{e}}_s - E_s \hat{\boldsymbol{e}}_p) \exp(iq_x x + iqy_y - w'' z), \tag{2.135}$$

where $(q_x, q_y) = q(\cos\phi, \sin\phi)$, $q, w'', \phi \in \mathbb{R}$ and $k^2 = n^2 k_v^2 = q^2 - w''^2$. The relevant unit vectors for calculating the Poynting vector are

$$\hat{\boldsymbol{e}}_k = \frac{1}{k} \begin{pmatrix} q\cos\phi \\ q\sin\phi \\ iw'' \end{pmatrix}, \quad \hat{\boldsymbol{e}}_p = \frac{1}{k} \begin{pmatrix} -iw''\cos\phi \\ -iw''\sin\phi \\ q \end{pmatrix}, \quad \hat{\boldsymbol{e}}_s = \begin{pmatrix} \sin\phi \\ -\cos\phi \\ 0 \end{pmatrix}. \tag{2.136}$$
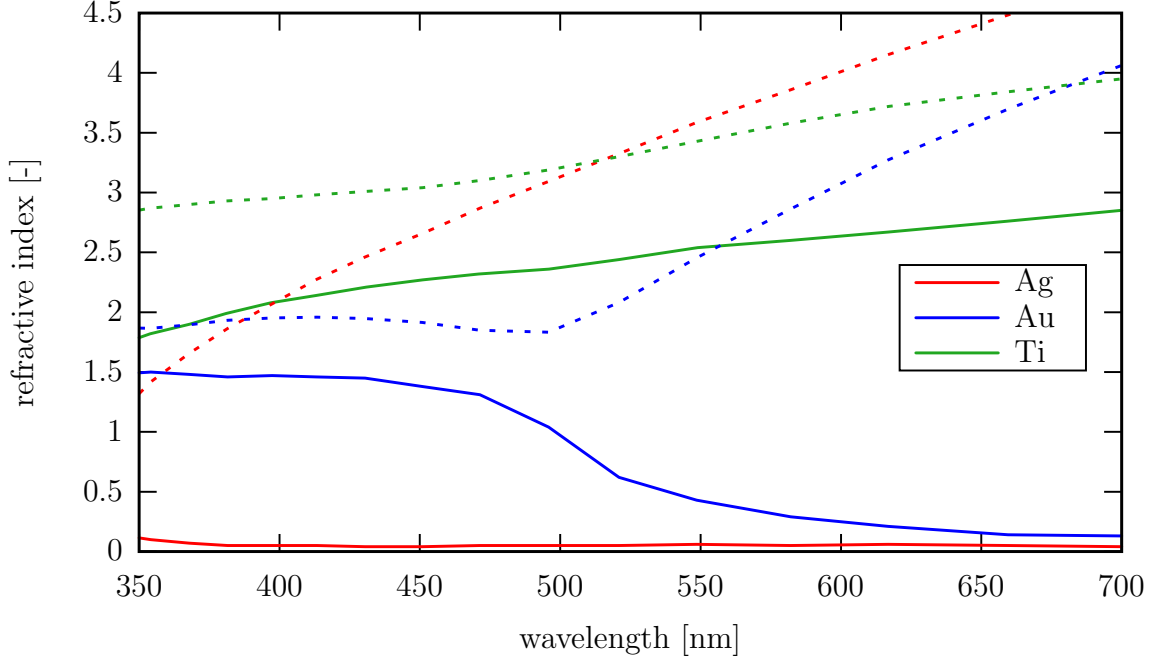
**Figure 2.14:** Refractive indices of silver, gold and titanium: real parts depicted by solid lines, imaginary parts by dashed lines. Data originally from [53, 54], downloaded from `https://refractiveindex.info`.

Thus, $\overline{\hat{e}_s} = \hat{e}_s$ is invariant under complex conjugation, while $\overline{\hat{e}_k} \neq \hat{e}_k$ and $\overline{\hat{e}_p} \neq \hat{e}_p$. The complex conjugated unit vectors are still orthonormal, however, $\overline{\hat{e}_p} \nparallel \hat{e}_p$:

$$\hat{e}_p \times \overline{\hat{e}_p} = \frac{1}{|k|^2} \begin{pmatrix} -2iw''q\sin\phi \\ 2iw''q\cos\phi \\ 0 \end{pmatrix}. \tag{2.137}$$

The time-averaged Poynting vector is then given by:

$$\begin{aligned}
\boldsymbol{S}_{\text{eva.}} &= \frac{c}{8\pi}\mathfrak{Re}\left(\boldsymbol{E} \times \overline{\boldsymbol{B}}\right) \\
&= \frac{c}{8\pi}\mathfrak{Re}\left[\overline{n}\mathrm{e}^{-2w''z}(E_p\hat{e}_p + E_s\hat{e}_s) \times (\overline{E_p}\hat{e}_s - \overline{E_s}\overline{\hat{e}_p})\right] \\
&= \frac{c}{8\pi}\mathfrak{Re}\left[\overline{n}\mathrm{e}^{-2w''z}(|E_p|^2\hat{e}_k - E_p\overline{E_s}(\hat{e}_p \times \overline{\hat{e}_p}) + |E_s|^2\overline{\hat{e}_k})\right] \\
&= \frac{cn}{8\pi}\mathrm{e}^{-2w''z}\left[(|E_p|^2 + |E_s|^2)\mathfrak{Re}(\hat{e}_k) - i\mathfrak{Im}(E_p\overline{E_s})(\hat{e}_p \times \overline{\hat{e}_p})\right], \tag{2.138}
\end{aligned}$$

where the last equality takes into account that $n \in \mathbb{R}$ and that $(\hat{e}_p \times \overline{\hat{e}_p})$ is purely imaginary. We can see that generally, the Poynting vector of an evanescent wave is not zero. However, its magnitude decays exponentially with the distance $z$ from the interface. Furthermore, by recognizing that the $z$-component of $(\hat{e}_p \times \overline{\hat{e}_p})$ equals zero and that the $z$-component of $\hat{e}_k$ is purely imaginary, we prove what we set out to show in this section: The $z$-component of $S_{\text{eva.}}$ vanishes.

The situation is quite different if we study the wave that results when a plane wave

encounters the interface between a lossless dielectric ($n_0 \in \mathbb{R}$) and a metal ($n = n' + in''$). In the lossless medium, the component of $\boldsymbol{k}$ parallel to the interface ($q$) is real. Since we saw that $q$ is constant on both sides of the interface, it is still real in the metal. For the $z$-component of $\boldsymbol{k}$, $w$, this means:

$$w = \sqrt{n^2 k_v^2 - q^2} = \sqrt{(n'^2 k_v^2 - n''^2 k_v^2 - q^2) + 2in'n''k_v^2} = w' + iw''. \qquad (2.139)$$

For metals in wavelength regions typical for fluorescence microscopy, the real and imaginary parts of the refractive index fulfill $0 < n' < n''$, as can be seen for some examples in figure 2.14. Thus, the term under the square root has a negative real and a positive imaginary part, which means that $0 < w' < w''$[13]. Using the same argument as for (2.138), but replacing $iw''$ by $w' + iw''$, we find for the flux along the $z$-direction in a metal:

$$\begin{aligned}
\boldsymbol{S}_{\mathrm{metal},z} &= \frac{c}{8\pi}\mathfrak{Re}\left[\overline{n}\mathrm{e}^{-2w''z}(|E_p|^2\hat{\boldsymbol{e}}_k - E_p\overline{E_s}(\hat{\boldsymbol{e}}_p \times \overline{\hat{\boldsymbol{e}}_p}) + |E_s|^2\overline{\hat{\boldsymbol{e}}_k})\right]_z \\
&= \frac{c}{8\pi k_v}\mathrm{e}^{-2w''z}\,\mathfrak{Re}\left[\frac{\overline{n}w}{n}|E_p|^2 + \frac{\overline{nw}}{\overline{n}}|E_s|^2\right] \\
&= \frac{c}{8\pi k_v}\mathrm{e}^{-2w''z}\left[\frac{w'(n'^2 - n''^2) + 2w''n'n''}{|n|^2}|E_p|^2 + w'|E_s|^2\right]. \qquad (2.140)
\end{aligned}$$

As a simple consistency check, we want to remark here that for real $n$, i.e. $n'' = 0$, the result for $\boldsymbol{S}_{\mathrm{metal},z}$ equals the projection $\hat{\boldsymbol{e}}_z \cdot \boldsymbol{S}_{\mathrm{PW}}$, see equation (2.130). Since we assume that the wave described here was excited by a plane wave impinging on the boundary between a lossless dielectric and the metal, the flux along $z$ should be larger than or equal to zero. The flux of the $s$-wave is always positive since $w' > 0$. For the $p$-wave, this is slightly harder to see, thus we investigate the $q$-dependence of the term. Since $q$ cannot be negative, we start with $q = 0$. Then, $w = \sqrt{n^2 k_v^2 - q^2} = nk_v$. In this case, the term in front of $|E_p|^2$ reduces to $k_v n'(n'^2 + n''^2)/|n|^2$, which is larger than zero. For increasing $q$, $w''$ increases while $w'$ approaches zero. Thus, the prefactor remains positive.

A further condition is that, in the absence of energy sources or sinks, the energy flux along one direction should be constant. The expression (2.140), however, decreases exponentially with distance $z$ from the interface. Thus, we have to assume that there are energy sinks in a metal. What are these sinks? The electric field exerts a force on the unbound conduction electrons, which are forced to oscillate. Thus, by accelerating electrons the electric field performs work. The electrons themselves interact with the lattice and ultimately transform their kinetic energy to heat. This is summed up by the continuity equation – also known as *Poynting's theorem* – which links the change in energy density to the energy flux and the work done by the fields:

$$\frac{\partial \mathcal{U}}{\partial t} = -\nabla \cdot \boldsymbol{S} - \boldsymbol{j} \cdot \boldsymbol{E}. \qquad (2.141)$$

---

[13]Since the square root of the complex variable $z = |z|\mathrm{e}^{i\phi}$ is given by $\sqrt{z} = \sqrt{|z|}\mathrm{e}^{i\phi/2}$, an initial angle $\pi/2 < \phi < \pi$ leads to a final angle $\pi/4 < \phi/2 < \pi/2$.

Here, $\boldsymbol{j}$ is a general electric current density. For a plane wave hitting a metal, $\boldsymbol{j}$ is the current induced by the field $\boldsymbol{E}$ itself. Since the energy density $\mathcal{U}$ (averaged over one oscillation period) is constant over time, the spatial variance of the Poynting vector is caused by the work done on the electrons. This is the reason why electromagnetic fields only penetrate a short distance into metals.

### 2.4.3 Energy flux of a plane wave near a planar interface

When a plane wave strikes a planar interface between two media with refractive indices $n_1$ and $n_2$, some energy gets reflected back into the first medium, while some gets transmitted to the second medium. In this section, we want to derive the energy balance for a number of different materials and angles of incidence. We will start in the upper halfspace by considering the initial field $\boldsymbol{E}_1^+$ and the reflected field $\boldsymbol{E}_1^-$:

$$
\begin{aligned}
\boldsymbol{E}_1^+ &= (\quad E_p \hat{\boldsymbol{e}}_p^+ + \quad E_s \hat{\boldsymbol{e}}_s) \exp(iq_x x + iq_y y + iw_1 z), \\
\boldsymbol{E}_1^- &= (r_p E_p \hat{\boldsymbol{e}}_p^- + r_s E_s \hat{\boldsymbol{e}}_s) \exp(iq_x x + iq_y y - iw_1 z).
\end{aligned}
\tag{2.142}
$$

Since the energy flux is related to the square of the amplitude of the electric field, we cannot simply calculate $\boldsymbol{S}$ for the incident and reflected field separately and then add them up. Instead, we have to determine the total electric field $\boldsymbol{E}_1 = \boldsymbol{E}_1^+ + \boldsymbol{E}_1^-$ and the total magnetic field $\boldsymbol{B}_1 = (\nabla \times \boldsymbol{E}_1)/ik_v$ and use them to find the Poynting vector:

$$
\begin{aligned}
\boldsymbol{S}_1 &= \frac{c}{8\pi k_v} \mathfrak{Re} \left\{ \boldsymbol{E}_1 \times \overline{\left[ \boldsymbol{k}_1^+ \times (E_p \hat{\boldsymbol{e}}_p^+ + E_s \hat{\boldsymbol{e}}_s) + \boldsymbol{k}_1^- \times (r_p E_p \hat{\boldsymbol{e}}_p^- + r_s E_s \hat{\boldsymbol{e}}_s) \right] e^{iq_x x + iq_y y + iw_1 z}} \right\} \\
&= \frac{c}{8\pi k_v} \mathfrak{Re} \left\{ \boldsymbol{E}_1 \times \overline{k_1 \left[ E_p \hat{\boldsymbol{e}}_s - E_s \hat{\boldsymbol{e}}_p^+ + r_p E_p \hat{\boldsymbol{e}}_s - r_s E_s \hat{\boldsymbol{e}}_p^- \right] e^{iq_x x + iq_y y + iw_1 z}} \right\}.
\end{aligned}
\tag{2.143}
$$

We assume that the first medium is a lossless dielectric ($n_1 \in \mathbb{R}$), $q$ is real and $w_1$ is either real (propagating PW) or purely imaginary (evanescent wave). By inserting the definitions

$$
\hat{\boldsymbol{e}}_p^\pm = \frac{1}{k} \begin{pmatrix} \mp w \cos\phi \\ \mp w \sin\phi \\ q \end{pmatrix}, \quad \hat{\boldsymbol{e}}_s = \begin{pmatrix} \sin\phi \\ -\cos\phi \\ 0 \end{pmatrix}, \quad \hat{\boldsymbol{e}}_k^\pm = \frac{1}{k} \begin{pmatrix} q \cos\phi \\ q \sin\phi \\ \pm w \end{pmatrix},
\tag{2.144}
$$

one can easily show that:

$$
\hat{\boldsymbol{e}}_p^+ \times \overline{\hat{\boldsymbol{e}}_p^+} = \frac{1}{|k|^2} \begin{pmatrix} q \sin\phi (\overline{w} - w) \\ q \cos\phi (w - \overline{w}) \\ 0 \end{pmatrix},
$$

$$
\hat{\boldsymbol{e}}_p^- \times \overline{\hat{\boldsymbol{e}}_p^-} = \frac{1}{|k|^2} \begin{pmatrix} q \sin\phi (w - \overline{w}) \\ q \cos\phi (\overline{w} - w) \\ 0 \end{pmatrix},
$$

$$
\hat{\boldsymbol{e}}_p^+ \times \overline{\hat{\boldsymbol{e}}_p^-} = \frac{1}{|k|^2} \begin{pmatrix} 0 \\ q \cos\phi (\overline{w} + w) \\ 0 \end{pmatrix},
$$

$$
\hat{\boldsymbol{e}}_p^\pm \times \hat{\boldsymbol{e}}_s = \hat{\boldsymbol{e}}_k^\pm,
\tag{2.145}
$$

and we already know that $\overline{\hat{e}_s} = \hat{e}_s$. We are interested in the fraction of energy that is reflected or transmitted at the interface, i.e. in the flux perpendicular to the surface. Therefore, we now proceed to determine the $z$-component of the Poynting vector. As we have just seen, the $z$-components of all mixed $p$-waves are zero, while e.g.

$$\hat{e}_z \cdot (\hat{e}_p^+ \times \overline{\hat{e}_s}) = \hat{e}_z \cdot (\hat{e}_p^+ \times \hat{e}_s) = \quad \hat{e}_z \cdot \hat{e}_k^+ = w/k,$$
$$\hat{e}_z \cdot (\hat{e}_s \times \overline{\hat{e}_p^-}) = \hat{e}_z \cdot (\overline{\hat{e}_s \times \hat{e}_p^-}) = -\hat{e}_z \cdot \overline{\hat{e}_k^-} = \overline{w/k} \text{ etc.} \qquad (2.146)$$

Inserting into equation (2.143), this leads to the much simpler formula:

$$S_{1,z} = \frac{c}{8\pi k_v} \mathrm{e}^{-2\mathfrak{Im}(w_1)z} \, \mathfrak{Re} \left\{ E_p \overline{k_1(E_p + r_p E_p)} \frac{w_1}{k_1} + E_s \overline{k_1(E_s - r_s E_s)} \frac{\overline{w_1}}{\overline{k_1}} \right.$$
$$\left. - r_p E_p \overline{k_1(E_p + r_p E_p)} \frac{w_1}{k_1} + r_s E_s \overline{k_1(E_s - r_s E_s)} \frac{\overline{w_1}}{\overline{k_1}} \right\}$$
$$= \frac{c}{8\pi k_v} \mathrm{e}^{-2\mathfrak{Im}(w_1)z} \, \mathfrak{Re} \left\{ w_1 |E_p|^2 \left( 1 - |r_p|^2 - 2i\mathfrak{Im}[r_p] \right) \right.$$
$$\left. + \overline{w_1} |E_s|^2 \left( 1 - |r_s|^2 + 2i\mathfrak{Im}[r_s] \right) \right\}, \qquad (2.147)$$

where we took into account that $k_1 \in \mathbb{R}$. Let us now interpret this result, first for pure propagating $p$- or $s$-waves with $w_1 \in \mathbb{R}$, $E_s = 0$ or $E_p = 0$. We start by comparing with the flux of the same wave in medium 1 without the presence of the interface, $S_{z,p/s}^{\mathrm{FS}} = (cw_1)/(8\pi k_v)|E_{p/s}|^2$. Then,

$$S_{1,z,p/s}^{\mathrm{prop}} = (1 - |r_{p/s}|^2) S_{z,p/s}^{\mathrm{FS}}. \qquad (2.148)$$

This is an expression for the net flux in $z$-direction, but we can interpret it as the sum of the fluxes of two waves: The incident wave with flux $S_{z,p/s}^{\mathrm{FS}}$ in the positive $z$-direction and the reflected wave with flux $|r_{p/s}|^2 S_{z,p/s}^{\mathrm{FS}}$ in the negative $z$-direction. Thus, the *reflectance* $R$ – the ratio of the reflected to the initial intensity – of a propagating $p$- or $s$-wave is given by

$$R_{p/s} = |r_{p,s}|^2. \qquad (2.149)$$

On the other hand, if we assume that medium 2 is a lossless material, too, and that $n_2 > n_1$ (i.e. $w_2 \in \mathbb{R}$), the flux in medium 2 is given by:

$$\boldsymbol{S}_2^{\mathrm{prop}} = \frac{cn_2}{8\pi} \left( |t_p E_p|^2 + |t_s E_s|^2 \right) \hat{e}_{k2}, \qquad (2.150)$$

which we found by setting $\boldsymbol{E}_0 = t_p E_p \hat{e}_{p2} + t_s E_s \hat{e}_{s2}$ in equation (2.130). When comparing the total amplitudes of the Poynting vectors of the incident and the transmitted waves, i.e. the total intensities of the plane waves, we find the *transmittance* T:

$$T_{p/s} = |t_{p,s,}|^2. \qquad (2.151)$$

Since there are no energy sources or sinks directly at the interface, the flux through the surface – i.e. the $z$-component of the fluxes – has to be identical in both media. In the second medium, this flux is given for pure $p$- and $s$-waves by

$$S_{2,z,p/s}^{\text{prop}} = \frac{c}{8\pi k_v} w_2 |t_{p,s}|^2 \, |E_{p,s}|^2 = \frac{w_2}{w_1}|t_{p,s}|^2 \, S_{z,p/s}^{\text{FS}}. \tag{2.152}$$

Thus, the condition $S_{1,z,p/s}^{\text{prop}} \overset{!}{=} S_{2,z,p/s}^{\text{prop}}$ is identical to the relation

$$1 - R_{p/s} \overset{!}{=} \frac{w_2}{w_1}T_{p/s}. \tag{2.153}$$

This formula can be found in many standard textbooks. We can confirm that it is true by inserting the definitions of $r_{s/p}$ and $t_{s/p}$ from equation (2.103).
If the incident wave in medium 1 is evanescent, i.e. $w_1 = iw''$, $0 < w'' \in \mathbb{R}$, equation (2.147) gives a different result:

$$S_{1,z}^{\text{evan}} = \frac{cw''}{4\pi k_v}\mathrm{e}^{-2w''z}\left(|E_p|^2\mathfrak{Im}[r_p] + |E_s|^2\mathfrak{Im}[r_s]\right) \tag{2.154}$$

We can directly draw some interesting conclusions. Firstly, if no interface was present, both $r_p$ and $r_s$ would be zero, which would lead to $S_{1,z}^{\text{evan}} = 0$ – exactly the result we found for evanescent waves in section 2.4.2. Secondly, a closer examination of the expressions for the two reflection coefficients

$$r_p = \frac{(n_2/n_1) - (w_2/w_1)}{(n_2/n_1)^2 + (w_2/w_1)} \quad \text{and} \quad r_s = \frac{1 - (w_2/w_1)}{1 + (w_2/w_1)} \quad \text{(see 2.103)}$$

reveals that for real $n_1, n_2$ and imaginary $w_1$, the imaginary part of $r_p, r_s$ is only nonzero for real $w_2$. Physically, this means that the combination of an incident and a reflected evanescent wave can only transport energy along $z$ if the wave becomes a propagating wave in the second medium. This can also be seen by expressing the conservation of energy at the interface ($z = 0$) as:

$$2w''\mathfrak{Im}(r_{p/s}) \overset{!}{=} \mathfrak{Re}(w_2)|t_{p,s}|^2. \tag{2.155}$$

This formula can usually not be found in textbooks, since they seldomly treat evanescent waves except for total internal reflection. In this context, it does not make sense to talk about reflectivity, since the incident evanescent wave would not transport energy in free space, anyway. We illustrate the results found so far in figure 2.15, which shows the fluxes $S_{1,z}$ and $S_{2,z}$ at an air/water interface for a wide range of $q$-values. Up to the black vertical line at $q = k_1$, the waves are propagating in both air and water. Directly at this line, $w_1 = 0$ and the wave is parallel to the interface. In this case, no energy is transported perpendicular to the interface. Between the black and orange vertical lines, the waves are evanescent in air but propagating in water, leading to a large energy flux from one medium to the other. At even larger $q$, the waves are evanescent in both media and no energy is transported along $z$ anymore.
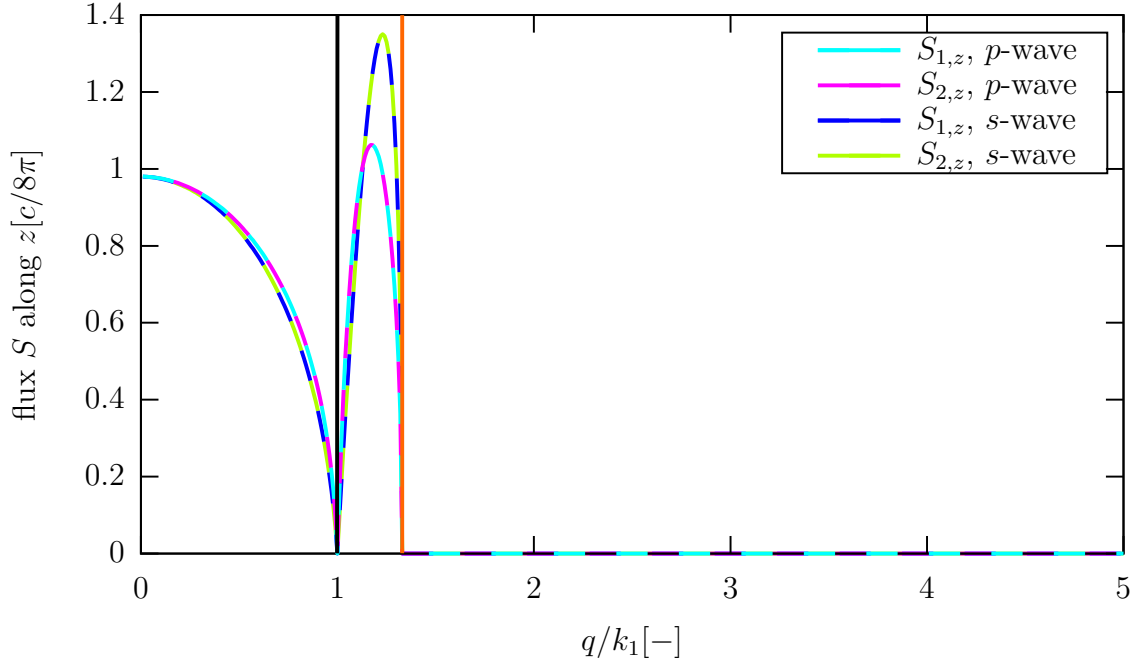
**(a)** Interface between air ($n_1 = 1$) and water ($n_2 = 1.33$).
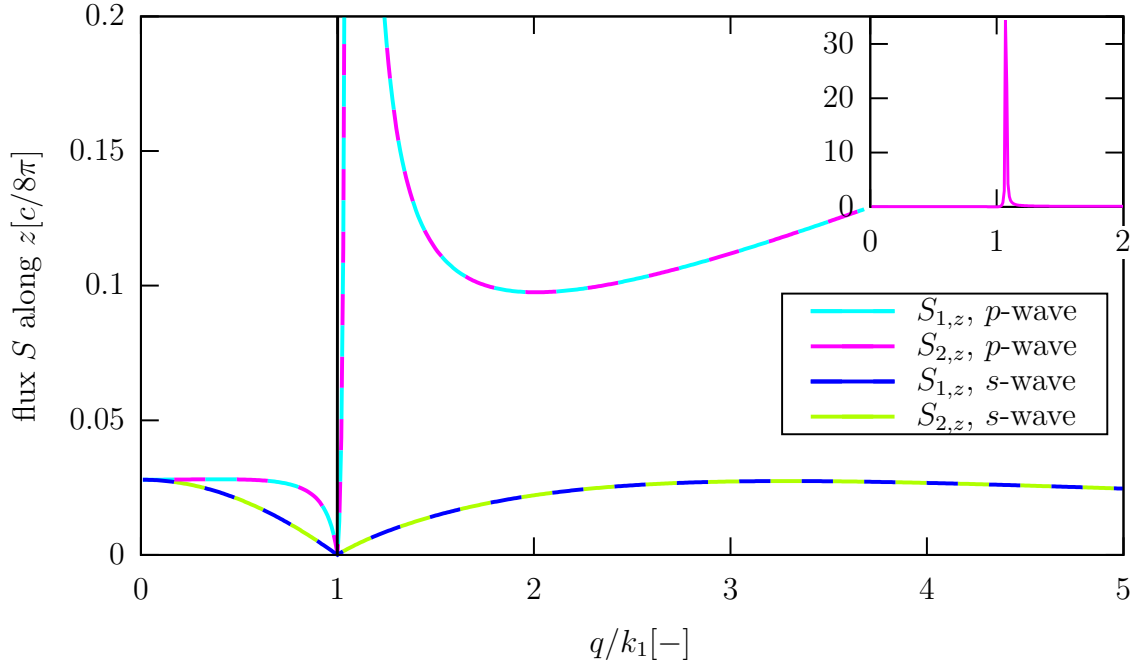


**(b)** Interface between air ($n_1 = 1$) and silver ($n_2 = 0.06 + 4.11i$ at $\lambda = 612\,\text{nm}$).

**Figure 2.15:** Energy flux $S_z$ parallel to the $z$-axis at both sides of an interface for $p$- and $s$-waves, plotted for various values of the component $q$ of the wavevector parallel to the interface. The perfect match of the corresponding pairs of fluxes in medium 1 and 2 proves the conservation of energy at the interface. The inset in (b) only shows $S_{2,z}$ because a dashed line would have been badly visible, $S_{1,z} = S_{2,z}$ in this case, too. The black vertical line at $q = k_1$ denotes the transition from propagating to evanescent waves in air. The orange vertical line in (a) denotes the transition from propagating to evanescent waves in water.

Finally, we want to investigate what happens if a plane wave reaches the interface between a lossless dielectric ($n_1 \in \mathbb{R}$) and a metal ($n_2 = n_2' + i n_2''$). In this case, $w_2 = w_2' + i w_2''$ is complex, $w_2', w_2'' \neq 0$. For propagating waves in the first medium, equations (2.148) and (2.149) which describe the reflectance are still valid. For evanescent waves in the same medium, i.e. $w_1 = i w_1''$, we find that $\mathfrak{Im}(r_{p/s})$ is always nonzero – *all* evanescent waves can couple to the metal. This is in contrast to a dielectric halfspace where energy can only be transported if a propagating wave exists. Here, energy can be dissipated as heat instead, as we described at the end of the previous section. An example can be found in figure 2.15, which shows the fluxes $S_{1,z}$ and $S_{2,z}$ at an air/silver interface. Again, propagating waves in medium 1 are those up to the black vertical line at $q = k_1$. Directly at $w_1 = 0$, all energy is transported parallel to the interface and no energy enters the metal, the fluxes are zero. All evanesent waves with $q > k_1$ can couple to the metal, unlike at the air/glass interface. One prominant feature appears around $q \approx 1.07 k_1$, where the flux increases dramatically (see inset of figure 2.15). This is due to the excitation of *surface plasmon polaritons* (SPPs). An introduction to this phenomenon can be found for example in [55, 56], we will restrict ourselves to a short description here:

At the interface between a metal and a dielectric, the "sea" of conduction electrons – which is sometimes interpreted as a plasma – can be forced to oscillate collectively, leading to an oscillation of the polarization. This resulted in the name surface plasmon polaritons. Mathematically, SPPs are described by electromagnetic waves which decay exponentially in both the positive and the negative $z$-direction. By considering the boundary conditions derived at the beginning of section 2.3, one finds that SPPs cannot exist for $s$-waves when both media are non-magnetic [56]. Therefore, no strong feature appears in the yellow/blue line representing the energy flux of $s$-waves in figure 2.15. For $p$-waves, a dispersion relation can be derived [56]:

$$q_{\text{SPP}}^2 = k_v^2 \frac{n_1^2 \cdot n_2^2}{n_1^2 + n_2^2}. \tag{2.156}$$

At the air/silver interface at $\lambda = 612\,\text{nm}$, this corresponds to $q_{\text{SPP}} \approx (1.61 + 0.004i)\,k_v \approx (1.07 + 0.002i)\,k_1$. Thus, evanescent waves with $q$ close to this value can couple to SPPs, leading to an enhanced energy flux. The fact that $q_{\text{SPP}}$ has a (small) imaginary component suggests that SPPs have a finite propagation length along the interface – again, the kinetic energy of the electrons ultimately gets converted to heat.

### 2.4.4 Energy flux of a dipole emitter in free space

Next, we want to calculate the energy flux for a dipole emitter in a homogeneous medium. Starting from the closed-form solution (2.75),

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{n^2} \left[ k^2 + \nabla(\nabla \cdot) \right] \boldsymbol{p} \frac{\exp(ikr)}{r},$$

since the curl of the gradient of a function is zero, we can derive

$$\nabla \times \boldsymbol{E} = k_v^2 \, \nabla \times \left[ \boldsymbol{p} \frac{\exp(ikr)}{r} \right] = k_v^2 \left( ik - \frac{1}{r} \right) \frac{\mathrm{e}^{ikr}}{r} \hat{\boldsymbol{r}} \times \boldsymbol{p}. \tag{2.157}$$

Thus, with the explicit form of $\boldsymbol{E}$ for a dipole in the origin (2.79),

$$\boldsymbol{E} = nk_v^3 \left[ \underbrace{\left( 1 + \frac{i}{kr} - \frac{1}{(kr)^2} \right)}_{=:a} \boldsymbol{p} + \underbrace{\left( -1 - \frac{3i}{kr} + \frac{3}{(kr)^2} \right)}_{=:b} \hat{\boldsymbol{r}}(\hat{\boldsymbol{r}} \cdot \boldsymbol{p}) \right] \frac{\exp(ikr)}{kr},$$

we arrive at:

$$\boldsymbol{E} \times \left( \overline{\nabla \times \boldsymbol{E}} \right) = \frac{k_v^4}{r^2} \left( -i\bar{k} - \frac{1}{r} \right) \left\{ a \left[ \hat{\boldsymbol{r}} |\boldsymbol{p}|^2 - \boldsymbol{p}(\hat{\boldsymbol{r}} \cdot \boldsymbol{p}) \right] + b(\hat{\boldsymbol{r}} \cdot \boldsymbol{p}) \left[ \hat{\boldsymbol{r}}(\hat{\boldsymbol{r}} \cdot \boldsymbol{p}) - \boldsymbol{p} \right] \right\} \tag{2.158}$$

Assuming that $n$ is real, we can expand all the terms of $a$ and $b$ and finally arrive at a much shorter form for the imaginary part of this expression:

$$\begin{aligned} \mathfrak{Im}\{ \boldsymbol{E} \times \left( \overline{\nabla \times \boldsymbol{E}} \right) \} &= \frac{k_v^4}{r^2} \left\{ \hat{\boldsymbol{r}} p^2 \left( -k + \frac{k}{(kr)^2} - \frac{1}{kr^2} \right) + \hat{\boldsymbol{r}}(\hat{\boldsymbol{r}} \cdot \boldsymbol{p})^2 \left( k - \frac{3k}{(kr)^2} + \frac{3}{kr^2} \right) \right. \\ &\quad \left. - \boldsymbol{p}(\hat{\boldsymbol{r}} \cdot \boldsymbol{p}) \left( -k + \frac{1}{kr^2} - \frac{1}{kr^2} + k - \frac{3}{kr^2} + \frac{3}{kr^2} \right) \right\} \\ &= \frac{nk_v^5}{r^2} \left[ (\hat{\boldsymbol{r}} \cdot \boldsymbol{p})^2 - p^2 \right] \, \hat{\boldsymbol{r}}. \end{aligned} \tag{2.159}$$

Thus we find the time-averaged Poynting vector of a dipole emitter in a homogeneous medium:

$$\boldsymbol{S}_D(\boldsymbol{r}) = \frac{cnk_v^4}{8\pi r^2} [p^2 - (\hat{\boldsymbol{r}} \cdot \boldsymbol{p})^2] \hat{\boldsymbol{r}} \tag{2.160}$$

From this expression, we can conclude that:

1. The energy flows radially away from the dipole emitter.

2. The total emission per time scales linearly with the refractive index of the medium, and is inversely proportional to the fourth power of the wavelength.

3. For constant observation direction $\hat{\boldsymbol{r}}$ but increasing distance $r$, the time-averaged energy flux decreases quadratically. This ensures conservation of energy: If $\boldsymbol{S}_D(\boldsymbol{r})$ is integrated over concentric spheres around the dipole, the flux through these spheres has to be constant, because the only energy source is the dipole emitter and because there are no energy sinks in a homogeneous dielectric material.

It is often important to know how much energy is emitted into certain directions in the far field (i.e. far away from the emitter). This is described by the *angular distribution of radiation* (ADR) $\mathrm{d}^2 S / \mathrm{d}\Omega^2$, which is the energy emitted into the solid angle element

$d\Omega^2 = \sin\theta d\theta d\phi$ along the direction given by the two angles $(\theta, \phi)$. Formally, one can derive the ADR by determining the surface integral of $\boldsymbol{S}_D(\boldsymbol{r})$ on an infinitesimally small part of a sphere of radius $r$, and then dividing by the solid angle element of this small part of the sphere:

$$
\begin{aligned}
\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) &= \lim_{\substack{\Delta\theta \to 0 \\ \Delta\phi \to 0}} \frac{1}{\sin\theta \Delta\theta \Delta\phi} \int_{\theta-\Delta\theta}^{\theta+\Delta\theta} \mathrm{d}\theta' \int_{\phi-\Delta\phi}^{\phi+\Delta\phi} \mathrm{d}\phi' \; r^2 \sin\theta' \; \hat{\boldsymbol{r}} \cdot \boldsymbol{S}_D(r, \theta', \phi') \\
&= \lim_{\substack{\Delta\theta \to 0 \\ \Delta\phi \to 0}} \frac{1}{\sin\theta \Delta\theta \Delta\phi} \int_{\theta-\Delta\theta}^{\theta+\Delta\theta} \mathrm{d}\theta' \int_{\phi-\Delta\phi}^{\phi+\Delta\phi} \mathrm{d}\phi' \; \frac{cnk_v^4}{8\pi} p^2 (\sin\theta')^3 \\
&= \frac{cnk_v^4}{8\pi} p^2 \sin^2(\theta) \;,
\end{aligned}
\tag{2.161}
$$

where we assumed that the dipole moment $\boldsymbol{p}$ is oriented parallel to the $z$-axis such that $\hat{\boldsymbol{r}} \cdot \boldsymbol{p} = p \cos\theta$. Thus, most of the energy is emitted in the direction perpendicular to the dipole moment, while no energy is emitted directly along the dipole axis. The *total power of emission* is obtained by integrating the ADR over all angular directions,

$$
\boxed{S_{\text{tot}} = \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \; \sin\theta \; \frac{cnk_v^4}{8\pi} p^2 \sin^2(\theta) = \frac{1}{3} cnk_v^4 p^2 \;.}
\tag{2.162}
$$

### 2.4.5 Energy flux of a dipole emitter near a planar interface

The findings from the previous sections will now be combined to answer a question that is essential for this work: How exactly does the energy emission of a dipole change when it comes close to a planar interface? We have seen that the field of a dipole emitter can be decomposed into plane waves. Together with the knowledge of the behaviour of plane waves at planar interfaces, this allowed us to calculate the electric field of a dipole close to such an interface in chapter 2.3.2. Rather than deriving the Poynting vector from this complicated field and integrating it over a small solid angle element to find the ADR, we choose a different approach. Recall that we have found in section 2.4.1 that the energy flux of a single plane wave is along only one direction, namely $\hat{\boldsymbol{e}}_k$, and that the magnitude of the energy flux is proportional to the square of the absolute amplitude of the plane wave. Therefore, it seems plausible that the PW-decomposition of a field can be used to find the angular distribution of radiation of said field: the ratio $[\mathrm{d}^2 S/\mathrm{d}\Omega^2(\theta_1, \phi_1)] \; / \; [\mathrm{d}^2 S/\mathrm{d}\Omega^2(\theta_2, \phi_2)]$ should equal the ratio $|A(\theta_1, \phi_1)|^2 \; / \; |A(\theta_2, \phi_2)|^2$, where $A(\theta_i, \phi_i)$ is the amplitude of the plane wave $\exp[i\boldsymbol{k}(\theta_i, \phi_i)z]$ in the PW decomposition. Then, the actual value of $\mathrm{d}^2 S/\mathrm{d}\Omega^2(\theta, \phi)$ should be proportional to $|A(\theta, \phi)|^2$. To test if this is the case, we return to the Weyl representation of a dipole emitter in a homogeneous medium:

$$
\begin{aligned}
\boldsymbol{E}(\boldsymbol{r}) &= \frac{ink_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \; \sin\theta \, [\hat{\boldsymbol{e}}_p(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p}) + \hat{\boldsymbol{e}}_s(\hat{\boldsymbol{e}}_s \cdot \boldsymbol{p})] \\
&\qquad\qquad \cdot \exp\left(iq\rho_d \cos(\phi - \phi_d) + ik\cos\theta[z - z_0]\right)
\end{aligned}
$$

$$
\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > z_0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < z_0 \end{cases} \;.
\tag{2.163}
$$

The ADR is then given by[14]:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) = \sigma \frac{n^2 k_v^6}{4\pi^2} \left[ |\hat{e}_p(\theta, \phi) \cdot \boldsymbol{p}|^2 + |\hat{e}_s(\theta, \phi) \cdot \boldsymbol{p}|^2 \right] \tag{2.164}$$

with the yet unknown proportionality constant $\sigma$. Note that the evanescent waves from the Weyl decomposition do not appear in the ADR, since the latter describes the energy flux in the far field, where all evanescent waves have decayed. Furthermore, the position $\boldsymbol{r}_0$ does not enter the expression – mathematically because in the Weyl expansion, it only appeared in the complex exponential function, whose absolute value is one; physically because the distance between the origin and $\boldsymbol{r}_0$ can always be ignored if the far field is evaluated far enough from the dipole.

The proportionaly constant $\sigma$ can be found by comparing with the already known ADR of a dipole emitter in free space. In equation (2.161), we found the angular distribution of radiation for a dipole oriented parallel to the $z$-axis. In this situation, $\hat{e}_p \cdot \boldsymbol{p} = -p\sin\theta$ and $\hat{e}_s \cdot \boldsymbol{p} = 0$, and we arrive at:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) = \sigma \frac{n^2 k_v^6}{4\pi^2} p^2 \sin^2\theta. \tag{2.165}$$

This is indeed identical to the result (2.161) if we define $\sigma := (c\pi) \, / \, (2nk_v^2)$, or:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) = \frac{cnk_v^4}{8\pi} \left[ (\hat{e}_p(\theta, \phi) \cdot \boldsymbol{p})^2 + (\hat{e}_s(\theta, \phi) \cdot \boldsymbol{p})^2 \right]. \tag{2.166}$$

Let us now turn to the case of an interface between two media with refractive indices $n_1$ at $z > 0$ and $n_2$ at $z < 0$. For a dipole located at $\boldsymbol{r}_0 = (x_0, y_0, z_0 > 0)$ inside the first medium, we derived in section 2.2.3 that the reflected and transmitted fields $\boldsymbol{E}_R(x, y, z > 0)$ and $\boldsymbol{E}_T(x, y, z < 0)$ are given by:

$$\boldsymbol{E}_T(x, y, z < 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \, \sin\theta_1 \left[ T_p \hat{e}_{p2}(\hat{e}_{p1} \cdot \boldsymbol{p}) + T_s \hat{e}_{s2}(\hat{e}_{s1} \cdot \boldsymbol{p}) \right]$$
$$\cdot \exp\left( iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0 + ik_2 \cos\theta_2 z \right)$$

$$\boldsymbol{E}_R(x, y, z > 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \, \sin\theta_1 \left[ R_p \hat{e}_{p1}^*(\hat{e}_{p1} \cdot \boldsymbol{p}) + R_s \hat{e}_{s1}(\hat{e}_{s1} \cdot \boldsymbol{p}) \right]$$
$$\cdot \exp\left( iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0 + ik_1 \cos\theta_1^* z \right)$$

with the path $\quad \mathcal{C} = \pi/2 + i\infty \to \pi/2 \to \pi \quad$ and with $\theta_1^* := \pi - \theta_1$. $\tag{2.167}$

Both equations have to be transformed slightly to make them applicable for our approach. Firstly, we remark that the angle $\theta$ which appears in the definition of the solid angle element $\mathrm{d}\Omega^2 = \sin\theta \mathrm{d}\theta \mathrm{d}\phi$ has to be $\theta_1$ in the first and $\theta_2$ in the second medium. These two angles are linked by Snell's law, $n_1 \sin\theta_1 = n_2 \sin\theta_2$. However, in the above

---

[14]Since $\hat{e}_p$ and $\hat{e}_s$ are perpendicular to each other, $|A_p\hat{e}_p + A_s\hat{e}_s|^2 = (A_p\hat{e}_p + A_s\hat{e}_s) \cdot \overline{(A_p\hat{e}_p + A_s\hat{e}_s)} = |A_p|^2 + (A_p\overline{A_s} + \overline{A_p}A_s)(\hat{e}_p \cdot \hat{e}_s) + |A_s|^2 = |A_p|^2 + |A_s|^2$.

equations, we always integrate over $\theta_1$. For the transmitted field, we thus transform the integral over $\theta_1$ according to[15]

$$\int_{\mathcal{C}} \mathrm{d}\theta_1 \sin \theta_1 = \int_{\mathcal{D}} \mathrm{d}\theta_2 \ \sin \theta_2 \left(\frac{n_2}{n_1}\right)^2 \frac{\cos \theta_2}{\cos \theta_1} \quad \text{and } \mathcal{D} = \arcsin\left(\frac{n_1}{n_2}\sin[\mathcal{C}]\right). \quad (2.168)$$

Physically, the reflected field is composed of upwards-traveling plane waves. We emphasize this in the integral by formally exchanging $\theta_1$ and $\theta_1^* = \pi - \theta_1$:

$$\boldsymbol{E}_R(x,y,z>0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \ \sin\theta_1 \left[R_p \hat{\boldsymbol{e}}_{p1}(\hat{\boldsymbol{e}}_{p1}^* \cdot \boldsymbol{p}) + R_s \hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p})\right]$$
$$\cdot \exp\left(iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1^* z_0 + ik_1 \cos\theta_1 z\right)$$
$$\text{with the path} \quad \mathcal{C} = 0 \to \pi/2 \to \pi/2 - i\infty \quad \text{and with } \theta_1^* := \pi - \theta_1. \quad (2.169)$$

In order to find the total field in the first medium, the original dipole field (2.163) has to be added at all $z > 0$. Now, the position $\boldsymbol{r}_0$ of the dipole that we ignored in the free space case becomes important for the ADR because $z_0$ determines the relative phase difference between the original and the reflected waves. For the upper halfspace, i.e. for the angles $\theta \in [0, \pi/2]$, the ADR is then given by:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}\left(0 \leq \theta \leq \frac{\pi}{2}, \phi\right) = \frac{cn_1 k_v^4}{8\pi}\left[\left|(\hat{\boldsymbol{e}}_p(\theta, \phi) \cdot \boldsymbol{p})\mathrm{e}^{-ik_1 \cos\theta z_0} + (\hat{\boldsymbol{e}}_p(\pi - \theta, \phi) \cdot \boldsymbol{p})R_p \mathrm{e}^{ik_1 \cos\theta z_0}\right|^2\right.$$
$$\left. + \left|\left(\mathrm{e}^{-ik_1 \cos\theta z_0} + R_s \mathrm{e}^{ik_1 \cos\theta z_0}\right)(\hat{\boldsymbol{e}}_s(\theta, \phi) \cdot \boldsymbol{p})\right|^2\right]. \quad (2.170)$$

When determining the ADR for the lower halfspace, we have to remember that the proportionality constant $\sigma_i = (c\pi) / (2n_i k_v^2)$ is defined in the medium $i = 1, 2$ where the energy flux is determined, while the prefactor $(in_1 k_v^3) / (2\pi)$ from the Weyl decomposition is defined in the dipole's medium and thus stays the same. With this, and the changed integral (2.168), we find:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}\left(\frac{\pi}{2} \leq \theta_2 \leq \pi, \phi\right) = \frac{c\pi}{2n_2 k_v^2}\left(\frac{n_1 k_v^3}{2\pi}\right)^2 \left|\frac{n_2^2 \cos\theta_2}{n_1^2 \cos\theta_1}\right|^2$$
$$\cdot \left[|(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p})T_p \mathrm{e}^{-ik_1 \cos\theta_1 z_0}|^2 + |(\hat{\boldsymbol{e}}_s \cdot \boldsymbol{p})T_s \mathrm{e}^{-ik_1 \cos\theta_1 z_0}|^2\right]$$
$$= \frac{cn_2 k_v^4}{8\pi}\left|\frac{n_2 \cos\theta_2}{n_1 \cos\theta_1}\right|^2 \left[|T_p(\hat{\boldsymbol{e}}_p \cdot \boldsymbol{p})|^2 + |T_s(\hat{\boldsymbol{e}}_s \cdot \boldsymbol{p})|^2\right] \mathrm{e}^{2\mathfrak{Im}(k_1 \cos\theta_1)z_0}.$$
$$(2.171)$$

In the last step, we took into account that for $n_2 > n_1$ there exist some waves with $\theta_2$ close to $\pi/2$ that can propagate in medium two but that are evanescent in medium one (note that $\mathfrak{Im}(k_1 \cos\theta_1) \leq 0$ for $\pi/2 \leq \theta_2 \leq \pi$).

After these theoretical considerations, we will now discuss some examples. The top right panel of figure 2.16 shows a three-dimensional plot of the ADR of a dipole emitter immersed in water and oriented parallel to the $z$-axis, which is the case already derived in section 2.4.4. The distribution can be described as having a "doughnut shape", a

---

[15]Using $\theta_2 = \arcsin(\sin\theta_1 \cdot n_1/n_2)$, and then finding $f$ such that $\sin\theta_1 \mathrm{d}\theta_1 = f(\theta_2)\mathrm{d}\theta_2$.
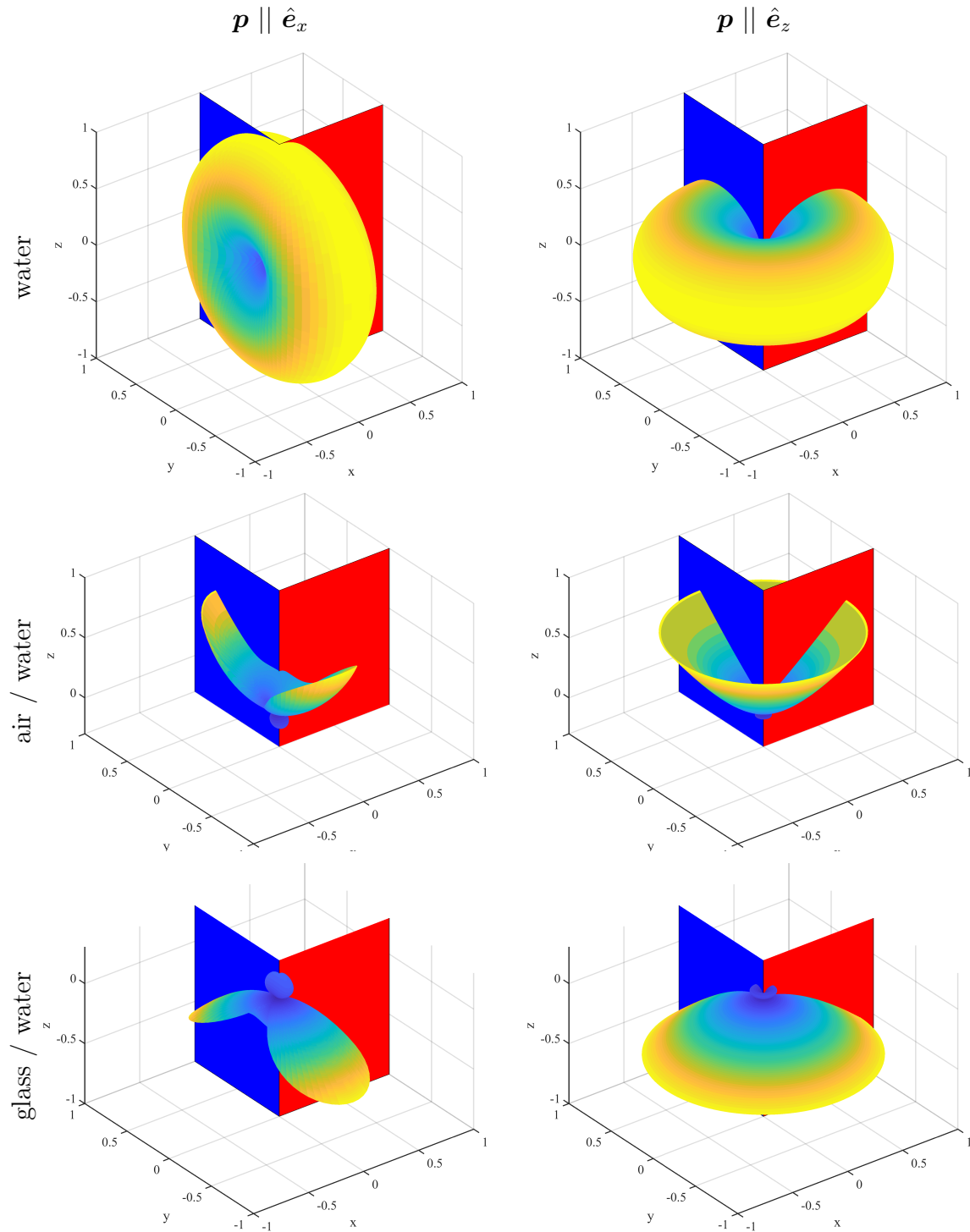
**Figure 2.16:** Angular distribution of radiation of a dipole emitter situated in water (top row), in water directly at the interface between water and air (central row) or in water at the interface between water and glass (bottom row). The dipole is oriented parallel to the $x$- or $z$-axis in the left and right column, respectively. The red (blue) plane shows a cut through $\phi = 0$ ($\phi = \pi/2$), this data is used for the curves on the right (left) side of the polar plots in figure 2.17. All distributions are normalized such that their maximum is one.
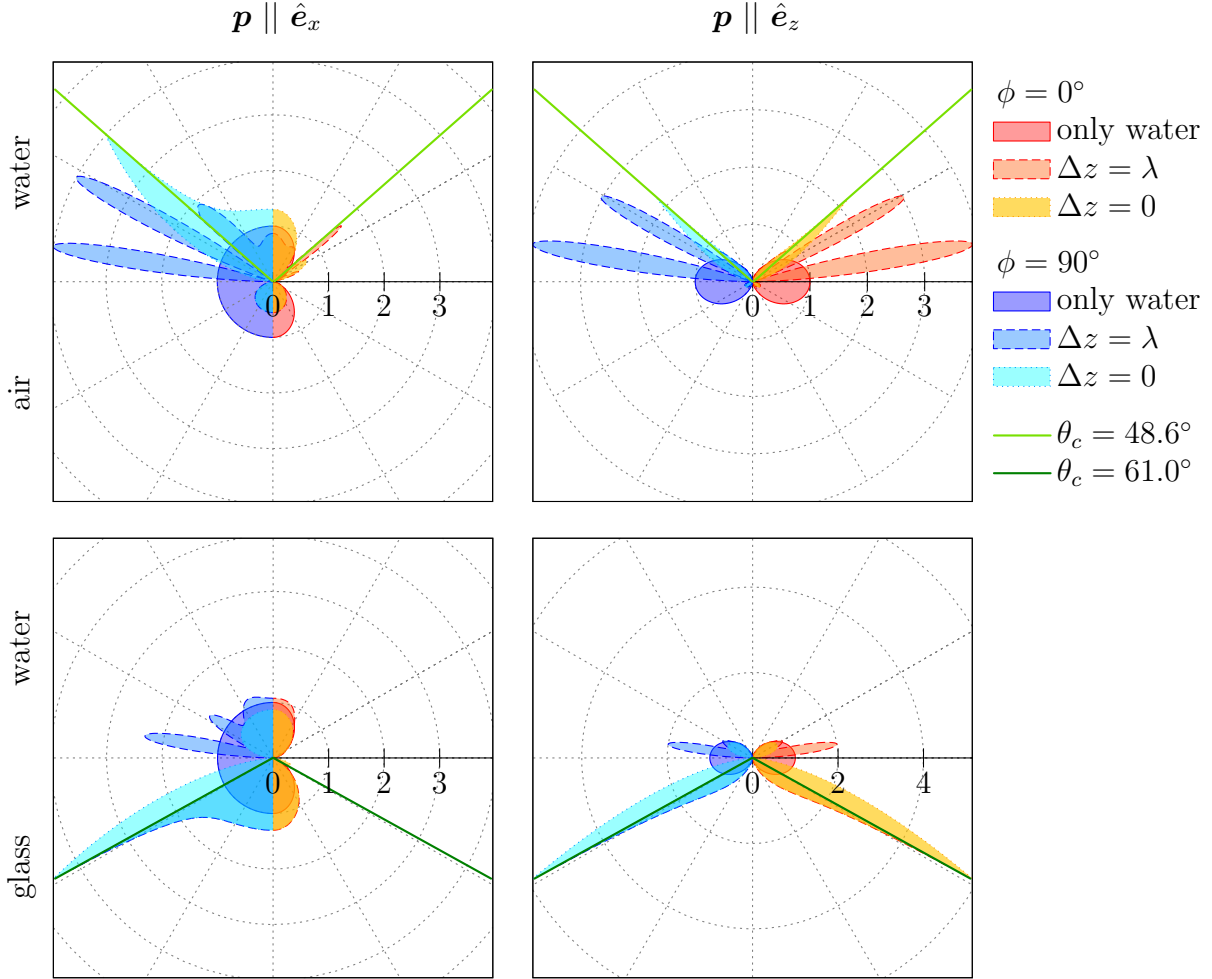
**Figure 2.17:** Normalized angular distribution of radiation of a dipole emitter situated in water at height $z$ over a water/air (top row) or water/glass (bottom row) interface. Shown are the patterns for a dipole parallel to the $x$-axis (left column) and parallel to the $z$-axis (right column). The radius of the polar plot at angle $\theta \in [0, \pi]$ corresponds to the flux in the direction $(\theta, \phi = 0)$ on the right side of each plot (red plane in figure 2.16), while the left side displays $\mathrm{d}^2 S/\mathrm{d}\Omega^2(\theta, \phi = \pi/2)$ (blue plane in figure 2.16). As a reference, the pattern for the same dipole only in water (top row in figure 2.16) is also included. Furthermore, the critical angle for total internal reflection $\theta_c$ is shown by the green lines, the stated values are measured between the line and the respective surface normal. Note that here, all curves are normalized to the maximum of $|\mathrm{d}^2 S/\mathrm{d}\Omega^2|$ of the water only case, while in figure 2.16 each single ADR is normalized such that its maximum is one. This allows to compare the total amount of emitted energy in different situations in this figure, while figure 2.16 illustrates the overall shape of the ADR.

vertical cut through the center reveals the typical "dumbbell shape" of $\sin^2(\theta)$. This illustrates that most of the energy is transported in the direction perpendicular to the dipole axis, while no energy at all is transported directly along the dipole axis. Due to the symmetry of the system, the pattern of a dipole oriented parallel to the $x$-axis is exactly the same, just rotated by 90° (top left panel of the figure). However, the symmetry gets broken when an interface is present. The central and bottom rows of figure 2.16 show the ADR for the same dipoles, but now situated in water directly at the interface between water ($n = 1.33$) and air ($n = 1$) or between water and glass ($n = 1.52$), respectively. Two things become apparent: Firstly, most of the energy is emitted into the halfspace with larger refractive index. Secondly, for given $\phi$ the $\theta$-distribution becomes much sharper, concentrating a large part of the energy transport in a small angular region. In fact, the polar plots in figure 2.17 show that with decreasing distance between dipole and interface, more energy is emitted into angles that are larger than the critical angle for total internal reflection. This means that some waves which are evanescent in the water halfspace become propagating waves when they enter the glass halfspace. However, if the dipole is too far from the interface, the evanescent waves have already decayed before reaching it, which is why this effect only occurs at small dipole-interface distance $z$.

These results have some practical implications based on the fact that it is always desirable to collect as many photons as possible in an experiment. For example, if we image single molecules that have been deposited on a glass cover slip, the situation is similar to the top row of figure 2.17 (just with glass instead of water). Then, most of the radiation is emitted into the glass halfspace, and it is advisable to image the molecules with an oil immersion objective through the glass rather than through an air objective from above. An oil-immersion objective with numerical aperture 1.49 and refractive index 1.52 of the immersion oil has a collection angle of 79°, which means that a large part of the lobes around 50° gets collected by the objective.

We follow this example with a short remark on metals. Metals have a complex refractive index, where the imaginary part describes absorption of energy by the metal. Thus, the electric field generated by an oscillating dipole decays more quickly in a metal than the $1/r$-law found for dielectric materials. Therefore it makes no sense to talk about "energy transport to the far field" if one halfspace consists of a metal. However, if only a several nanometers thin metal layer is present between two dielectric media, the ADR can be calculated exactly as described above – one only has to replace the Fresnel reflection and transmission coefficients with their effective values for the stack of layers. Figure 6.1 in the appendix shows the example of a 5 nm thick silver film between a water and a glass halfspace. It looks very similar to the case of a simple glass/water interface in the bottom row of figure 2.17.

Finally, we come to a quantity that will play a major role in the following chapters: the total power of emission $S_{\text{tot}}$ of the dipole. In the previous section, we calculated this quantity as the integral of the ADR over the full solid angle range, see eq. (2.162). This approach is valid if no absorbing materials are present, so it can also be used for a dipole emitter near a water/glass or water/air interface, as shown in figure 2.18a. However, if a metal structure is close to an oscillating dipole, the energy actually emitted by the latter will differ significantly from the energy transported to the far field. Then, the Poynting vector has to be integrated over a closed surface that completely encloses
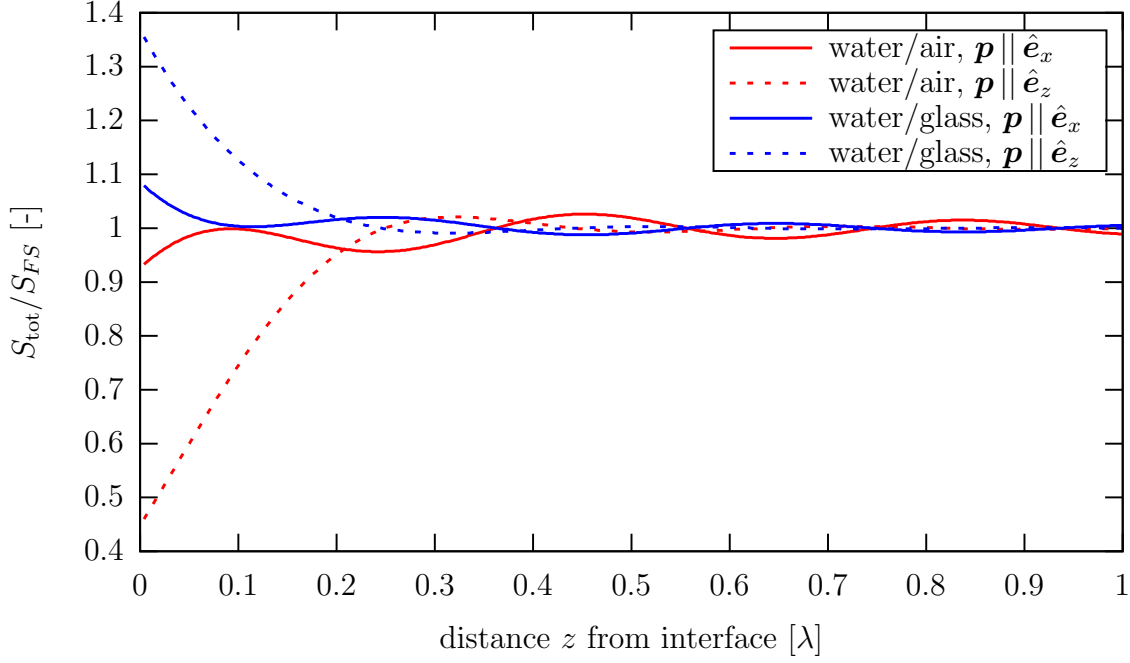
**(a)** Planar interface between water and air (red curves) or water and glass (blue curves).
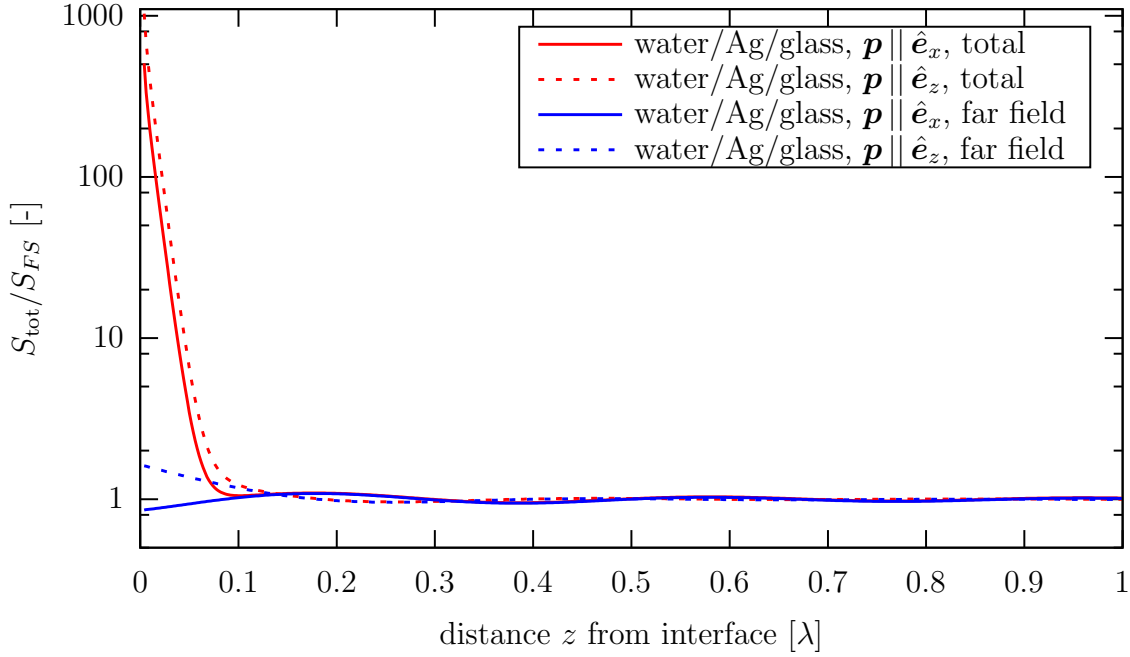


**(b)** 5 nm thick silver film on glass. Red curves are the "true" values of $S_{\text{tot}}(z)$, blue curves show the integral $\iint \mathrm{d}\theta \mathrm{d}\phi \, \sin\theta[\mathrm{d}^2 S(\theta,\phi,z)/\mathrm{d}\Omega^2]$, i.e. the energy transported to the far field.

**Figure 2.18:** Total power of emission $S_{\text{tot}}$ of a dipole emitter located in water at varying distance $z$ from a planar interface. Shown are both the case of a dipole parallel to the $x$- and $z$-axis (solid and dashed lines, respectively). All values were normalized to the total power of emission of a dipole in water, $S_{FS} = \frac{1}{3}cnk_v^4 p^2$ with $n = 1.33$.

the emitter but does not contain any absorbing materials. An efficient way to do this will be presented later, in section 3.4. For now, we restrict ourselves to a discussion of the results, shown in figure 2.18. For an emitter that is located far away from the interface, the total power of emission is the same as when there is no interface present at all. This confirms our observation from the previous paragraphs that evanescent waves play an important role in the interaction with an interface. If the emitter is close to the interface, $S_{\text{tot}}$ decreases if the neighbouring material has a lower refractive index than the host material of the dipole, and increases if the neighbouring material has a higher refractive index. This increase is particularly pronounced if a metal is close by. In order to explain this effect, we have to consider what happens to the energy emitted by the dipole. We do this on the basis of the results of section 2.4.3 and especially figure 2.15, where we investigated the energy transport of plane waves at an interface. If the adjoining medium is transparent ($n \in \mathbb{R}$), energy can only be transported in the form of propagating waves. If $n_2 > n_1$, then *some* evanescent waves from the host material turn into propagating waves, leading to an increase in total energy transport. However, many evanescent waves from the dipole's medium are still evanescent in the neighbouring medium and thus do not contribute to the energy transport. On the other hand, if a metal is present, then *all* evanescent waves that reach the metal can dissipate energy because of the absorption properties of the material. This leads to a much higher total energy emission.

### 2.4.6 Energy flux of a dipole emitter near a spherical interface

To conclude this chapter, we will now turn our attention to the energy flux of an oscillating dipole near a spherical interface. As we saw when discussing the field of a dipole emitter near a spherical interface, the application of the boundary conditions is simplified significantly by decomposing the field in vector spherical harmonics. However, the direction of energy transport of a VSH is not as easily found as for a plane wave, where we know that energy is only transported along $\hat{\boldsymbol{e}}_k$. Therefore, we will express the field of the dipole emitter as a linear combination of VSH, use this representation to describe the interaction with the spherical interface, and then transform the result to a PW decomposition in order to find the angular distribution of radiation. The first step is to decompose the original dipole field in VSH according to equation (2.80):

$$\boldsymbol{E}_{\text{VSH}}(\boldsymbol{r}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^{D} \boldsymbol{M}_{\ell m}^{f}(k\boldsymbol{r}) + b_{\ell m}^{D} \boldsymbol{N}_{\ell m}^{f}(k\boldsymbol{r})$$

$$\text{with} \quad a_{\ell m}^{D} = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{M}_{\ell m}^{g}(k\boldsymbol{r}_0)} \quad \text{and} \quad b_{\ell m}^{D} = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{N}_{\ell m}^{g}(k\boldsymbol{r}_0)} \,, \qquad (2.172)$$

with $f = j$ and $g = h$ for $r < r_0$ and the reverse for $r > r_0$. As the second step, the field is scattered at the sphere as described in sections 2.3.3 and 2.3.4, finding the expansion coefficients $a_{\ell m}^{sc}, b_{\ell m}^{sc}$ of the scattered field

$$\boldsymbol{E}_{sc}(\boldsymbol{r}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^{sc} \boldsymbol{M}_{\ell m}^{h} + b_{\ell m}^{sc} \boldsymbol{N}_{\ell m}^{h}. \qquad (2.173)$$

The third step is to calculate the PW decomposition of the resulting far field. This means that $r > r_0$ and thus $f = h$ and $g = j$ in equation (2.172). Thus, only VSH with spherical Hankel functions as radial terms are needed, and we can use equations (2.60) and (2.61) to find the plane wave decomposition of the field. This directly leads to the fourth step, determining the angular distribution of radiation. We use the result from the previous section that for a general electric field given as

$$\boldsymbol{E}(\boldsymbol{r}) = \int_{\mathcal{C}} \int_0^{2\pi} \mathrm{d}\theta \ \mathrm{d}\phi \ \sin\theta \ \boldsymbol{A}(\theta, \phi) \mathrm{e}^{i\boldsymbol{k}(\theta,\phi)\cdot\boldsymbol{r}}, \tag{2.174}$$

where $\mathcal{C}$ contains propagating ($\theta \in [0, \pi]$) and evanescent ($\theta = \pi/2 + i\theta'', \theta'' \in \mathbb{R}$) waves, the angular distribution of radiation is given by:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) = \frac{c\pi}{2nk_v^2} \ |\boldsymbol{A}(\theta, \phi)|^2. \tag{2.175}$$

For a single vector spherical harmonic, the plane-wave decompositions from section 2.1.4 can be used directly. If a dipole field is made up of several different VSH, it is important to add the plane wave decompositions *before* using them in the calculation of the flux, since $|\sum \boldsymbol{A}_i|^2 \neq \sum |\boldsymbol{A}_i|^2$. Combining these four steps leads to the final result:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}\Omega^2}(\theta, \phi) = \frac{c\pi}{2nk_v^2} \left\{ \left| \frac{1}{2\pi} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \mathrm{e}^{im\phi} \left[ (a_{\ell m}^D + a_{\ell m}^{sc})\pi_{\ell m} + (b_{\ell m}^D + b_{\ell m}^{sc})\tau_{\ell m} \right] \right|^2 \right.$$

$$\left. + \left| \frac{1}{2\pi} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \mathrm{e}^{im\phi} \left[ (a_{\ell m}^D + a_{\ell m}^{sc})\tau_{\ell m} + (b_{\ell m}^D + b_{\ell m}^{sc})\pi_{\ell m} \right] \right|^2 \right\}. \tag{2.176}$$

We use this formula to compare the ADR of a dipole situated close to a glass or silver sphere of radius $R = 200\,\mathrm{nm}$ – the same situation for which we calculated the electric field at the end of section 2.3.4 and in figure 2.13. Figure 2.19 shows the three-dimensional distribution for $\boldsymbol{r}_0 = (0, 0, R)$ and $\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_x$ or $\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_z$. A $z$-dipole positioned on the $z$-axis in free space does not emit any energy along the z-axis, while an $x$-dipole at the same position has its emission maximum in the whole $y$-$z$-plane. Therefore, it seems plausible that a glass sphere at the origin has a larger influence on the ADR of an $x$-dipole at $(0, 0, R)$ than on that of a $z$-dipole at $(0, 0, R)$. Indeed, that is what figure 2.19 shows: The distribution of the $z$-dipole is only slightly distorted, while that of the $x$-dipole displays a strong lobe in the direction $\theta = \pi$, i.e. towards the center of the sphere. This is further corroborated by the two-dimensional cuts along the red and blue half-planes at $\phi = 0$ and $\phi = \pi/2$, which are plotted in figure 2.20. There, the ADR is also shown for a dipole that is situated one wavelength away from the sphere. In that case, the distribution for a $z$-dipole is already almost identical to the free space ADR, while the pattern for an $x$-dipole still shows a preferential emission in the direction of the sphere. On the other hand, if the sphere is made out of silver, the distributions of radiation of both dipole orientations are strongly changed compared to the free space case. They have several emission maxima, one of them is even close to
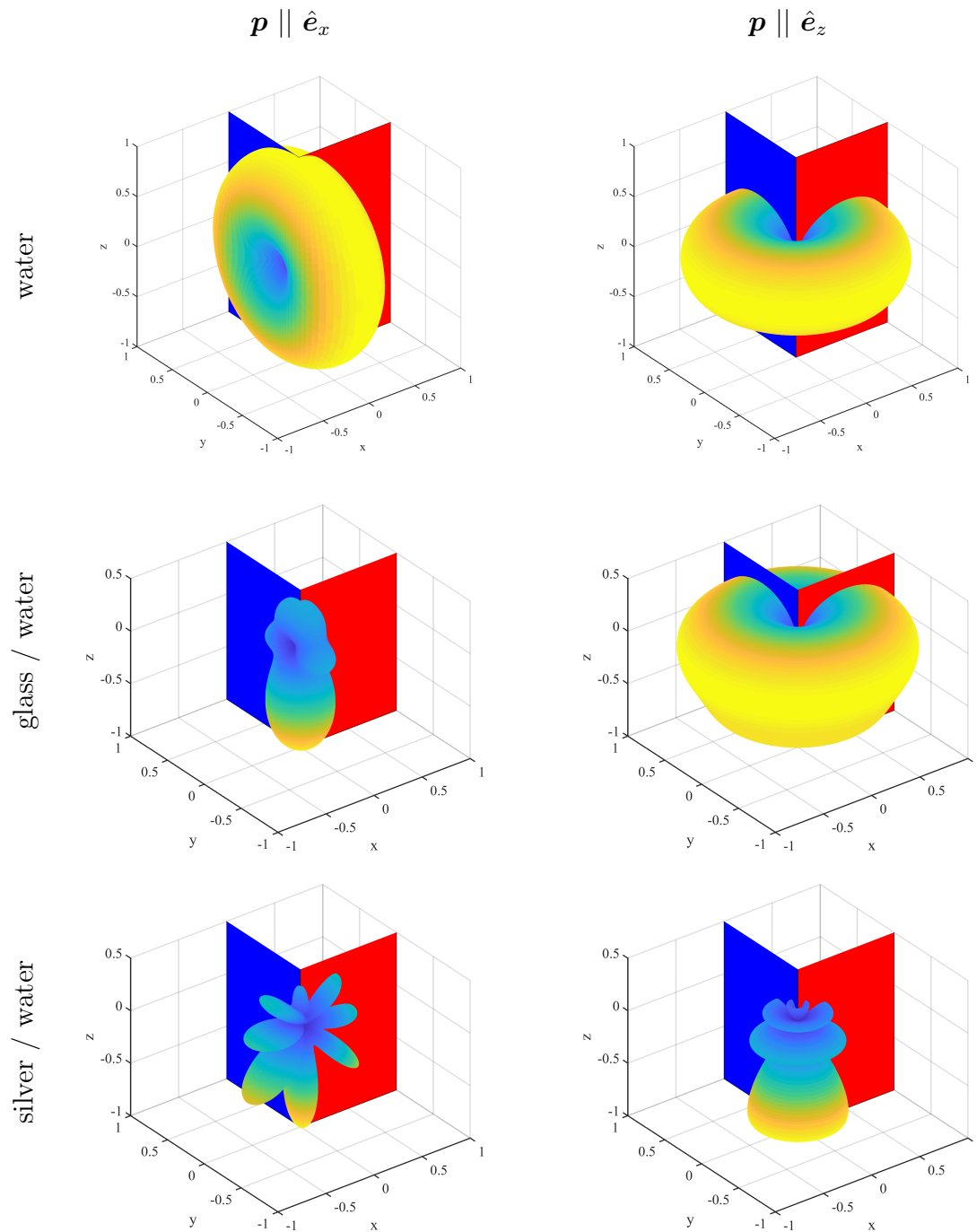
**Figure 2.19:** Angular distribution of radiation of a dipole emitter situated in water (top row) or on the $z$-axis directly at the surface of a glass or silver sphere embedded in water (central and bottom row). The dipole is oriented parallel to the $x$- or $z$-axis in the left and right column, respectively. The red (blue) plane shows a cut through $\phi = 0$ ($\phi = \pi/2$), this data is used for the curves on the right (left) side of the polar plots in figure 2.20. The wavelength is $\lambda = 500\,\mathrm{nm}$, the radius of the spheres is $R = 200\,\mathrm{nm}$, and all distributions are normalized such that their maximum is one.
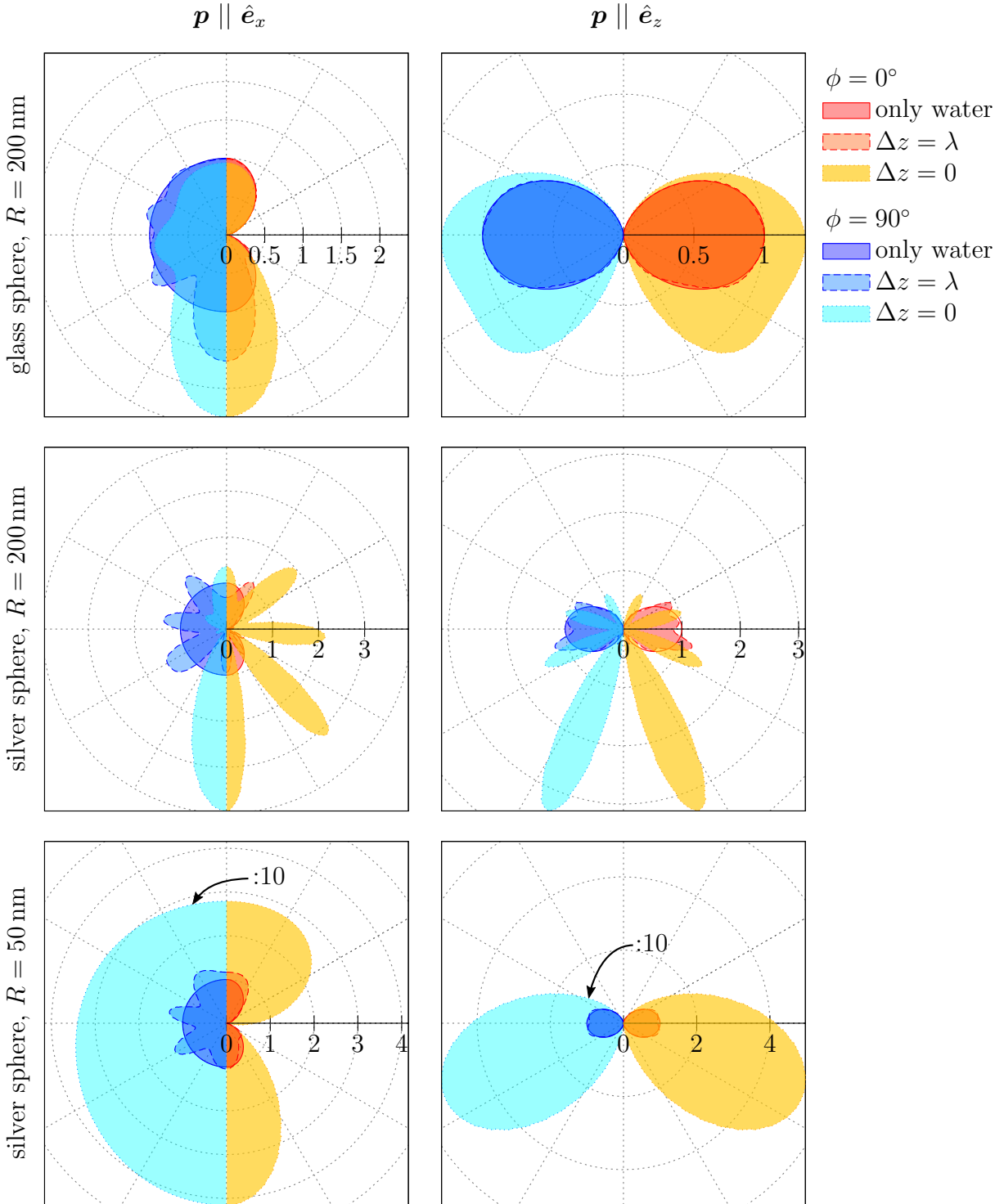
**Figure 2.20:** ADR of a dipole emitter oriented parallel to the $x$- or $z$-axis (left or right column) and placed in water at different height on the $z$-axis close to a sphere made of glass or silver (top or central/bottom row). The wavelength is $\lambda = 500\,\mathrm{nm}$, the radius of the sphere is $R$, the distance $\Delta z$ between the surface of the sphere and the dipole is zero (dotted line) or $\lambda$ (dashed line). As a reference, a dipole in water without sphere is shown with a solid line. All values are normalized to the maximum value of the water only case. For $R = 50\,\mathrm{nm}$, the values for zero distance are so large that they have been divided by 10 to fit inside the figure.

the formerly forbidden direction $\theta = \pi/2$ for $\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_x$. Now, also the radius of the sphere plays an important role: If $R$ is changed to $50\,\mathrm{nm}$, the emission pattern resembles the free space one with slight bias towards the negative halfspace, but is strongly enhanced (note that the radius for $\Delta z = 0$ in the bottom row of figure 2.20 has to be multiplied by 10). In contrast, if a glass sphere with radius $50\,\mathrm{nm}$ is brought close to the dipole, the pattern remains almost unchanged compared to free space (see figure 6.2 in the appendix).

The dramatic change in the amplitude of the angular distribution of radiation for different properties of the sphere raises the question how the total power of emission is affected by these factors. Since we want to calculate $S_{\mathrm{tot}}$ for both transparent and absorbing spheres, we have to be able to integrate the Poynting vector over spherical surfaces[16]. Therefore, it is advantageous to do this calculation in vector spherical harmonics instead of plane waves. As we saw in section 2.1.3, VSH have the very useful property that $\nabla \times \boldsymbol{M}_{\ell m}^{f} = k\boldsymbol{N}_{\ell m}^{f}$ and $\nabla \times \boldsymbol{N}_{\ell m}^{f} = k\boldsymbol{M}_{\ell m}^{f}$. Thus, if an electric field is given by the decomposition

$$\boldsymbol{E} = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}\boldsymbol{M}_{\ell m}^{f} + b_{\ell m}\boldsymbol{N}_{\ell m}^{f} \ , \tag{2.177}$$

the time-averaged Poynting vector is

$$\boldsymbol{S} = -\frac{c}{8\pi k_v}\mathfrak{Im}\left\{\boldsymbol{E} \times \left(\overline{\nabla \times \boldsymbol{E}}\right)\right\}$$

$$= -\frac{c}{8\pi k_v}\mathfrak{Im}\left\{\left(\sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}\boldsymbol{M}_{\ell m}^{f} + b_{\ell m}\boldsymbol{N}_{\ell m}^{f}\right) \times \overline{\left(k \sum_{o=1}^{\infty} \sum_{p=-o}^{o} a_{op}\boldsymbol{N}_{op}^{f} + b_{op}\boldsymbol{M}_{op}^{f}\right)}\right\} \ . \tag{2.178}$$

$S_{\mathrm{tot}}$ can now be found by integrating the time-averaged Poynting vector over the surface of a volume that contains the dipole but no other energy sources or sinks. In the simplest case of a dipole in free space, this can simply be a sphere with radius $R > r_0$ centered at the origin (surface area element $R^2 \sin\theta \mathrm{d}\theta \mathrm{d}\phi \hat{\boldsymbol{e}}_r$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$). Then we only need to find the $r$-component of $\boldsymbol{S}$. To do that, consider

$$\left(\boldsymbol{M}_{\ell m}^{f} \times \overline{\boldsymbol{M}_{op}^{f}}\right) \cdot \hat{\boldsymbol{e}}_r = -i\left(\pi_{\ell m}\overline{\tau_{op}} + \tau_{\ell m}\overline{\pi_{op}}\right)\mathrm{e}^{im\phi - ip\phi}f_{\ell}\overline{f_o} \ ,$$

$$\left(\boldsymbol{M}_{\ell m}^{f} \times \overline{\boldsymbol{N}_{op}^{f}}\right) \cdot \hat{\boldsymbol{e}}_r = \quad \left(\pi_{\ell m}\overline{\pi_{op}} + \tau_{\ell m}\overline{\tau_{op}}\right)\mathrm{e}^{im\phi - ip\phi}f_{\ell}\overline{\frac{1}{kr}\frac{\mathrm{d}(rf_o)}{\mathrm{d}r}} \ ,$$

$$\left(\boldsymbol{N}_{\ell m}^{f} \times \overline{\boldsymbol{M}_{op}^{f}}\right) \cdot \hat{\boldsymbol{e}}_r = -\left(\tau_{\ell m}\overline{\tau_{op}} + \pi_{\ell m}\overline{\pi_{op}}\right)\mathrm{e}^{im\phi - ip\phi}\frac{1}{kr}\frac{\mathrm{d}(rf_{\ell})}{\mathrm{d}r}\overline{f_o} \ ,$$

$$\left(\boldsymbol{N}_{\ell m}^{f} \times \overline{\boldsymbol{N}_{op}^{f}}\right) \cdot \hat{\boldsymbol{e}}_r = -i\left(\tau_{\ell m}\overline{\pi_{op}} + \pi_{\ell m}\overline{\tau_{op}}\right)\mathrm{e}^{im\phi - ip\phi}\frac{1}{|k|^2r^2}\frac{\mathrm{d}(rf_{\ell})}{\mathrm{d}r}\overline{\frac{\mathrm{d}(rf_o)}{\mathrm{d}r}} \ , \tag{2.179}$$

---

[16]Since transparent spheres do not contain any energy sinks, $S_{\mathrm{tot}}$ could in this case be determined alternatively by integrating the ADR over the full solid angle range, as we did for a dipole in free space in equation (2.162). For metal spheres, the total emission into the far field no longer equals the total emission, as already discussed for a dipole close to a planar metallic region.

where the complex conjugation of $\pi_{op}$ and $\tau_{op}$ could also be omitted since these functions are real, anyway. If any of these expressions is integrated over $\phi \in [0, 2\pi]$, the result will be zero unless $m = p$ due to the complex exponential function. When proving the orthogonality of the VSH in section 6.1.1 in the appendix, we have shown in equation (6.4) that

$$\int_0^\pi \mathrm{d}\theta \sin\theta \left( i\pi_{\ell m}(\theta)\tau_{om}(\theta) + i\pi_{om}\tau_{\ell m} \right) = 0, \tag{2.180}$$

and in equation (6.5) that

$$\int_0^\pi \mathrm{d}\theta \sin\theta \left( \pi_{\ell m}\pi_{om} + \tau_{\ell m}\tau_{om} \right) = \delta_{\ell o}\frac{\ell(\ell+1)}{2\pi}. \tag{2.181}$$

Therefore, of all the terms in equation (2.178), only those where $\boldsymbol{M}$ is multiplied with $\boldsymbol{N}$ and where $\ell = o$ and $m = p$ give a contribution to the total power of emission $S_{\text{tot}}$. This leads us to:

$$S_{\text{tot}} = -\frac{cR^2}{8\pi k_v} \sum_{\ell=1}^\infty \sum_{m=-\ell}^\ell \mathfrak{Im}\left\{ \ell(\ell+1)\overline{k} \left[ |a_{\ell m}|^2 f_\ell \overline{\frac{1}{kr}\frac{\mathrm{d}(rf_\ell)}{\mathrm{d}r}} - |b_{\ell m}|^2 \overline{f_\ell}\frac{1}{kr}\frac{\mathrm{d}(rf_\ell)}{\mathrm{d}r} \right] \right\}. \tag{2.182}$$

Several conclusions can be drawn from this expression. Firstly, since $j_\ell(kr) \in \mathbb{R}$, the total power of emission of a VSH with spherical Bessel function is zero[17]. Secondly, since ([46], §19)

$$h_\ell(kr) \xrightarrow{kr \gg 1} \frac{1}{kr}\mathrm{e}^{ikr - i(\ell+1)\pi/2}$$

$$\text{and} \quad \frac{1}{kr}\frac{\mathrm{d}[rh_\ell(kr)]}{\mathrm{d}r} \xrightarrow{kr \gg 1} \frac{i}{kr}\mathrm{e}^{ikr - i(\ell+1)\pi/2}, \tag{2.183}$$

we find for large values of $R$ and for real $k$:

$$S_{\text{tot}} = \frac{c}{8\pi n k_v^2} \sum_{\ell=1}^\infty \sum_{m=-\ell}^\ell \ell(\ell+1) \left[ |a_{\ell m}|^2 + |b_{\ell m}|^2 \right] \tag{2.184}$$

Therefore, a field that only consists of VSH with spherical Hankel functions has a positive net flux towards the far-field. As a consistency check, consider the VSH decomposition of a dipole situated at the origin, eq. (2.82). There, we found $b_{1,0} = ink_v^3 p\sqrt{4\pi/3}$ as the only non-zero coefficient for a $z$-dipole. Thus, the total power of emission is

$$S_{\text{tot}}(\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_z, \boldsymbol{r}_0 = 0) = \frac{c}{8\pi n k_v^2}\, 2\, \left(nk_v^3 p\right)^2 \frac{4\pi}{3} = \frac{1}{3}cnk_v^4 p^2, \tag{2.185}$$

which is exactly the same as found in equation (2.162) from the closed-form solution of the dipole field.

---

[17]Unless it is evaluated in a lossy medium where $k$ has an imaginary component and $f_\ell(kr) \notin \mathbb{R}$, but that is not the case in the situations studied by us.

Up to now, we assumed that it was possible to enclose the dipole in a (virtual) sphere centered at the origin without including any other energy sources or sinks. However, if a metal sphere is placed at the origin, this is no longer the case. Similar to the approach from the previous section, we then have to change the integration surfaces. In order to still be able to exploit the properties of the VSH with respect to integration over concentric spheres, we simply choose two spherical surfaces: one with a large radius $r_{\text{out}} > r_0$, and one that is between the dipole and the surface of the (physical) metal sphere, $r_0 > r_{\text{in}} > R$. $S_{\text{tot}}$ is then given by the sum of the two fluxes[18]. Since the field used to calculate the Poynting vector is now the sum of the original dipole field and the scattered field, where the latter always uses $h(kr)$ and the former uses $j(kr)$ for $r < r_0$ and $h(kr)$ for $r > r_0$, finding the flux through the inner sphere necessitates treating a field of the form

$$\boldsymbol{E} = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^{j} \boldsymbol{M}_{\ell m}^{j} + a_{\ell m}^{h} \boldsymbol{M}_{\ell m}^{h} + b_{\ell m}^{j} \boldsymbol{N}_{\ell m}^{j} + b_{\ell m}^{h} \boldsymbol{N}_{\ell m}^{h}. \tag{2.186}$$

Then, mixed radial terms appear in the result:

$$
\begin{aligned}
S_{\text{in}} = -\frac{cR^2}{8\pi k_v} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \mathfrak{Im} \Bigg\{ \ell(\ell+1)\overline{k} \Bigg[ &|a_{\ell m}^{j}|^2 j_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rj_\ell)}{\mathrm{d}r}} + a_{\ell m}^{j} \overline{a_{\ell m}^{h}} j_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r}} \\
+\ & a_{\ell m}^{h} \overline{a_{\ell m}^{j}} h_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rj_\ell)}{\mathrm{d}r}} + |a_{\ell m}^{h}|^2 h_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r}} - |b_{\ell m}^{j}|^2 \frac{1}{kr} \frac{\mathrm{d}(rj_\ell)}{\mathrm{d}r} \overline{j_\ell} \\
-\ & b_{\ell m}^{j} \overline{b_{\ell m}^{h}} \frac{1}{kr} \frac{\mathrm{d}(rj_\ell)}{\mathrm{d}r} \overline{h_\ell} - b_{\ell m}^{h} \overline{b_{\ell m}^{j}} \frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r} \overline{j_\ell} - |b_{\ell m}^{h}|^2 \frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r} \overline{h_\ell} \Bigg] \Bigg\}.
\end{aligned} \tag{2.187}
$$

Does it matter where exactly we place the virtual sphere, i.e. which value we choose for $r_{\text{in}}$? If we again assume that the medium containing the dipole is transparent (i.e. $n \in \mathbb{R}$), the terms that only contain versions of $j_\ell$ are purely real and thus do not contribute to the final result. Applied to the example of the dipole, this means that if there is no metal particle present but a dipole is located at $\boldsymbol{r}_0 \neq \boldsymbol{0}$ away from the origin, there is no net flux through a sphere that does not contain the dipole. For the terms that contain only versions of $h_\ell$, we just saw that the $r$-dependence has the form $r^{-2}$, leading to a radius-independent flux through a concentric sphere. But what about the mixed terms? We define two auxiliary functions,

$$
\begin{aligned}
f_\ell(\rho) &:= j_\ell \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho h_\ell)}{\mathrm{d}\rho}} + h_\ell \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho j_\ell)}{\mathrm{d}\rho}} \\
\text{and} \quad g_\ell(\rho) &:= j_\ell \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho h_\ell)}{\mathrm{d}\rho}} - h_\ell \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho j_\ell)}{\mathrm{d}\rho}}.
\end{aligned} \tag{2.188}
$$

---

[18]Here, one has to take into account that the vector $\hat{\boldsymbol{e}}_r$, which we used as the normal vector of the sphere, always points outwards, while we are interested in the *inwards* flux through the inner sphere. Therefore, the result for the flux through the inner sphere has to be multiplied by -1.

Then, for $n \in \mathbb{R}$, the total flux through the inner sphere is given by

$$S_{\mathrm{in}} = -\frac{cnR^2}{8\pi} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \ell(\ell+1) \left\{ (|a_{\ell m}^h|^2 + |b_{\ell m}^h|^2) \mathfrak{Im} \left[ h_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r}} \right] \right.$$
$$\left. + \mathfrak{Re} \left[ a_{\ell m}^j \overline{a_{\ell m}^h} + b_{\ell m}^j \overline{b_{\ell m}^h} \right] \cdot \mathfrak{Im}[f_\ell] + \mathfrak{Im} \left[ a_{\ell m}^j \overline{a_{\ell m}^h} + b_{\ell m}^j \overline{b_{\ell m}^h} \right] \cdot \mathfrak{Re}[g_\ell] \right\}, \quad (2.189)$$

where we exploited the fact that $\mathfrak{Re}(x\overline{y}) = \mathfrak{Re}(\overline{x}y)$ and $\mathfrak{Im}(x\overline{y}) = -\mathfrak{Im}(\overline{x}y)$. Since spherical Hankel functions can be expressed as a sum of (real) spherical Bessel and Neumann functions, $h_\ell = j_\ell + iy_\ell$, we find for the real part of $g_\ell$[19]:

$$\mathfrak{Re}[g_\ell(\rho)] = j_\ell \left( \frac{j_\ell}{\rho} + \frac{\mathrm{d}j_\ell}{\mathrm{d}\rho} \right) - j_\ell \left( \frac{j_\ell}{\rho} + \frac{\mathrm{d}j_\ell}{\mathrm{d}\rho} \right) = 0. \quad (2.190)$$

Furthermore, we find the identity

$$\mathfrak{Im} \left[ h_\ell \overline{\frac{1}{kr} \frac{\mathrm{d}(rh_\ell)}{\mathrm{d}r}} \right] = j_\ell \frac{1}{kr} \frac{\mathrm{d}(-ry_\ell)}{\mathrm{d}r} + y_\ell \frac{1}{kr} \frac{\mathrm{d}(rj_\ell)}{\mathrm{d}r} = \mathfrak{Im}[f_\ell(kr)]. \quad (2.191)$$

Therefore, we only need to study the behaviour of $\mathfrak{Im}[f_\ell(kr)]$ for different $r$. In the limit of large radius, the asymptotic expression from equation (2.58) tell us that

$$f_\ell(\rho) \xrightarrow{\rho \gg 1} \frac{1}{\rho^2} \left[ -i + (-1)^{\ell+1} \cdot \sin(2\rho) \right]. \quad (2.192)$$

Thus, for large $r$ the imaginary part of $f$ is simply $-(kr)^{-2}$. In fact, this is true for *all* $r$. Since a proof could not be found in the literature, a proof by induction is provided in section 6.1.2 in the appendix of this work. To sum up, this means that the integral over a mixed VSH-decomposition with coefficients $a_{\ell m}^j$, $a_{\ell m}^h$, $b_{\ell m}^j$ and $b_{\ell m}^h$ is simply given by:

$$S_{j,h} = \frac{c}{8\pi n k_v^2} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \ell(\ell+1) \left\{ |a_{\ell m}^h|^2 + |b_{\ell m}^h|^2 + \mathfrak{Re} \left[ a_{\ell m}^j \overline{a_{\ell m}^h} + b_{\ell m}^j \overline{b_{\ell m}^h} \right] \right\}. \quad (2.193)$$

These results can now be used to find the total flux for a dipole emitter close to a nanosphere. With the notation from equations (2.172,2.173) and the correct sign of the flux through the inner sphere, this is:

$$\boxed{\begin{aligned} S_{\mathrm{tot}} = \frac{c}{8\pi n k_v^2} \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \ell(\ell+1) & \left\{ \left[ |a_{\ell m}^{D,h} + a_{\ell m}^{\mathrm{sc}}|^2 + |b_{\ell m}^{D,h} + b_{\ell m}^{\mathrm{sc}}|^2 \right] \right. \\ & \left. - \left[ |a_{\ell m}^{\mathrm{sc}}|^2 + |b_{\ell m}^{\mathrm{sc}}|^2 + \mathfrak{Re} \left( a_{\ell m}^{D,j} \overline{a_{\ell m}^{\mathrm{sc}}} + b_{\ell m}^{D,j} \overline{b_{\ell m}^{\mathrm{sc}}} \right) \right] \right\}. \end{aligned}} \quad (2.194)$$

We now provide some examples to illustrate these results. Figure 2.21 shows the total power of emission of an $x$- as well as a $z$-dipole on the $z$-axis above a glass or silver sphere. It is immediately obvious that, for a large distance between dipole and sphere,

---

[19]Again assuming a transparent medium, i.e. $k \in \mathbb{R}$ and thus $\rho \in \mathbb{R}$.

$S_{\mathrm{tot}}$ approaches the value of the free space case $S_{FS} = cnk_v^4p^2/3$. This is reasonable since the absence of a sphere is the same as the limit of an infinite distance betwen sphere and emitter. Remarkably, close to a glass sphere, $S_{\mathrm{tot}}/S_{FS}$ becomes larger than one for a $z$-dipole but smaller than one for an $x$-dipole – unless the sphere is large, then an $x$-dipole has a higher total power of emission, too. This can be explained by studying the polar plots of $\mathrm{d}^2S/\mathrm{d}\Omega^2$ in figures 2.20 and 6.2: For a $z$-dipole, all values for $\Delta z = 0$ are larger than the free space values. Contrarily, the $x$-dipole curves for $\Delta z = 0$ are completely within the free space curves for $R = 50$. For $R = 200\,\mathrm{nm}$, the dipole still emits less energy than in free space towards the upper halfspace, however, this is overcompensated by a strong emission towards $\theta = \pi$. The behaviour is different close to a silver sphere: Directly at the surface, $S_{\mathrm{tot}}$ increases significantly, more so for a smaller sphere than a larger one. At a distance of roughly one wavelength from the interface, figure 2.20 shows that the shape of the ADR is still distorted, but that an increased emission in some directions is compensated by a decrease in others, leading to $S_{\mathrm{tot}} \approx S_{FS}$.
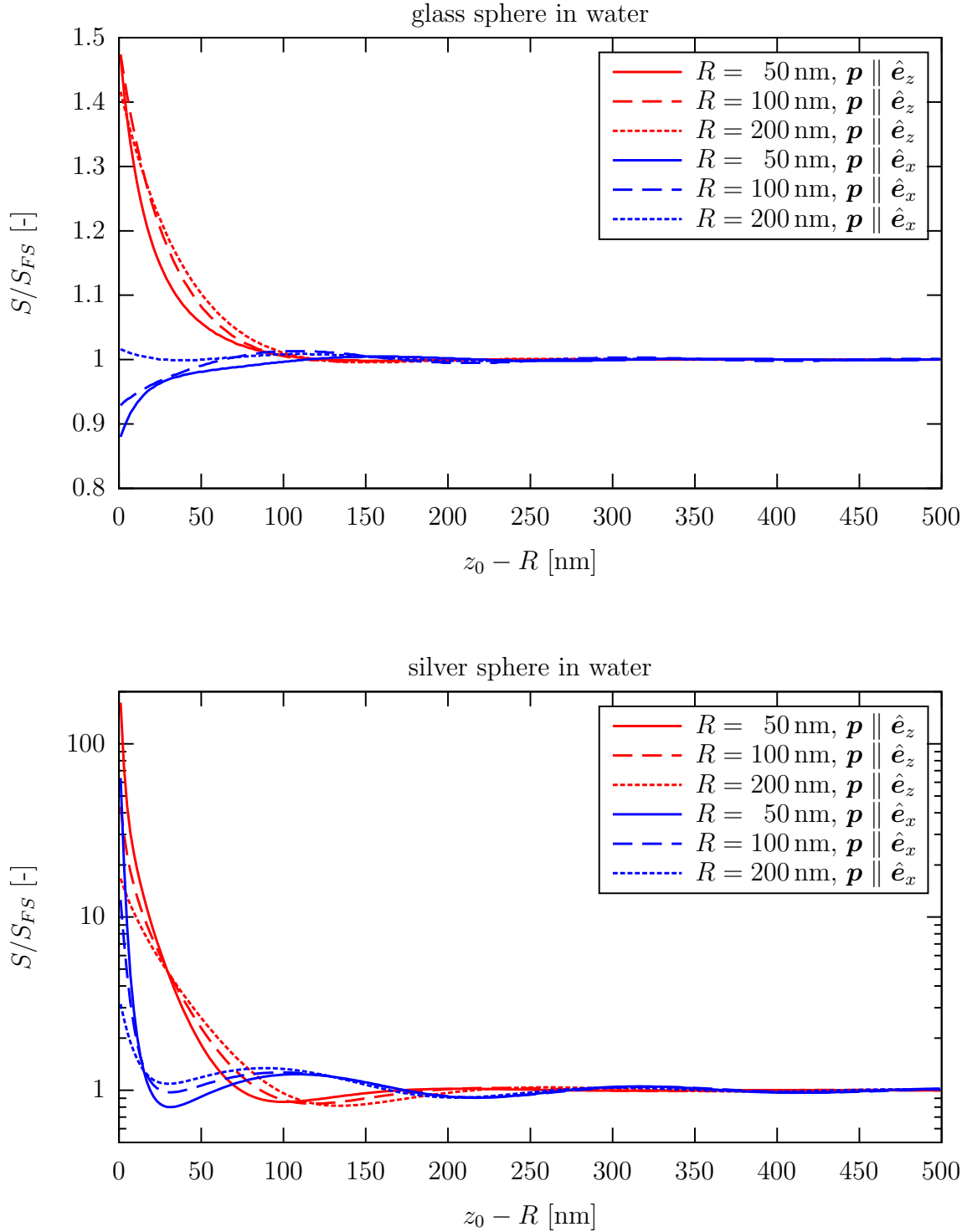
**Figure 2.21:** Total power of emission $S_{\text{tot}}$ of a water-immersed dipole located on the $z$-axis at $\boldsymbol{r}_0 = (0, 0, z_0)$ and oriented parallel to the $x$- or $z$-axis (blue and red curves, respectively). The upper image shows the results for a glass sphere with radius $R = 50, 100$ or $200\,\mathrm{nm}$, centered at the origin, while the lower plot displays the same for a silver sphere. Due to the different orders of magnitude, the upper plot uses a linear and the lower plot a semilogarithmic scale. All curves are normalized to the free space value $S_{FS} = cnk_v^4 p^2/3$.

# 2.5 Fluorescence lifetime

Up to now, we have always considered a generic "oscillating dipole emitter" without specifying what type of molecule this might be in praxis. Amazingly, even though light emission by real molecules is described rigorously by quantum mechanics, most light-emitting molecules can be modeled consistently by the classical electrodynamical models presented in the previous sections. In this chapter, we will introduce several concepts, namely fluorescence, fluorescence lifetime and fluorescence lifetime imaging, and explain how they can be combined with the techniques developed in the previous sections to experimentally collect information about dipole emitters.

## 2.5.1 Fluorescence

Some hands-on science museums offer an interesting experiment: Visitors are given a flashlight before entering a small room that is illuminated by a yellow lamp. Several paintings are hung on the wall, all seemingly only in shades of yellow and black. However, as soon as the visitor turns on the flashlight and points it on the paintings, they suddenly explode in bright shades of red, blue and green. How can this be? When monochromatic light (i.e. light with a single wavelength) reaches our eyes, we perceive it as coloured. Short wavelengths between roughly 400 nm and 490 nm appear blue, followed by green up to about 560 nm, narrower regions of yellow and orange light, and finally red between 620 nm and 710 nm. Shorter or longer wavelengths are invisible to humans. However, outside of a laboratory it is rare to observe monochromatic light – light sources such as the sun, incandescent light bulbs or even modern white light-emitting diodes create light with several different wavelengths simultaneously. In everyday life, the colour of objects is determined by the absorption and reflection spectra of their components. For example, black materials absorb all visible light, while white objects reflect this light. Coloured substances such as pigments selectively absorb some wavelengths, reflecting or scattering all others. This reflected spectrum is then perceived as colour, for example "red" if all short wavelengths are absorbed. The important point to note here is that such substances can only subtract from the original spectrum, not add any new wavelength components. This was exploited in the mentioned experiment: The yellow lamp is a sodium-vapor lamp which produces monochromatic light at a wavelength of $\lambda = 589$ nm. This light is absorbed or scattered to varying degrees by the pigments in the painting's oil colours, but the resulting spectrum still only consists of light with $\lambda = 589$ nm, thus appearing yellow. As soon as the white light from the flashlight illuminates the paintings, the different pigments produce varying reflected spectra and thus the appearance of red, blue or other colours.

There is, however, a class of substances with strongly different behaviour. When these materials are illuminated with monochromatic light, it is partly absorbed and partly transmitted, but additionally, light with a different (usually longer) wavelength is emitted a short time later. This process is called *fluorescence*. Substances with this unique property have a wide range of applications. For example, the aptly named fluorescent lamp is a low-pressure mercury-vapor gas-discharge lamp that would normally only produce invisible short-wavelength light. However, a fluorescent phosphor coating on the inside of the lamp absorbs the high-frequency photons and emits light in the

visible wavelenght range, making the lamp useful for lighting purposes. Another field where fluorescent substances have found widespread use is light microscopy. Like the pigments in the oil colours mentioned previously, fluorescent substances can be used as "dyes" to colour something. In microcopy, they can be employed to selectively label structures of interest. We will elaborate on this application at the end of this section. Fluorescence is described both in many textbooks, such as [57], and reviews on fluorescence mircroscopy and spectroscopy, such as [58]. Therefore, we restrict ourselves to a brief summary here, which is based on the two mentioned sources.

In reality, fluorescent substances, also called *fluorophores*, are usually molecules with many atoms and complex molecular electronic orbitals. However, the underlying principle can be well understood by considering just two electronic states, the ground state $S_0$ and the excited state $S_1$ with energies $E_0 < E_1$, which are each split into several sublevels of similar energy by vibronic and rotational states[20]. The system with all relevant processes is described by the Jabłoński diagram in figure 2.22. A few findings from quantum mechanics are needed to comprehend all relevant aspects of the situation. Firstly, all elementary particles, including electrons, carry an intrinsic form of angular momentum called *spin*. While the magnitude of this angular momentum only depends on the type of particle and cannot be changed, the direction of the spin can take on different values. To be more precise: The component of the angular momentum measured along any direction can only take on certain values. For electrons, one commonly chooses the $z$-direction, and finds that the component of the spin along $z$ is either $+\hbar/2$ or $-\hbar/2$. This is then referred to as "spin up" or "spin down" and described by the quantum number $\pm 1/2$. Secondly, according to the *Pauli exclusion principle*, identical elementary particles with half-integer spin cannot simultaneously exist in the same state. Here, the term "state" includes properties such as position, velocity, and several quantum numbers that also determine the energy level. As a consequence, two electrons can only occupy the same orbital and be on exactly the same energy level if they carry opposite spin. If the spin of one of the electrons were flipped, the electron would have to transition to a different energy level. Thus, in the ground state, all electrons exist in pairs of opposite spin, resulting in zero net spin – the ground state is a so-called singlet state and is therefore denoted by $S_0$. Without any external influences, the steady-state distribution of the occupation of the energy levels with energy $E_i$ is given by the Boltzmann distribution, $p_i \propto \exp(-E_i/k_B T)$ with the absolute temperature $T$ and Boltzmann's constant $k$. At room temperature, most molecules occupy one of the lower vibronic levels of the electronic ground state. They can absorb a photon if its energy ($E_\lambda = hc/\lambda_{\mathrm{abs}}$ with Planck's constant $h$) matches the energy difference between the current and any higher energy level. Generally, this process cannot flip the spin of the electron, therefore the transition from $S_0$ to the triplet state $T_1$ is called spin-forbidden. Instead, the excited state $S_1$ is also a

---

[20]The field of quantum chemistry studies – among other problems – the properties of molecular orbitals. One of the concepts is that functions called orbitals describe the probability to find an electron at certain positions in space. These orbitals are associated with energy levels, and rules exist that determine which orbitals are occupied by electrons. Generally, the orbitals of lowest energy are occupied first, with higher-energy orbitals remaining empty. Then, the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) assume the roles of ground state and excited state, respectively. For more details, see e.g. [59].
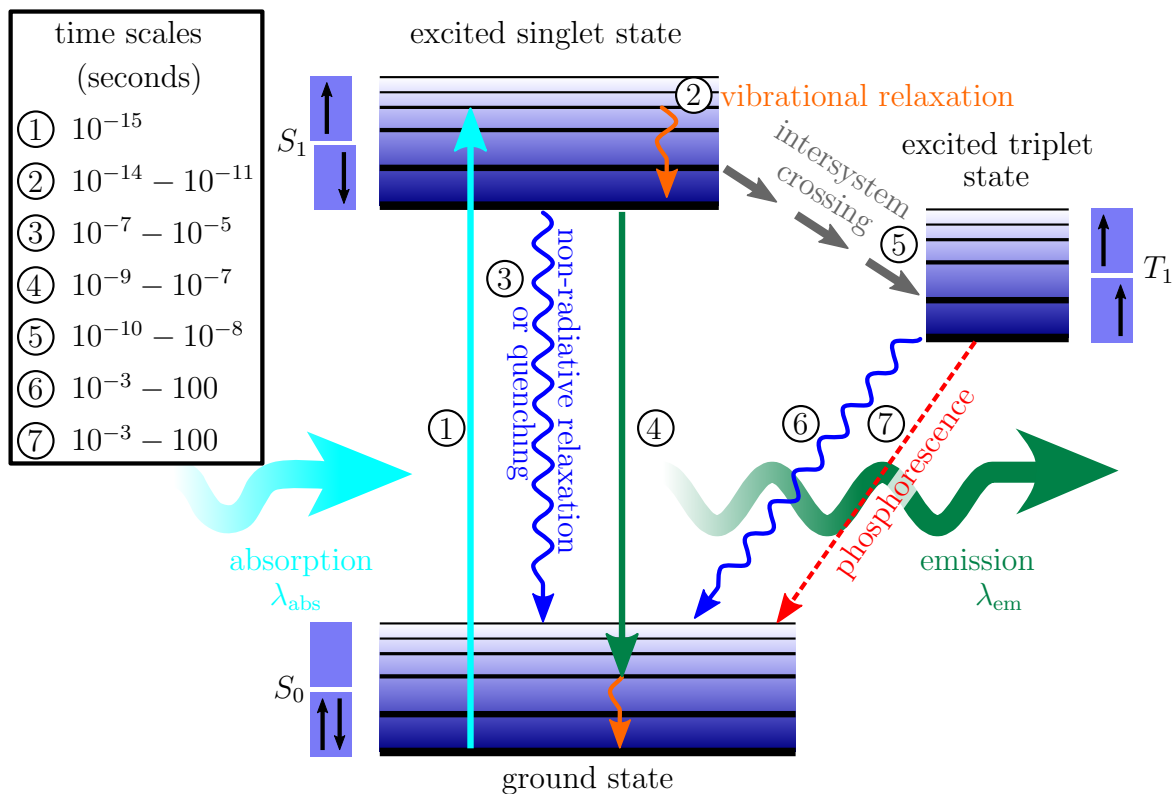
**Figure 2.22:** Jabłoński diagram, adapted from [58]. Shown are the electronic singlet ($S_0$, $S_1$) and triplet ($T_1$) states with vibronic sublevels (black lines), as well as possible transition pathways between them. Typical timescales for these processes are listed in the top left corner. The spins in opposite or parallel directions are shown schematically as small arrows next to $S_0$, $S_1$ and $T_1$. For details on the transitions see main text.

singlet state. Any of the vibronic energy levels of the excited electronic level can be reached, depending on the exact energy of the photon and on the overlap between the initial and final vibrational wave functions (*Franck-Condon principle*). Because of this, when the molecule is irradiated with a continuous spectrum of different wavelengths, some are more efficiently absorbed than others. This is illustrated by the *absorption spectrum* of the fluorophore, see figure 2.23. In a process called *vibrational relaxation*, the molecule then quickly ($\sim 10^{-12}$ s) relaxes to the lowest vibrational level of the excited electronic state, converting the excess energy to heat, which is transferred via collisions to neighbouring molecules. From this lowest level of $S_1$, the electron can then relax to one of the vibrational levels of $S_0$ while emitting a photon whose wavelength $\lambda_{em}$ is again determined by the energy difference of the two levels. Since the structure of the vibronic sublevels of $S_0$ and $S_1$ is often similar, the spectrum of the emitted light is then a near-mirror image of the absorption spectrum (*mirror image rule*). Because of the vibrational relaxation, the emission spectrum is usually independent of the excitation wavelength (*Kasha's rule*), and the wavelength of the emitted light is longer than that of the absorbed light (*Stokes shift*). Besides this radiative transition, the molecule can also relax to the ground state non-radiatively, for example through collisions with solvent molecules. The radiative and non-radiative relaxation channels are described by
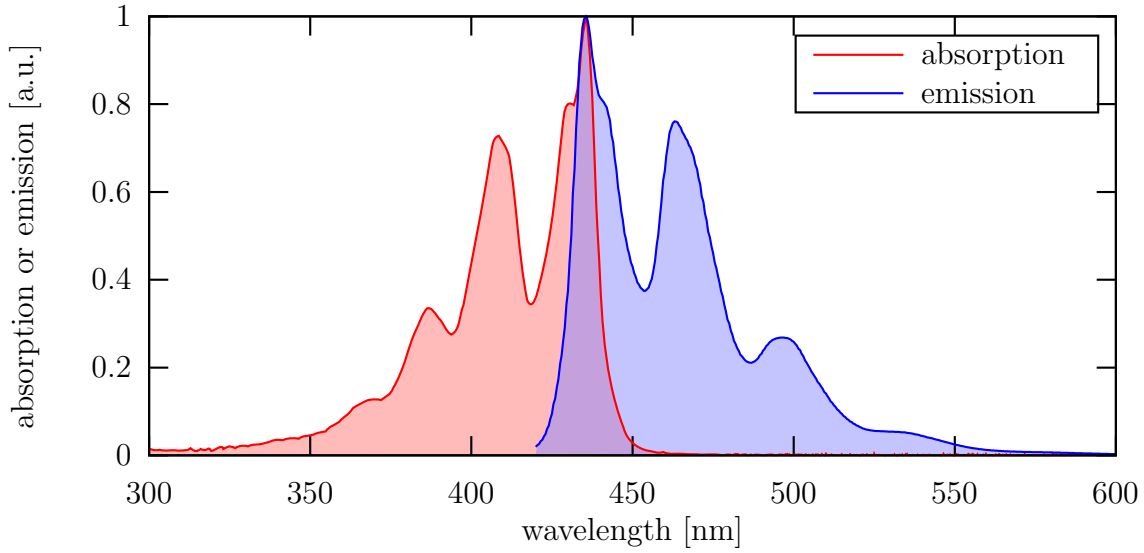
**Figure 2.23:** Normalized absorption and emission spectra of perylene, data from [60]. Both spectra are almost mirror images of each other. For this molecule, the Stokes shift – i.e. the wavelength difference between the absorption and emission maximum – is zero.

their rate constants $k_\mathrm{r}$ and $k_\mathrm{nr}$, respectively. If an ensemble of identical molecules is excited at time $t = 0$, the number $N(t)$ of molecules still in the excited state at time $t$ is described by

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = -k_\mathrm{r}N(t) - k_\mathrm{nr}N(t) \quad \Rightarrow \quad N(t) = N(0) \cdot \mathrm{e}^{-(k_\mathrm{r}+k_\mathrm{nr})t}. \tag{2.195}$$

Equivalently, for one single molecule that was excited at time $t = 0$, the probability $p(t)$ to still be in the excited state at time $t$ is given by

$$p(t) = k\mathrm{e}^{-kt}, \tag{2.196}$$

where we introduced the total rate $k := k_\mathrm{r} + k_\mathrm{nr}$. This rate determines the average time a molecule spends in the excited state, called the *excited state lifetime*, *fluorescence lifetime* or simply *lifetime* and denoted by $\tau$:

$$\tau := \int_0^\infty \mathrm{d}t \, t \cdot p(t) = -k\frac{\partial}{\partial k}\int_0^\infty \mathrm{d}t \, \mathrm{e}^{-kt} = -k\frac{\partial}{\partial k}\frac{1}{k} = \frac{1}{k}. \tag{2.197}$$

Thus, a change in either the radiative or non-radiative relaxation rate has a direct impact on the fluorescence lifetime. From an experimental point of view, it is interesting to know how many emitted photons to expect for a certain number of absorption events. This ratio, called the *quantum yield* (QY) $\Phi$, is given by:

$$\Phi = \frac{k_\mathrm{r}}{k} = k_\mathrm{r}\tau. \tag{2.198}$$

Finally, there is a small probability for an *intersystem crossing*: The excited singlet state can pass non-radiatively to the excited triplet state, where it then stays for a time that is orders of magnitude longer than $\tau$ because the return to $S_0$ is again spin-forbidden. If the relaxation occurs radiatively, the process is called *phosphorescence*, otherwise the term *triplet-state blinking* is used to describe the long dark period. To be exact, the rate of intersystem crossing should be included in the total deexcitation rate $k = k_\text{r} + k_\text{nr} + k_\text{IC}$. However, since the probability of an intersystem crossing is so small, or equivalently the transition rate $k_\text{IC} \ll k_\text{r}, k_\text{nr}$, this process is usually ignored when calculating $k$.

One important point still has to be addressed to complete this summary: What are the advantages of fluorophores compared to other dyes? Due to the nature of this thesis, we answer this question for the life sciences. While some specialised applications such as automated sequencing of DNA by the chain termination method [61] or DNA detection in agarose gel electrophoresis [62] exist, the most widely known field is fluorescence imaging. Gaining a deeper understanding of the processes inside single cells is an important aim of modern biology. One of the key players of these processes are proteins, polypeptides that are made up of one or several chains of amino acids. Unraveling their structure and dynamics is a crucial step towards deducing the complex interplay of the components necessary for cellular processes. There is, however, one property of proteins that makes them hard to study: They are small, usually just a few nanometers in diameter [63], and have a small absorption coefficient, making them effectively transparent in conventional widefield microscpy. Here, fluorescent labels offer the possibility to discriminate between a structure of interest and all other molecules present in the sample. Since the label emits light of a different wavelength than the one used to illuminate the sample, it is easily distinguished from scattered illumination light of other parts of the sample, allowing to localize the emitter.

There are several types of fluorescent markers, among them organic dyes, fluorescent proteins or semiconductor nanocrystals called quantum dots (QDs). Each of these have their advantages and disadvantages, for example, while organic dyes have a size of roughly 1 nm, fluorescent proteins measure about 3 nm to 5 nm and coated QDs have a diameter of 10 nm-15 nm [64]. Larger labels may interfere with the function of the labeled protein or worsen the localization precision. The labels also differ in emission properties. On average, quantum dots are brighter than organic dyes and fluorescent proteins [65]. Furthermore, QDs can be imaged over arbitrarily long time periods, whereas both organic dyes and fluorescent proteins become nonemissive after prolonged excitation – they are said to *photobleach*. However, both organic dyes and quantum dots have to be linked to the target protein via (at least) one other macromolecule, such as an antibody or a streptavidin-biotin bond. This increases the distance between the detectable fluorescent marker and the protein whose location is of interest. Also, antibodies can be labeled with several dye molecules, which on the one hand increases the signal but on the other hand hinders exact counting of molecules. Fluorescent proteins are covalently bound to the protein of interest during expression of the protein, therefore the labeling is highly specific and monovalent. Furthermore, organic dyes and QDs can only enter the cell if the cell membrane has been permeabilized, which is often only possible after fixation of the cell and which might actually change the morphology of the cell and thus the structures of interest. Fluorescent proteins, on the other hand,

are ideally suited for life-cell imaging.

But fluorescent labels can do more than just mark specific structures: Some of them have environment-specific properties such as solvent-dependent emission spectra [66], pH-dependent fluorescence lifetime and quantum yield [67, 68], viscosity-dependent fluorescence lifetime [69] or even quenching of the fluorescene by the presence of certain substances ([57], chapter 9). This makes them suitable as sensors for a variety of microscopic properties in their immediate surrounding which might be hard to access otherwise, for example intracellular ion concentrations [70] or intracellular pH [71]. Furthermore, the fluorescence lifetime of an emitter can be changed by the presence of metal or dielectric nanostructures, mirrors or simply interfaces between regions with different refractive indices. This will be explained in detail in section 2.5.4 and is the main driving force behind this thesis. Before coming to this exciting topic, however, we will describe how exactly fluorescence lifetimes can be measured.

## 2.5.2 Fluorescence microscopy and fluorescence lifetime imaging

In fluorescence microscopy, one is interested in a certain aspect of a sample, for example the shape and size of a fluorescently labeled structure, or the exact position of a single fluorophore that has been attached to a small object. These fluorescence intensity patterns carrying relevant information are called the *signal*. Detected intensity that arises from other sources, such as the excitation light, contaminations of the sample, or out-of-focus fluorescence is denoted as *background* or background signal. Discriminating between signal and background is essential to recover the information one wants to gain from the sample.

This process, and also the subsequent interpretation of the signal, is complicated by the fact that every fluorescence intensity measurement is inherently noisy. In the previous section, we showed that a single fluorophore emits light with a constant rate $k_r$. However, since the emission of a photon is a binary process[21], one can find sufficiently small time intervals of duration $\Delta t$ in which either one photon is generated or no photons are emitted. Then, the probability for light emission within the short time interval $\Delta t$ is given by $k_r \Delta t$. The angular distribution of radiation determines which fraction of emitted light can be collected by the objective. Together with losses in the optical path between objective and detector, as well as a finite detection efficiency of the detector itself, this means that the probability to detect a photon within the time interval $\Delta t$ is given by $k_{det} \Delta t$, where $k_{det}$ is proportional to $k_r$[22]. Then, the number $N$ of detection events in a longer time interval $T$ (i.e. the measured intensity for exposure time $T$) is a random number and follows a *Poisson distribution* with mean $k_{det}T$ ([72], §16.2.6.2, p. 833):

$$P(N) = e^{-k_{det}T} \frac{(k_{det}T)^N}{N!}. \tag{2.199}$$

---

[21]At each given time point, the event either occurs or does not occur, but a molecule cannot emit "0.3 photons".

[22]In a rigorous treatment, the probability that the fluorophore is even excited within a certain time interval also has to be taken into account. Since the main message, namely that the detection probability is proportional to the length of the time interval, is not changed by this aspect, it was omitted here for simplicity.
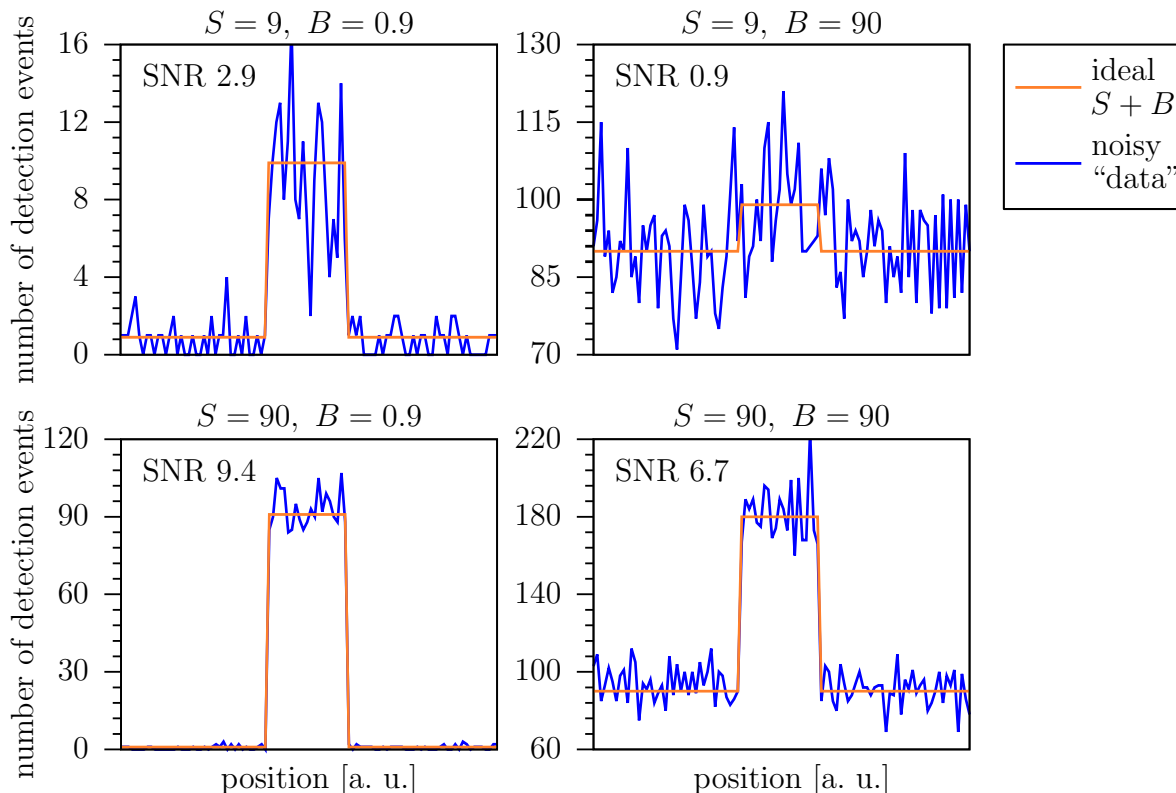
**Figure 2.24:** Simulation of a one-dimensional intensity measurement with shot noise. The signal $S$ of a fluorescently labeled structure is superposed by a constant background $B$, the resulting ideal intensity is shown as a red line. A measurement would result in noisy data (blue line), where the signal in each pixel follows a Poisson distribution with mean $S + B$. Different combinations of signal $S$ and background $B$ lead to different signal-to-noise ratios $\text{SNR} = S/\sqrt{S + B}$. The fluorescently labeled structure can be distinguished more clearly from the background for a high SNR compared to a low SNR, while it is *de facto* invisible for a SNR smaller than one.

Thus, even if the sample itself does not change, the number of detection events will vary between consecutive measurements – we say that the data is noisy. Mathematically, this is quantified by the *variance* $\sigma^2$, the expectation of the squared deviation of $N$ from $k_{\text{det}}\Delta t$. The *standard deviation* $\sigma$, which is the square root of the variance, is commonly called the *noise*, it has the same dimension as the signal itself. For a Poisson distributed quantity, the variance is simply equal to the mean, and thus the standard deviation equals the square root of the mean. This has implications for the *signal-to-noise ratio* (SNR), the ratio of the desired signal to the noise:

$$\text{SNR} := \frac{\text{signal}}{\text{noise}} \overset{\text{Poisson}}{\underset{\text{noise}}{=\!=}} \frac{N}{\sqrt{N}}. \tag{2.200}$$

This type of noise is called Poisson noise or shot noise[23] in order to distinguish it

from other types of noise. The latter arise mainly in the detector itself, for example, a detector can report a detection event even though no photon was present (*dark noise*). However, in an ideal setup, shot noise is the dominating form of noise. Therefore, we illustrate its effect in figure 2.24 that simulates a one-dimensional intensity scan. In all four panels, we assume a constant background $B$ that is superposed by the signal $S$ from a fluorescently labeled one-dimensional structure. In red, the total ideal intensity $B + S$ is shown, while the blue curves are simulations of shot-noise limited measurements. When calculating the signal-to-noise ratio, one has to take into account that the total number of photons equals $B + S$, while the signal is only given by $S$:

$$\text{SNR} = \frac{S}{\sqrt{B + S}} \tag{2.201}$$

Therefore, increasing the signal increases the SNR, while increasing the background decreases the SNR. The aim of a measurement of this system could, for example, be to localize the labeled structure and determine its size. As can be seen in the figure, this is easier for a higher SNR. Equivalently, the reliability of fits of model functions to curves such as these is increased by a higher SNR, too. To sum up, it is desirable to achieve a high SNR experimentally. This challenging task is one of the main considerations when designing an experimental setup for fluorescence microcopy, as we will now demonstrate exemplarily on a setup used in our laboratory (see figure 2.25).

As stressed in the previous section, the main feature of fluorescence is the different wavelength of the emitted radiation compared to the excitation light. This effect can be used to separate fluorescence and scattered or transmitted excitation light by specialized optical components. The simplest possibility are absorptive filters, glass or plastic plates that contain pigments which selectively absorb some wavelengths, letting all others pass. More sophisticated components are *dichroic filters* (also called *dichroic mirrors*), glass plates that are coated with a series of thin films with different refractive indices. Due to multiple reflections at these many interfaces, some wavelengths interfere constructively behind the filter and can thus pass it, while others interfere destructively and are reflected instead. Depending on the type, thickness and number of coatings, there is usually a so-called cutoff wavelength below or above which all light is transmitted (shortpass or longpass filter, respectively), all other light is reflected. Furthermore, there exist bandpass filters that only allow light within a certain wavelength range (e.g. the fluorescence light) to pass. Since dichroic filters split the light beam into a transmitted and a reflected component, they are also called *beam splitters*. If the cutoff wavelength of the filter is chosen to lie between the wavelengths of the excitation light and the fluorescence, the two types of radiation can be efficiently separated, allowing to detect only the fluorescence emission and thus considerably reduce the background. However, since these filters are never perfect, but always have an – albeit very small – non-zero transmission coefficient in the reflective wavelength range (usually on the order of $10^{-6}$), and since the excitation intensity is usually much higher than the fluorescence intensity, it is desirable to let as little excitation light as possible enter the detection pathway. This is achieved in our setup by using an inverted microscope, where both excitation

---

[23]This is a translation of the German "Schroteffekt", which Walter Schottky used in the first description of this type of noise for current fluctuations in vacuum tubes [73].
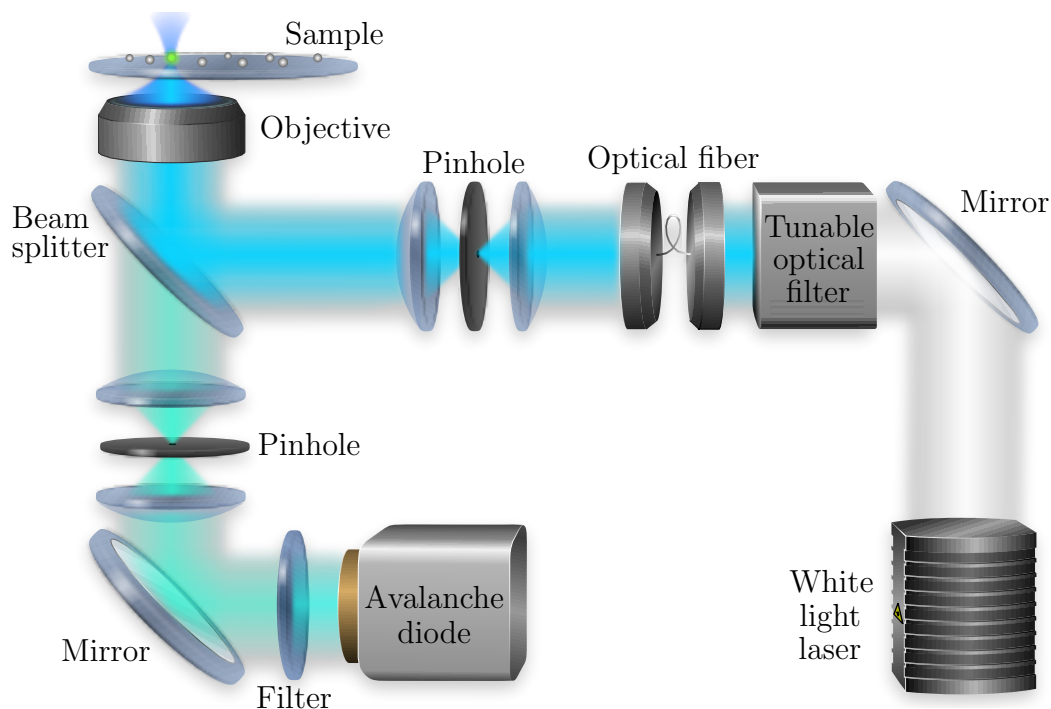
**Figure 2.25:** Schematic of the experimental setup for confocal fluorescence lifetime imaging. A monochromatic light source (here a white light laser with a tunable optical filter to allow rapid switching between wavelengths for multi-colour measurements) is focused through a pinhole onto the sample. The fluorescence light from the small focal spot is then focused onto a second pinhole, residual excitation light is blocked by an additional filter. Adapted with permission from an image created by Alexey Chizhik.

and detection are done from below the sample (see figure 2.25). Then, the beam splitter and any further filters which might be added directly before the detector only have to block scattered excitation light, instead of the full power of the initial beam.

We said above that in an ideal setup, shot noise is the dominating form of noise. This is true in a real setup, too, if the detector has a low noise level. For fluorescence imaging, complementary metal-oxide semiconductors (CMOS) or electron-multiplying charge-coupled devices (EMCCDs) with a cooled chip are usually used, they both measure the intensity of the incoming light during a certain exposure time (usually on the order of milliseconds). Both types of cameras are used in our laboratory. However, the setup shown in figure 2.25 is used to determine fluorescent lifetimes. As we will see below, this requires the determination of the arrival times of single photons with an accuracy on the order of picoseconds. This is achieved using *single photon avalanche diodes* (SPADs) [74], which were used for the experiments presented in chapter 4.

The shot noise itself can be reduced by lowering the background, as demonstrated in figure 2.24. For a setup that is shielded from ambient light and that uses good filters, this background signal is mainly fluorescence from out-of-focus planes. This is minimized in our setup by using a *confocal microscope*. The general idea of confocal microcopy is to illuminate only one point in the sample, focus the emitted light onto a small pinhole and detect only the light that manages to pass the pinhole. Light that originates from outside the focus is smeared out over the plane of the pinhole and thus

does not reach the detector, reducing the background signal. Since only one point is illuminated and imaged at a time, this point has to be scanned over the sample in order to collect a whole image. This is in contrast to standard *widefield microscopy* where a large field of view is imaged at the same time.

In equation (2.201), we showed that increasing the signal also increases the SNR. This can be achieved by raising the number of emitters (i.e. by increasing the labeling density) or by boosting the number of photons collected from each single emitter. The latter is implemented in our setup by using a high-power excitation laser that efficiently excites the fluorophores. Lasers are monochromatic, which facilitates the excitation of only specific types of fluorophores and the selective blocking of the excitation light in the detection pathway. Furthermore, they have a high energy flux in a small solid angle element, allowing to collimate the beam and focus it on the sample. Last but not least, lasers offer the probability to be pulsed, and their amplitude can be electronically regulated, properties that are needed for determining lifetimes, as will be explained below.

Up to now, we mainly discussed experimental methods and optical components for fluorescence microscopy, where only the intensity of fluorescence at a point in space is of interest. However, as presented in the previous sections, the lifetime of a fluorophore is an observable that offers many interesting insights, too. The general term *fluorescence lifetime imaging microscopy* (FLIM) encompasses different techniques to measure fluorescent lifetimes, which can be divided into frequency-domain and time-domain methods [57]. In a frequency-domain measurement, the sample is excited with a periodically modulated light beam, where the modulation frequency $\omega_{\mathrm{mod}}$ is on the same order of magnitude as the inverse decay time of the excited state. The intensity of the detected fluorescence is then modulated with the same frequency, however, due to the finite lifetime of the excited state, the emission is delayed relative to the excitation. This leads to both a phase shift $\varphi$ and a demodulation $m$ of the fluorescence signal, see the left side of figure 2.26a. For a *monoexponential decay* (i.e. a decay with a single decay time, as in equation (2.196)), the decay time $\tau$ can be directly recovered from $\varphi$ or $m$:

$$\tau_\varphi = \frac{1}{\omega_{\mathrm{mod}}} \tan \varphi \quad \text{or} \quad \tau_m = \frac{1}{\omega_{\mathrm{mod}}} \sqrt{\frac{1}{m^2} - 1}. \tag{2.202}$$

However, a sample can also contain species with different lifetimes, for example different types of fluorophores, fluorophores in different environments, or even fluorophores that exhibit multiple emitting states with different lifetimes [75]. Then, the total number $N(t)$ of molecules that is still in the excited state at a certain time $t$ (compare with equation (2.195)) is described by a linear combination of exponential functions with different decay constants. Therefore, such a decay is called a *multiexponential decay*. When measuring such a sample, the apparent lifetimes $\tau_\varphi$ and $\tau_m$ "represent a complex weighted average of the decay components" [57]. Then, measurements have to be made for a wide range of modulation frequencies, and the data has to be fitted with appropriate models. An example of real data taken from [57] is shown on the left side of figure 2.26b. Here, phase and demodulation were measured over a range of light modulation frequencies for the tryptophan derivative N-acetyl-L-tryptophanamide
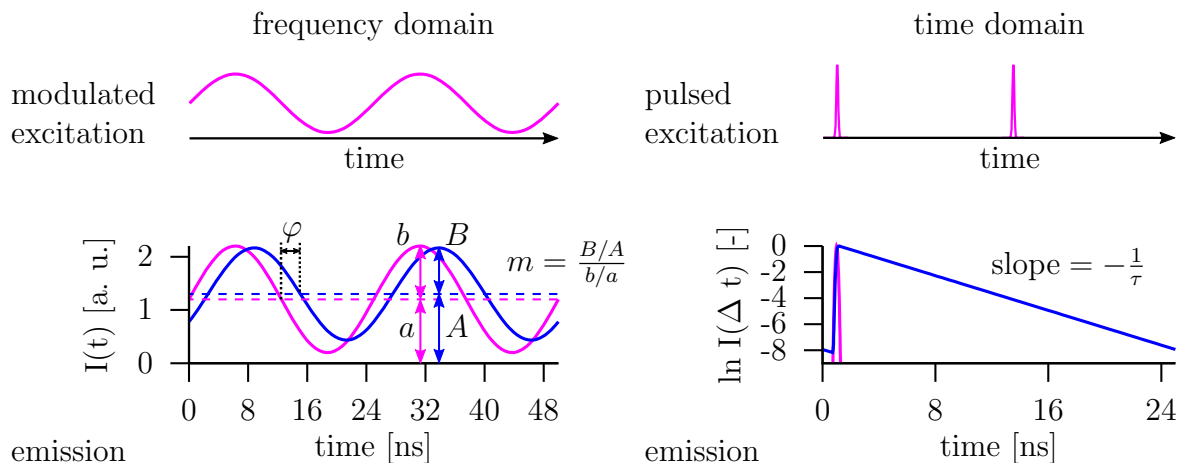
**(a)** General principles of frequency-domain (left) and time-domain (right) methods for determining the fluorescence lifetime. The excitation, shown in magenta, is either periodically modulated with sine-like functions (i.e. roughly equal amount of time with high or low intensity) or is a train of short pulses. In the frequency domain, one measures the phase shift $\varphi$ and the demodulation $m$ of the fluorescence signal, shown in blue. In the time domain, a mono- or multiexponential decay law is fitted to the intensity distribution, also in blue. For a monoexponential decay, the slope of the semilogarithmic plot equals the inverse of the lifetime. Adapted from [57], p. 98, with a fluorescence lifetime of $\tau = 3\,\text{ns}$, a modulation frequency / pulse repetition rate of $40\,\text{MHz}$, $a = 1.2$, $b = 1$, $A = 1.3$, and Gaussian pulses with a full width at half maximum of $150\,\text{ps}$.
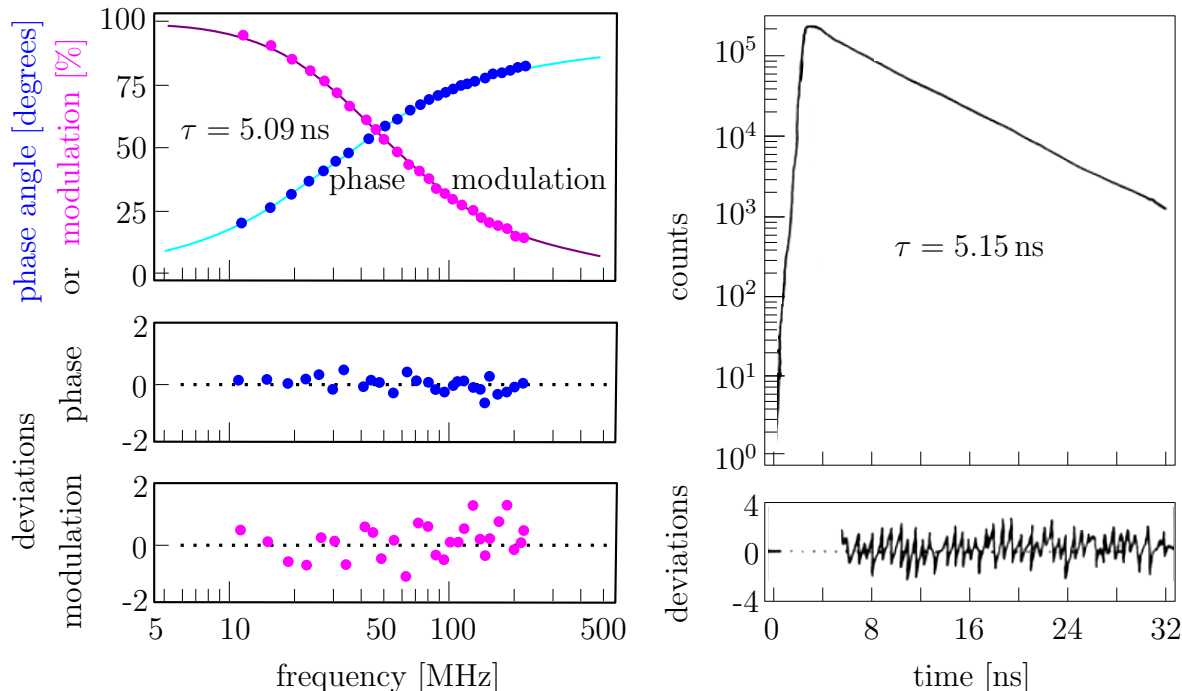


**(b)** Comparison of frequency-domain (left) and time-domain (right) decay time measurements of N-acetyl-L-tryptophanamide. Reproduced from [57], p. 100. For the time-domain measurements, the fit is not separately visibile because it overlaps strongly with the measured data. The lower panels show the deviations between measured data and fit, weighted by the standard deviations of each measurement.

**Figure 2.26:** Comparison of frequency-domain and time-domain measurements of fluorescence lifetimes, from [57].

(NATA). A least-squares fit of the data to equations (2.202) resulted in a lifetime of 5.09 ns. The deviations between data and fit shown in the two bottom left panels of figure 2.26b are random, which indicates that the decay truly is monoexponential.

In time-domain experiments, the sample is excited by very short light pulses, and the time-dependent intensity is recorded as illustrated on the right side of figure 2.26a. The probability that a single molecule that was excited at time $t$ is still in its excited state at time $t + \Delta t$ was described in equation (2.196) as

$$p(\Delta t) = \frac{1}{\tau} \mathrm{e}^{-\Delta t/\tau} \tag{2.203}$$

for a monoexponential decay. Consequently, the probability that this single molecule decays radiatively at time $t + \Delta t$ is proportional to the time derivative of $p(\Delta t)$. Under the assumption that all excited molecules follow the same law, the intensity $I(t + \Delta t)$ measured at the time $t + \Delta t$ is also proportional to $\mathrm{e}^{-\Delta t/\tau}$. This is only valid for times $\Delta t$ that are smaller than the time between consecutive excitation pulses, though. If excitation pulses occur periodically at times $t = 0, T, 2T, \ldots$, the time-dependent intensity is also periodic with period $T$. Therefore, one can sum the intensities period-wise and obtain a curve of the form

$$I(\Delta t) := \sum_{\{\tau \ \mid \ \tau \bmod T = \Delta t\}} I(\tau) \quad \propto \mathrm{e}^{-\Delta t/\tau} \qquad \forall\, 0 \leq \Delta t \leq T. \tag{2.204}$$

In practice, one has to take into account two properties of detectors: Firstly, a detector needs some time to recuperate between consecutive detection events. This time period, where the detector is effectively "blind", is called the *dead time* of the detector. Because the dead time is usually much longer than the fluorescence lifetime [6], only one photon can be detected per excitation pulse. Therefore, the excitation power has to be low enough that on average, less than one photon reaches the detector per excitation pulse – otherwise, only the first photon would be detected, leading to a shift of the average measured arrival time to shorter times. Alternatively, we have developed a method to correct this effect in a post-processing step, see [6] and section 6.3 in the appendix. Secondly, detectors only have a finite time resolution. Thus, instead of recording continuous values of arrival times, $\Delta t$ can only take on certain values, and $I(\Delta t)$ is actually a histogram.

This concept is implemented in *time-correlated single photon counting* (TCSPC). There, the sample is excited with a periodic train of short laser pulses, and the relative delay $\Delta t$ between the detection of a photon and the exciting pulse is recorded. The resulting photon number histogram $I(\Delta t)$, the so-called *TCSPC curve*, can then be fitted with a mono- or multiexponential model function to recover the fluorescent lifetime(s). Details on the fitting routines used by us are presented in section 3.5. The same example as shown for frequency domain measurements was also measured in the time domain, see the right side of figure 2.26b. A monoexponential fit of the data resulted in a lifetime of 5.15 ns, which is close to the result of the frequency domain measurement. The deviation between data and fit in the bottom right panel again seems random, which is an indicator for a correct choice of the model.

### 2.5.3 Lifetime changes close to interfaces

We showed in section 2.3 that the electric field created by an osillating electric dipole is changed close to interfaces between media with different refractive indices. This was followed by the demonstration how both the angular distribution of radiation and the total amount of energy emitted per time are influenced by the interface in section 2.4. Especially the last point suggests that there might also be a variation in fluorescent lifetime close to an interface: If more energy per time is emitted, but the energy of one single photon stays the same, should there not be more emission events per time? This question was answered theoretically by Edward Purcell in 1946 in the context of nuclear magnetic moment transitions at radio frequencies. He predicted that the spontaneous emission rate of atoms inside a resonant cavity should be increased, with the magnitude of the effect depending on certain properties of the cavity [36]. For transitions in the visible range, this *Purcell effect* was demonstrated experimentally by Karl-Heinz Drexhage in 1970 [37]. He used a technique known since the 1930s called *Langmuir-Blodgett films* (LB films) [76] to create a layer of known thickness $d$ on top of a silver mirror. With this method, a monolayer of fatty acid molecules is formed in a bath by dropping a small amount of fatty acid onto the liquid. Then, the solid substrate on which one wishes to deposit the film is dipped into the liquid and carefully removed again, leaving behind a monolayer of fatty acid on the substrate. Since the thickness of one monolayer is known, by repeating the process one can create layers of (almost) arbitrary thickness. Drexhage doped the top layer with europium-dibenzoylmethane complexes, knowing that $Eu^{3+}$ ions absorb in the ultraviolet and fluoresce at about 612 nm with a decay time of about one millisecond. He measured the fluorescence lifetime for different film thicknesses $d$ and observed a strong variation of $\tau(d)$.

Several models exist which explain these lifetime changes, and which all lead to the same predictions. Since the experiment concerns single ions, it seems appropriate to start with a quantum mechanical approach. As we already pointed out when introducing fluorescence in section 2.5.1, the excited fluorophore can be approximated as a two-state system. Let us assume the initial (excited) state has the energy $\mathcal{E}_i$, while the final (ground) state has the energy $\mathcal{E}_f$. Both are eigenstates of the Hamiltonian $H_0$ that describes the fluorophore in free space. If an electric field $\boldsymbol{E}$ with angular frequency $\omega = (\mathcal{E}_i - \mathcal{E}_f)/\hbar$[24] is acting on the system, the total Hamiltonian $H$ is given by $H = H_0 + H'$, where $H' = -\boldsymbol{p} \cdot \boldsymbol{E}$ with the dipole moment $\boldsymbol{p}$. Then, quantum mechanical perturbation theory provides an expression for the transition rate $\Gamma = \tau^{-1}$ between the initial state $i$ and the final state $f$, known as *Fermi's golden rule* ([77], section 9.2):

$$\Gamma_{i \to f} = \frac{2\pi}{\hbar} |\langle f|H'|i \rangle|^2 \rho(\omega), \tag{2.205}$$

where $\langle f|H'|i \rangle$ is the matrix element of the perturbation $H'$ between the initial and final states and $\rho(\omega)$ is the density of final states (number of states per energy range and per volume at energy $\hbar\omega$). This rule immediately shows that the lifetime is inversely proportional to $\rho$. The important point to note is that $\rho(\omega)$ becomes position-dependent, $\rho(\omega, \boldsymbol{r})$, in the vicinity of metallic or dielectric interfaces. Thus, $\rho(\omega, \boldsymbol{r})$ is called the

---

[24]$\hbar = h/2\pi$ is the reduced Planck constant.

*local density of states* (LDOS).

On the other hand, it is known that vacuum fluctuations can excite a dipole ([77], section 9.2). In thermal equilibrium, the dipole has to on average emit the same amount of energy as it absorbs, which determines the spontaneous emission rate [78]. If the density of vacuum modes is doubled, the spontaneous emission has to be doubled, too, which leads to the same dependency on $\rho(\omega, \boldsymbol{r})$ as Fermi's golden rule.

The drawback of this approach is the difficulty in calculating the LDOS for geometries that are more complicated than a rectangular box in vacuum. One possibility is described by Greffet and co-workers [56, 79]. It is based on what they call "fluctuational electrodynamics". The idea is that, in thermal equilibrium, electrons in the material fluctuate, producing randomly fluctuating currents, which in turn generate radiation. The statistical properties of these currents can be gained from the *Fluctuation-Dissipation Theorem* (FDT). The electromagnetic fields that are generated by the currents are changed due to the presence of the interface, as described in section 2.4[25]. Greffet's argument is that the average density of electromagnetic energy at the position $\boldsymbol{r}$ is given by

$$\langle \mathcal{U} \rangle(\boldsymbol{r}) = \frac{1}{8\pi} \langle |\boldsymbol{E}(\boldsymbol{r}, t)|^2 \rangle_t + \frac{1}{8\pi} \langle |\boldsymbol{B}(\boldsymbol{r}, t)|^2 \rangle_t = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} u(\boldsymbol{r}, \omega) \tag{2.206}$$

with the spectral energy density $u$. Alternatively, an energy density can be understood as the product of the local density of states and the mean energy of each state at temperature $T$:

$$u(\boldsymbol{r}, \omega) = \rho(\boldsymbol{r}, \omega) \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1}. \tag{2.207}$$

As auxiliary quantities, Greffet *et al.* introduce the electric- and magnetic-field correlation functions $\mathcal{E}$ and $\mathcal{B}$:

$$\mathcal{E}_{ij}(\boldsymbol{r}, \boldsymbol{r}', t - t') = \mathfrak{Re} \left[ \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \mathcal{E}_{i,j}(\boldsymbol{r}, \boldsymbol{r}', \omega) \mathrm{e}^{-i\omega(t-t')} \right] = \langle E_i(\boldsymbol{r}, t) \overline{E_j(\boldsymbol{r}', t')} \rangle,$$

$$\mathcal{B}_{ij}(\boldsymbol{r}, \boldsymbol{r}', t - t') = \mathfrak{Re} \left[ \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \mathcal{B}_{i,j}(\boldsymbol{r}, \boldsymbol{r}', \omega) \mathrm{e}^{-i\omega(t-t')} \right] = \langle B_i(\boldsymbol{r}, t) \overline{B_j(\boldsymbol{r}', t')} \rangle. \tag{2.208}$$

Then, by setting $\boldsymbol{r} = \boldsymbol{r}'$ and $t = t'$, one finds that the spectral energy density is given by

$$u(\boldsymbol{r}, \omega) = \frac{1}{8\pi} \sum_{i=1:3} \mathcal{E}_{ii}(\boldsymbol{r}, \boldsymbol{r}, \omega) + \mathcal{B}_{ii}(\boldsymbol{r}, \boldsymbol{r}, \omega). \tag{2.209}$$

Finally, the FDT relates $\mathcal{E}_{ij}(\boldsymbol{r}, \boldsymbol{r}', \omega)$ and $\mathcal{B}_{ij}(\boldsymbol{r}, \boldsymbol{r}', \omega)$ to the Green's function of the system, thus allowing to derive the LDOS $\rho(\boldsymbol{r}, \omega)$ for different geometries by comparing (2.207) and (2.209). However, this approach was mentioned here only for the sake of completeness and will not be used in the following.

---

[25]Alternatively, the influence of the interface can also be treated using the Green's tensor of the system as in [56, 79].

**Theory**

Historically, the first suggestion that accurately reproduced the experiment [80, 81] was a mechanical model, namely a classical damped oscillator that is driven by an electric field. In this model, the excited state of the dipole emitter is described by a harmonically bound, oscillating charge $e$ with effective mass $m$. Without external influences and without damping, this would be described by

$$m\ddot{\boldsymbol{x}} = -k\boldsymbol{x} \quad \Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{x}_0^+ \mathrm{e}^{i\omega t} + \boldsymbol{x}_0^- \mathrm{e}^{-i\omega t}, \tag{2.210}$$

where $k$ is the elastic force constant, $\boldsymbol{x}$ is the three-dimensional position of the charge and $\omega = \sqrt{k/m}$ is the oscillation frequency. Instead of using this general solution, we choose only the term proportional to $\mathrm{e}^{-i\omega t}$ and keep in mind that the physical solution can be found by taking the real part, as was done in all other previous chapters. By multiplying the whole equation with the charge $e$, we find an expression for the dipole moment $\boldsymbol{\mu}(t) = e\boldsymbol{x}(t)$:

$$\ddot{\boldsymbol{\mu}} = -\omega^2 \boldsymbol{\mu} \quad \Rightarrow \quad \boldsymbol{\mu}(t) = \boldsymbol{\mu}_0 \mathrm{e}^{-i\omega t}. \tag{2.211}$$

Without damping, the dipole never stops oscillating, which corresponds to an infinite lifetime of the excited state. When damping with a damping constant $b$ is introduced, the equation becomes

$$\ddot{\boldsymbol{\mu}} = -\omega^2 \boldsymbol{\mu} - b\dot{\boldsymbol{\mu}} \quad \Rightarrow \quad \boldsymbol{\mu}(t) = \boldsymbol{\mu}_0 \mathrm{e}^{-i\omega' t}\mathrm{e}^{-\frac{b}{2}t}, \tag{2.212}$$

with $\omega' = \sqrt{\omega^2 - b^2/4}$. The oscillation frequency $\omega'$ is thus shorter than the unperturbed frequency $\omega$. Furthermore, the dipole moment decays exponentially with decay rate $b/2$. Since the amplitude of the electric field generated by the dipole is proportional to $|\boldsymbol{\mu}|$, and the intensity in turn is proportional to the absolute square of the electric field, this tells us that the intensity of this classical oscillator decays with decay rate $b$. Therefore, in this model, the excited state lifetime $\tau$ is identical to $1/b$. For visible light with wavelength $\lambda \approx 380\,\mathrm{nm} - 780\,\mathrm{nm}$, the oscillation frequency is in the range $\omega = 2\pi c/\lambda \approx (2 \text{ to } 5) \cdot 10^{15}\,\mathrm{s}^{-1}$. On the other hand, the excited state typically lasts a few nanoseconds, or $b \approx 10^9\,\mathrm{s}^{-1}$. Thus, the shift in oscillation frequency caused by the damping is negligible, $\omega' \approx \omega$.

Now we come to the reason for these considerations, the interaction of the reflected field with the dipole. Any electric field $\boldsymbol{E}$, regardless of its origin, exerts a force $e\boldsymbol{E}$ on the charge, resulting in:

$$\ddot{\boldsymbol{\mu}} = -\omega^2 \boldsymbol{\mu} - b\dot{\boldsymbol{\mu}} + \frac{e^2}{m}\boldsymbol{E}. \tag{2.213}$$

However, here we reach a limit of the model: If $\boldsymbol{E}$ has components that are perpendicular to the original dipole axis, it should force the dipole to change the direction of its oscillation. In reality, the orientation of the dipole moment is fixed relative to the dye molecule itself, and we assume that the molecule (with its much larger mass compared

to a single electron) is not rotated or translated by the electric field. Therefore, in the following we project everything on the dipole axis and solve a one-dimensional equation:

$$\ddot{\mu} = -\omega^2\mu - b\dot{\mu} + \frac{e^2}{m}\,E_\parallel \tag{2.214}$$

Since we know that the electric field is generated by the dipole, both oscillate with the same frequency. Inserting the ansatz $E_\parallel(t) = E_0\exp(-i\Omega t)$, $\mu(t) = \mu_0\exp(-i\Omega t)$ in equation (2.214), we find:

$$0 = \Omega^2 - \omega^2 + ib\Omega + \frac{e^2}{m\mu_0}E_0. \tag{2.215}$$

Under the premise that the frequency $\Omega$ should approach $\omega$ in the absence of any damping or external fields, we take the postive square root:

$$\Omega = -i\frac{b}{2} + \sqrt{-\frac{b^2}{4} + \omega^2 - \frac{e^2}{m\mu_0}E_0}. \tag{2.216}$$

This expression can be simplified. As shown above, $b \ll \omega$. As a first-order approximation of $E_0$, consider the closed-form solution for the electric field of a dipole in eq. (2.79): For small $r$ and $\boldsymbol{r} \perp \boldsymbol{p}$, the magnitude of the field is approximately equal to $\mu/(n^2 r^3)$. In the cgs-system, the elementary charge has the magnitude $4.8 \cdot 10^{-10}\,\sqrt{\text{g}}\sqrt{\text{cm}}^3/\text{s}$, while the mass of an electron is $9.1 \cdot 10^{-28}\,\text{g}$. Thus, $e^2/m \approx 2.5 \cdot 10^8\,\text{cm}^3/\text{s}^2$. With this, the term $e^2/(m\mu_0)E_0$ takes on values in the range $10^{26}\,\text{s}^{-2}$ ($r = 10\,\text{nm}$) or less. Compared to $\omega^2 \approx 10^{30}\,\text{s}^{-2}$, this is small enough to warrant a linear approximation of the square root[26]:

$$\Omega \approx -i\frac{b}{2} + \omega - \frac{b^2}{8\omega} - \frac{e^2}{2m\mu_0\omega}E_0. \tag{2.217}$$

Now we interpret this result. Since we used the ansatz $\mu(t) = \mu_0\exp(-i\Omega t)$, all real components of $\Omega$ determine the oscillation frequency of the dipole,

$$\omega' = \omega - \frac{b^2}{8\omega} - \frac{e^2}{m\mu_0\omega}\,\mathfrak{Re}(E_0) \approx \omega, \tag{2.218}$$

which is not significantly changed compared to the situation without external field. Thus, the wavelength of the radiation stays basically constant. On the other hand, the imaginary part of $\Omega$ describes the damping and thus the radiative deexcitation rate:

$$\frac{b'}{2} = \frac{b}{2} + \frac{e^2}{2m\mu_0\omega}\,\mathfrak{Im}(E_0)$$

---

[26]Since we use a mechanical model to describe what is actually a quantum mechanical transition dipole, using the mass and charge of an electron is initially just a reasonable guess. However, since this model was successfully used to explain experimental results [81], the assumption of small $e^2/(m\mu_0)E_0$ seems justified.

$$\text{or} \quad \tau' = \left[\frac{1}{\tau} + \frac{e^2}{m\mu_0\omega}\ \mathfrak{Im}(E_0)\right]^{-1}. \tag{2.219}$$

If the dipole is close to the mirror, the damping due to $\mathfrak{Im}(E_0)$ can be much larger than the intrinsic damping $b$ – the lifetime of the dipole is drastically reduced compared to the free space case. In the original works [80, 81], the damping constant $b$ was already interpreted as the sum of two processes, radial damping and thermal damping, with $b = b_\mathrm{r} + b_\mathrm{nr}$ and the quantum yield $\Phi = b_\mathrm{r}/b$, and a classical expression for the radiative damping constant:

$$b_\mathrm{r} = \frac{2e^2\omega^3 n}{3c^3 m}, \tag{2.220}$$

with the refractive index $n$ of the medium surrounding the dipole, and the speed of light $c$. Then, the expression for the lifetime in the presence of the interface could be rewritten as:

$$\frac{\tau'}{\tau} = \left[1 + \Phi\frac{3c^3}{2\mu_0\omega^3 n}\ \mathfrak{Im}(E_0)\right]^{-1}. \tag{2.221}$$

Since $E_0$ is proportional to $\mu_0$, the dipole moment cancels. The speed of light, the wavelength, all refractive indices and even the distance between dipole and mirror were known, such that the quantum yield $\Phi$ was the only free fit parameter. Reasonably good fits could be obtained for all distances $r$ in [81], which was seen as a confirmation of this model.

A completely different but still semi-classical approach was suggested by Drexhage to explain his experimental results [37]. It is based on the idea mentioned at the beginning of this section: The total amount of radiation emitted per time by a classical oscillating dipole $S_\mathrm{tot}$ changes in the vicinity of interfaces or nanostructures. Since the energy emitted per deexcitation event is determined solely by the wavelength (energy $= hc/\lambda$ with Planck's constant $h$) and thus fixed, "more energy emisison per time" can only be realised through "more deexcitation events per time" and thus a shorter lifetime. In his calculations, Drexhage assumed that no energy can enter the mirror, and thus only integrated the energy flux over the top halfspace. Furthermore, he stated that a doubled total emitted energy per time leads to a halved lifetime, both with respect to the values in free space (i.e. in the same medium but without presence of the interface), which can be easily determined. With this approach, Drexhage was able to explain the lifetime variation at $nd/\lambda \gtrsim 0.6$. For shorter distances, his results predicted an increased lifetime, while the measurement showed that the lifetime strongly decreased. Drexhage correctly assumed that this was due to the energy transfer to the metal mirror, which he had not taken into account. This effect was included in the later calculations by Chance, Prock and Silbey [82], who integrated the energy flux over two planes parallel to the mirror, one above and one below the dipole, as we did in section 2.4.5. In [82] it is said that "$F_\uparrow$, divided by the energy of the dipole (which we take as $|\mu|^2$ for convenience) is simply the rate constant $b_\uparrow$ associated with energy loss through the upper plane", where $F_\uparrow$ is the part of $S_\mathrm{tot}$ that is emitted into the upper halfspace. In other words, if an oscillating dipole has some initial energy $\varepsilon_0$ and then emits energy with a rate $S_\mathrm{tot}$,

the energy is "used up" after the time $\tau_r = \varepsilon_0 / S_{\text{tot}}$. According to this assumption, the radiative deexcitation rate should simply be given by $S_{\text{tot}} / |\mu|^2$. However, it is not clear why the initial energy of the dipole should be $\varepsilon = |\mu|^2$. After all, the energy difference between the excited state and the ground state is $hc/\lambda$. Here, a connection between the classical mechanical model with an oscillating charge that results in the dipole moment $\mu$ and the quantum-mechanical system with the transition matrix element $\mu$ would have to be found. However, later on the authors renormalise the rate to that of a dipole emitter in free space, which is the same as Drexhage's approach and which seems much more robust. Therefore, our own calculations in the subsequent chapters use Chance, Prock and Silbey's idea to determine $S_{\text{tot}}$ with Drexhage's suggestion to compare with the free space values $\tau_0$ and $S_0$:

$$\boxed{\frac{\tau(d)}{\tau_0} = \frac{S_0}{S_{\text{tot}}(d)}.}$$

(2.222)

The calculations can be sped up significantly by exploiting a peculiarity of the total energy flux. As will be shown in section 3.4, the total flux through an infinite plane[27] generated by a dipole emitter with dipole moment $\boldsymbol{p}$ is given by an integral over the angles $\theta'$ and $\phi'$, the latter from zero to $2\pi$. The integrand is the sum of two terms, one proportional to $|\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_p|^2$ and the other proportional to $|\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_s|^2$. These two terms have the property that

$$\int_0^{2\pi} \mathrm{d}\phi' \, |\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_p|^2 = \pi \left[ \sin^2\theta \cos^2\theta' + \cos^2\theta \sin^2\theta' \right]$$

$$\text{and} \quad \int_0^{2\pi} \mathrm{d}\phi' \, |\boldsymbol{p} \cdot \hat{\boldsymbol{e}}_s|^2 = \pi \sin^2\theta$$

(2.223)

for a dipole moment given by

$$\boldsymbol{p} = p \begin{pmatrix} \sin\theta \cos\phi \\ \sin\theta \sin\phi \\ \cos\phi \end{pmatrix}.$$

(2.224)

Since $\boldsymbol{p}$ does not appear anywhere else in the expression for the total flux, we can draw two conclusions from this result. Firstly, the flux does not depend on the angle $\phi$ between $\boldsymbol{p}$ and the $x$-axis, therefore we define the total flux generated by any dipole oriented parallel to the interface as $S_{\text{tot},\parallel}$. Secondly, when the above terms are evaluated for a dipole parallel ($\theta = \pi/2$) or perpendicular ($\theta = 0$) to the plane, either $\sin^2\theta$ or $\cos^2\theta$ equals one, while the other is zero. Hence, the flux of an arbitrarily oriented dipole emitter can be expressed as a linear combination of the flux $S_{\text{tot},\perp}$ of a vertical dipole with $\theta = 0$ and the flux $S_{\text{tot},\parallel}$:

$$\boxed{S_{\text{tot},\theta}(d) = \sin^2\theta \cdot S_{\text{tot},\parallel}(d) + \cos^2\theta \cdot S_{\text{tot},\perp}(d),}$$

(2.225)

---

[27]Which is assumed to be parallel to the $x$-$y$-plane.

where we now again included the dependence on the distance $d$ between the dipole emitter and the planar interface. Thus, only $S_{tot,\parallel}(d)$ and $S_{tot,\perp}(d)$ have to be determined explicitly, then all other fluxes can be quickly derived from them.

Following Drexhage's experimental proof of Purcell's prediction [36, 37] and the publication of various theoretical explanations ([80–84], some of those presented above), the Purcell effect was studied in a number of experimental situations in the 1980s and 1990s. These included fluorophores close to silver films of varying thickness [85], between two closely spaced mirrors (i.e. inside a planar cavity) [86, 87], in multilayered thin-film samples [88], close to silver island films consisting of many individual metal spheroids [89], or in dielectric nanospheres [90, 91]. Parallel to the advances in nanotechnology in the 2000s, the Purcell effect was studied experimentally for emitters in a variety of previously inaccessible environments, such as photonic crystals [92] and electrospun polymer nanofibers [93]. These works were accompanied by a large number of theoretical treatments, such as predictions of the radiative rate inside nanoscopic dielectric particles of varying size and shape [94] and inside multilayered nanoscopic spheres [95]. In 2010, the Purcell effect was met with renewed interest, when it was applied to the rapidly growing field of super-resolution microscopy in the form of metal-induced energy transfer microscopy. This technique will be treated in detail in the next section.
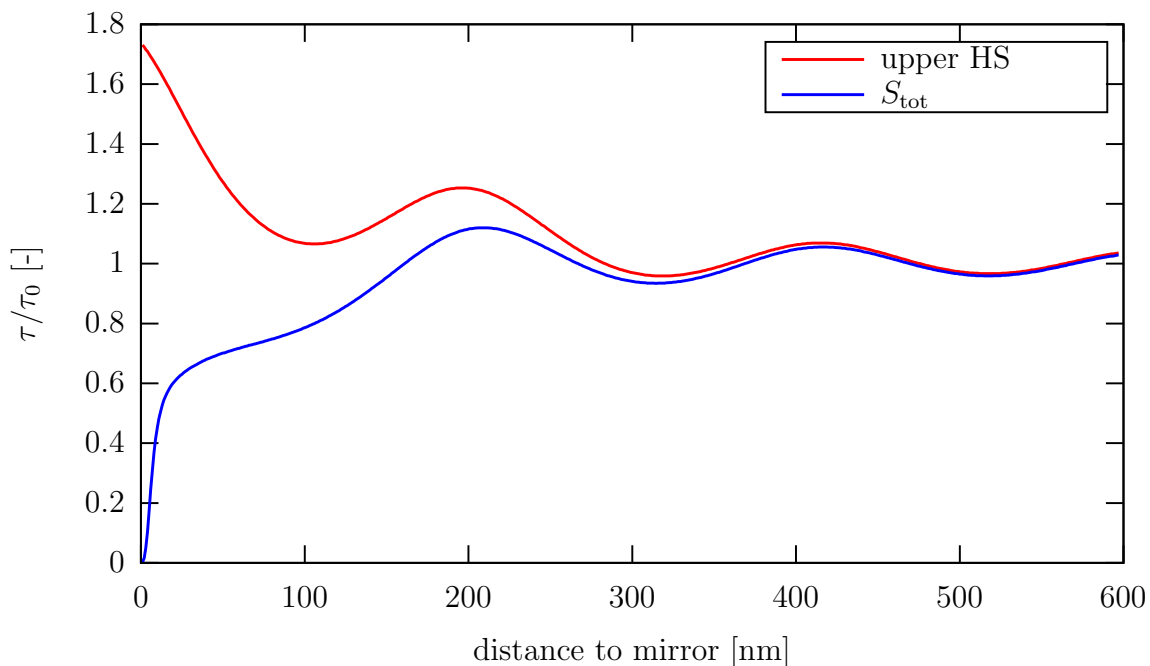


**Figure 2.27:** Relative fluorescence lifetime of $Eu^{3+}$ ions in an LB-film close to a silver mirror (more accurately: a silver halfspace). Same parameters as used by Drexhage in [37]: $\lambda = 612$ nm, $\Phi = 0.7$, $n_{Ag} = 0.06 + 4.11i$, $n_{LB} = 1.54$, random orientation of molecules. The red curve shows the result if only the flux into the uppper halfspace is taken into account, the blue curve shows the result including energy transfer to the metal.

## 2.5.4 Metal-induced energy transfer

One of the central findings of this thesis is that the fluorescence lifetime $\tau$ of a dipole emitter changes dramatically, but in a well-defined manner, close to an interface. We now show an example to demonstrate the dependence of this lifetime change on the distance $z$ between the emitter and the interface. Figure 2.27 depicts $\tau(z)$ for the same system as used by Drexhage in [37], europium ions in an LB-film on top of a thick silver mirror. As already mentioned above, Drexhage only integrated the energy flux over the upper halfspace, which leads to the red curve in the figure. When $S_{\text{tot}}$ is used instead, the blue curve results. It becomes clear that, at large distances from the mirror, the emission into the upper halfspace is the dominating effect. Then, the lifetime oscillates in $z$ with a period of $0.5\lambda/n$ and with decreasing amplitude. This can be explained by either of the two classical models, assuming for simplicity a perfect mirror with reflection coefficient $r = \exp(i\pi)$ and transmission coefficient $t = 0$:

As Drexhage already pointed out [37], if the distance between the emitter and the mirror is given by $z = \lambda/n \cdot (2j+1)/4, j \in \mathbb{N}$, the reflected light interferes constructively with the light emitted by the dipole. Since the intensity is proportional to the square of the amplitude, the energy radiated into the upper halfspace per time is quadrupled. Due to the perfect mirror, no energy is emitted into the lower halfspace. Together, this leads to a doubling of $S_{\text{tot}}$. Contrarily, if the path difference leads to destructive interference, no energy is emitted at all and the lifetime would become infinite[28]. However, this is only true for all emission angles if the molecule is located in the center of a spherical mirror whose top half has been removed, such that the path difference is independent of the emission angle. For a real planar mirror, both the path difference and the reflection coefficient do depend on the emission angle. Thus, while some waves might interfere constructively, others interfere destructively. This is why the amplitude of the modulation is relatively small. Since the field of a dipole decreases with distance, the effect of adding the reflected field also gets weaker with larger $z$, leading to an asymptotic convergence $\tau \to \tau_0$.

Alternatively, the model of an electrical charge oscillating in a harmonic potential can also be used to explain the far-field behaviour, as done by Kuhn in [80]. From this perspective, the reflected electric field can be interpreted as the field of a virtual "mirror image" oscillator identical to the real emitter and oscillating out of phase (due to the phase of the reflection coefficient $r$). The field of this mirror image oscillator reaches the real oscillator with a time delay $\Delta t = 2zn/c$ due to the finite velocity of propagation of the fields. The electric field exerts a force on the charge, therefore it can accelerate or decelerate the charge, depending on the magnitude of the time delay. Since in this model the lifetime is interpreted as the time until the oscillator stops, a deceleration of the charge results in a decreased lifetime, while an acceleration of the charge increases the lifetime.

However, the more interesting case for us is $z \lesssim \lambda/2$. In this regime, the lifetime change is much more pronounced than at larger distances. As the comparison between the red and the blue curve in figure 2.27 shows, the behaviour at small distances is mainly governed by energy transfer to the lower halfspace. We call this *Metal-Induced Energy Transfer* (MIET).

---

[28]Assuming in both cases that absolutely no other relaxation pathways exist.

As already explained at the end of section 2.4.5, all evanescent waves generated by the dipole emitter that enter the silver halfspace will get absorbed. Of course, with increasing distance between the emitter and the metal surface, the amplitude of the evanescent waves when they reach the interface decreases. Thus, also the amount of energy transported to the metal decreases. At an intermediate distance range, the absorption by the metal is dominated by surface plasmon polaritons, which we introduced in section 2.4.3. This is illustrated in figure 2.28(a), which shows the relative contributions of SPPs, general absorption by the metal and radiation into the fatty acid halfspace to the total energy flux[29].

A more detailed understanding of this behaviour can be gained from the logarithmic plot of the different fluxes in figure 2.28(b). We can see that the flux due to SPP decreases monoexponentially due to the sharp resonance at a single $q$-value explained in section 2.4.3. The remaining flux into the metal is caused by all other evanescent waves and thus shows a multiexponential decrease, which is very steep at short distances. In contrast, at large distances from the metal, evanescent waves do not play an important role anymore. For propagating waves, the reflectivity of the silver/fatty acid interface is larger than 96% for all angles of incidence. Thus, almost all energy that the dipole emits is either directly radiated or reflected into the upper halfspace. To sum up, if an emitter is located within about 150 nm from a metal surface (where the exact limit depends on the dipole's orientation), there is a strong energy transfer from the emitter to the metal, which can be detected as a decrease in fluorescence lifetime.

This phenomenon can now be exploited to overcome a severe limitation of conventional confocal microscopy, namely the low axial resolution, by "inverting" Drexhage's experiment: Instead of placing dipole emitters at known positions and measuring their lifetimes, one can infer the unknown axial positions of dipole emitters from their fluorescence lifetimes. This technique is called *MIET microscopy* after its main physical mechanism. As the lifetime-*versus*-height curves (from now on simply called *MIET curves*) for different sample geometries in figure 2.29 show, there is a window of opportunity in the first 200 nm where $\tau(z)$ inreases monotonically. This region is interesting for studying single molecules, but also the basal (lower) membrane of cells, or cellular structures that are close to the substrate. A number of examples are discussed in chapter 4. The main advantage of MIET is its excellent $z$-resolution on the order of a few nanometers, two orders of magnitude smaller than in confocal microscopy [96].

---

[29]As described by equation (2.156), surface plasmon polaritons obey the dispersion relation $q_{\mathrm{SPP}} = k_v n_1 n_2 / \sqrt{n_1^2 + n_2^2}$. However, evanescent waves with $q$ close to this value can also couple to SPPs, albeit with smaller efficiency. This was reflected in figure 2.15 as a sharp but not infinitely thin peak of $S_z(q)$ – the energy flux along the $z$ direction caused by waves with horizontal wave vector component $q$ – at $q_{\mathrm{SPP}}$. Mathematically, this resonance is due to corresponding peaks in the amplitudes of both the reflection and transmission coefficients of $p$-waves, $|r_p(q)|$ and $|t_p(q)|$, at these $q$-values. Unlike the concrete shape of $S_z(q)$, these two functions are independent of the fluorophore and its $z$-position. Therefore, they were used to estimate the $q$-range where evanescent waves couple to SPPs: We determined the mean $\langle |t_p(q)| \rangle_q$ over all $q > k_1$, i.e. over all evanescent waves. All $q$-values with $|t_p(q)| > 1.5 \cdot \langle |t_p(q)| \rangle_q$ were interpreted as corresponding to an excitation of SPPs. The fluxes presented in figure 2.28 were then obtained by integrating $S_z(q)$ over the corresponding $q$-ranges.
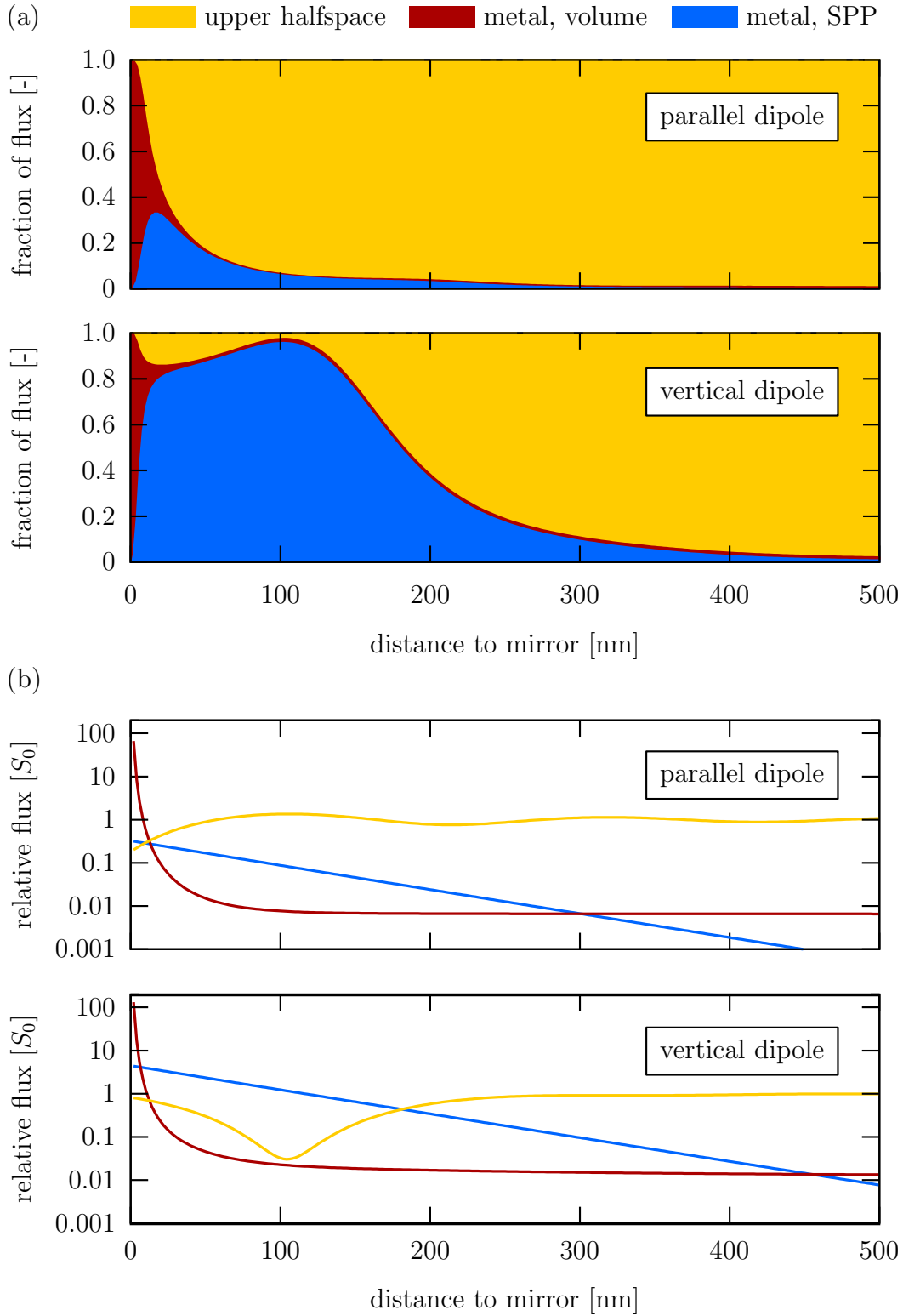
**Figure 2.28:** Comparison of the different radiative energy channels relevant for Eu$^{3+}$ ions close to a fatty acid/silver interface. (a) Relative contributions of SPPs, general absorption by the metal and radiation into the fatty acid halfspace to the total energy flux. (b) Fluxes normalized by the total emission of a dipole in free space $S_0 = cnk_v^4p^2/3$.
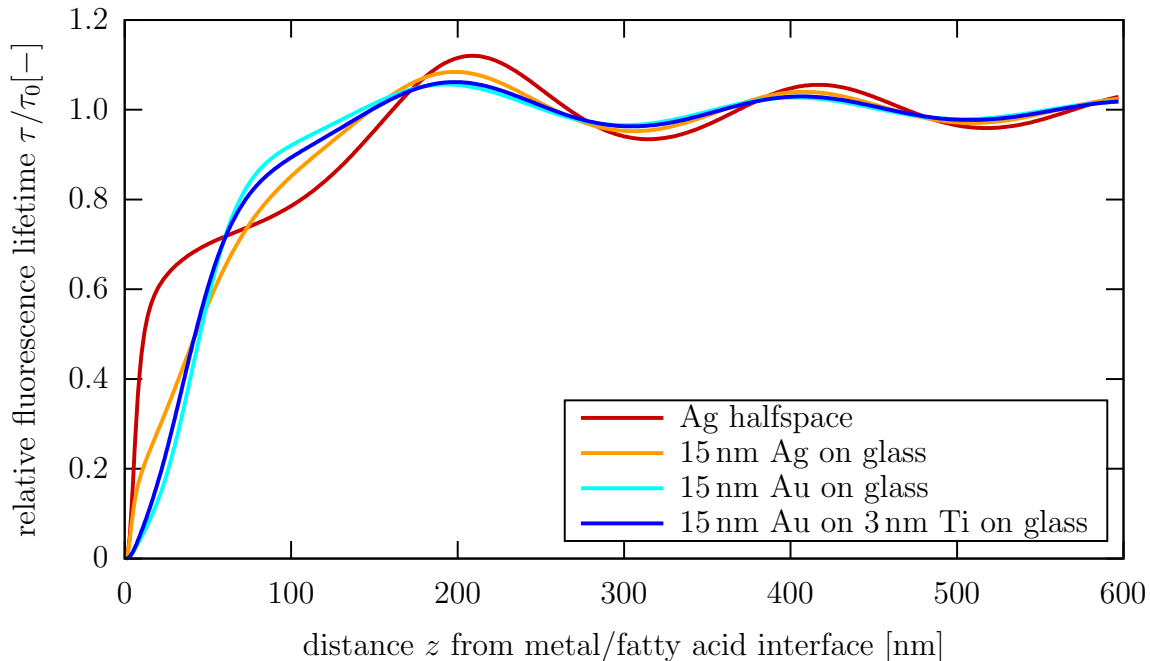
**Figure 2.29:** MIET lifetime-*versus*-height curves for different sample geometries. $Eu^{3+}$ ions are embedded with random orientation in an LB-film ($n = 1.54$) at varying distance from a metal/LB-film interface. The metal is either a complete silver halfspace (red curve), a thin silver film on a glass halfspace (orange curve), a thin gold film on a glass halfspace (light blue curve) or a thin gold film on a titanium layer on a glass halfspace (dark blue curve). The ions fluoresce at $\lambda = 612\,\text{nm}$ with a quantum yield of $\Phi = 0.7$.

It is mainly determined by the quality of the lifetime measurements, since a lifetime uncertainty $\Delta\tau$ can be directly converted to a height uncertainty $\Delta z$ via

$$\Delta z = \frac{\partial z(\tau)}{\partial \tau}\,\Delta\tau, \tag{2.226}$$

assuming that the MIET curve is locally linear. The implications will now be demonstrated for a concrete example, the first proof-of-principle study of MIET, which was conducted in 2010 [38]. There, microtubules were labeled with the fluorescent dye Alexa Fluor 488, and then attached to a gold-coated surface via different spacer molecules, see figure 2.30 (a)-(c). By measuring the fluorescence lifetime of the dye and comparing with the MIET calibration curve, the influence of the spacer molecules on the axial distance between gold and microtubules could be examined. For each experiment, only one type of spacer molecule was used, therefore the microtubules were aligned parallel to the surface. Thus, the lifetimes from all pixels could be averaged, leading to a standard error of the mean that was much smaller than the standard deviation of single pixels (0.04 ns-0.08 ns as compared to 0.15 ns, see the top panel of figure 2.30(d)). The diameter of the microtubules was assumed to be 25 nm, as determined by electron microscopy [38], and they were assumed to be evenly covered by the dye. Using this geometry, the average observable lifetime as a function of the distance between the top of the gold surface and the bottom of the microtubules was determined, see the bottom
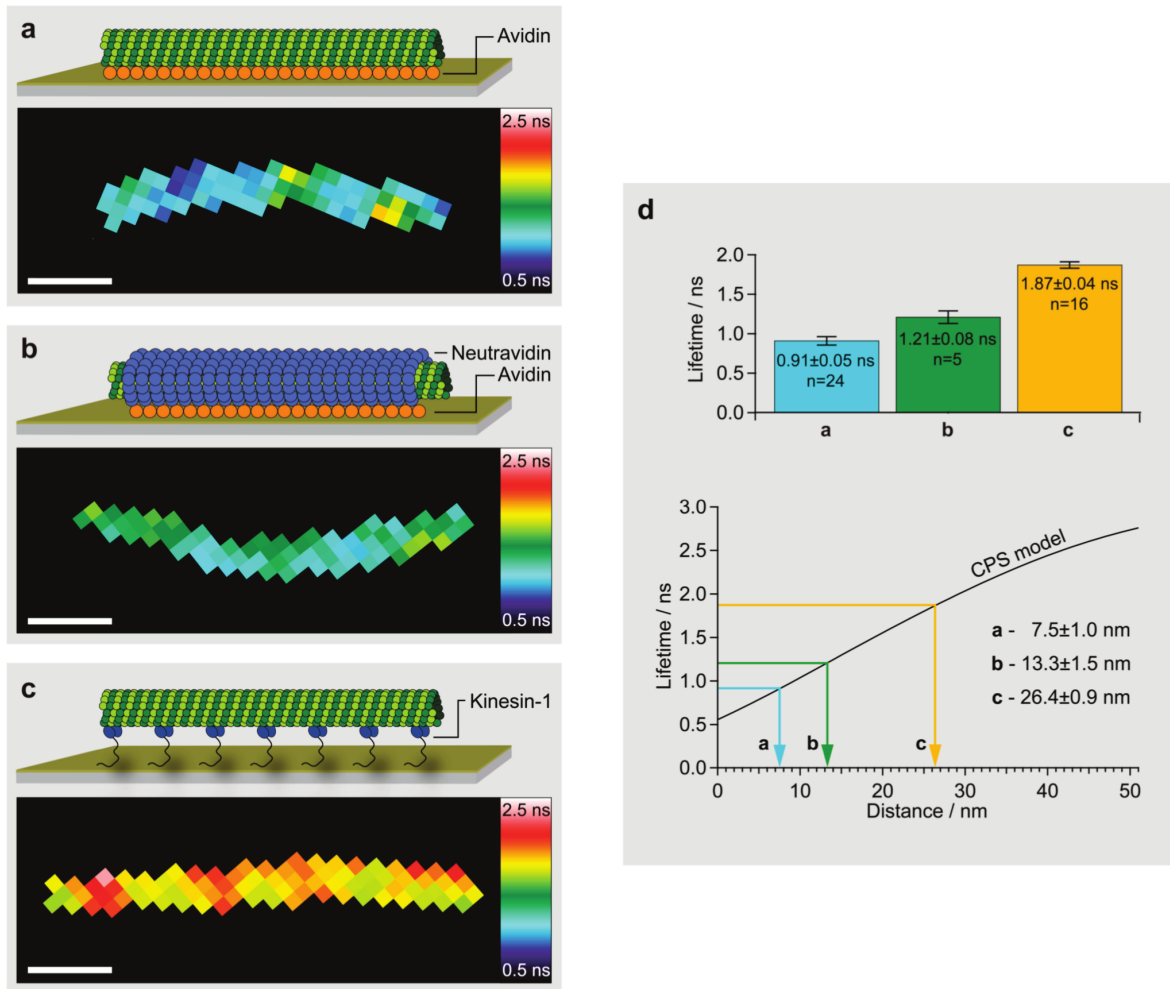
**Figure 2.30:** First proof-of-principle experiment of MIET microscopy, from [38]. Alexa Fluor 488-labeled microtubules were attached to a gold-coated cover slip using avidin (a), neutravidin and avidin (b), or kinesin-1 (c). Shown are experimental schematics and examplary lifetime images, scale bars correspond to 5 μm. Panel (d) depicts the average lifetimes of $n$ regions from different samples, which were converted to absolute distances between microtubules and gold surface using the MIET calibration curve (here called Chance-Prock-Silbey model).

panel of figure 2.30(d). This is just a variation of the MIET curve discussed previously. With the help of this curve, the lifetime values were converted to axial elevations of the microtubules above the gold surface. The obtained results, given in figure 2.30(d), were in very good agreement with the geometrical sizes of the spacer molecules avidin, neutravidin and kinesin-1 as obtained by other methods [38]. Due to the linearity of the MIET curve in the observed $z$-region (see figure 2.30(d)), the authors used the lifetime uncertainty $\Delta\tau = 0.15$ ns to estimate a position-independent height uncertainty $\Delta z \approx 3$ nm. This was significantly more accurate than most other axial localization techniques available at the time, for example 3D STORM with $\Delta z \approx 20$ nm $-$ 25 nm[30]

---

[30]Zhuang *et al.* state that their axial resolution is in the range 50 nm-60 nm [31], however, this is the full width at half maximum (FWHM) of their $z$-position distributions. In this work, the standard

[31] or TIRF with $\Delta z \approx 15\,\mathrm{nm} - 50\,\mathrm{nm}$ [97]. With *fluorescence interference contrast* (FLIC) microscopy, another technique with similar localization accuracy on the order of $1\,\mathrm{nm}$ had been available since the late 1990s [97, 98]. However, FLIC relies on accurate intensity measurements and is therefore susceptible to intensity fluctuations due to inhomogeneous labeling or photobleaching – effects that do not affect lifetime measurements. Thus, MIET microscopy proved to be an attractive new tool for studies of cells or single proteins.

In order to reliably achieve high $z$-resolution, some experimental parameters have been optimized. As shown in equation (2.226), a smaller lifetime uncertainty $\Delta \tau$ leads to a smaller $z$-uncertainty $\Delta z$. The former necessitates a high signal-to-noise ratio of the lifetime measurement. As explained in section 2.5.2, the SNR can be increased by reducing both residual excitation light as well as fluorescence from out-of-focus planes, which is achieved by using an inverted confocal microscope. However, as we have seen several times now, light only penetrates a short distance into metal before being absorbed, thus we cannot excite or detect fluorescence through a silver halfspace. If we tried to excite and detect from above, the thickness of the sample would result in a distorted focus and a lot of scattered light for samples that are more complex than an LB film, for example fluorescently labeled structures in a cell. Fortunately, the results of modeling the situation for different sample parameters offer a good alternative: As figure 2.29 shows, for $Eu^{3+}$ ions close to a thin silver film on top of a glass halfspace, the lifetime is still zero at $z = 0$, and increases monotonically up to roughly the same distance as for the silver halfspace, $z \approx 200\,\mathrm{nm}$. Such a thin metallic layer can simply be evaporated on top of a smooth glass cover slip, and enough light can pass through to allow accurate lifetime measurements. Another practical aspect that has to be taken into account is the stability of the metal film: Silver oxidizes quickly, which changes the refractive index and thus the MIET curve. It can also increase the surface roughness, making the assumption of a plane interface inaccurate and leading to unpredicted lifetime changes. A metal that is much more stable against oxidization is gold. Figure 2.29 also shows the MIET curve for a $15\,\mathrm{nm}$ thick gold film, which shares the main characteristics of the curve of the silver film. However, gold does not adhere very well to glass [99]. The risk of dissolution can be greatly lowered by including a $3\,\mathrm{nm}$ thick layer of titanium between the glass and the gold [99]. As figure 2.29 shows, this only has a small impact on the MIET curve. Because of these considerations, an optimal sample for MIET microscopy consists of a layered glass/titanium/gold structure. This is the standard sample used in the applications described in chapter 4.

Before introducing the various systems studied with MIET in the course of this work, we now present the concrete numerical implementation of some of the aspects described in this chapter.

---

deviation (STD) of such distributions is interpreted as the resolution. Since for a Gaussian distribution $\mathrm{FWHM} = \sqrt{8 \ln 2}\, \mathrm{STD} \approx 2.4\, \mathrm{STD}$, we arrive at an axial resolution of $20\,\mathrm{nm}$-$25\,\mathrm{nm}$.

# 3 Numerical Implementation

The analytical theory presented in the previous chapter allows to formulate some general principles, such as the orthogonality of electric and magnetic fields in the absence of free charges and currents, or the importance of evanescent waves close to metal structures. However, a large part of the theory cannot be used directly to interpret experimental data. That is because the integrals and infinite sums that are part of many of our equations (e.g. the Weyl representation of the electrical field of a dipole or the conversion formulas for VSH and plane waves) generally cannot be solved analytically. Instead, the analytical equations are approximated by numerial algorithms, which are implemented on a computer system. To this end, integrals are transformed into sums, and cutoffs have to be chosen whenever an integration or summation to infinity is required. The errors introduced by these approximations usually decrease when more integration points and larger cutoff values are used. This convergence behaviour is explored in section 3.1 for some of the formulas introduced in chapter 2.

Another aspect has to be taken into account when performing calculations on a computer: Numbers can only be stored with a finite accuracy and within a certain range. In all of our calculations, we use floating-point arithmetic. This means that all numbers are approximated by a representation of the form

$$\text{significand} \times \text{base}^{\text{exponent}}, \tag{3.1}$$

where the significand is a real number and base and exponent are integers. While in principle any integer equal to or larger than two can be used as a base, the most common bases are two and ten. These representations are called *binary* and *decimal*, respectively. The exact number of bits used to store significand, base and exponent has been defined by the IEEE Standard for Floating-Point Arithmetic (IEEE 754) [100], which is implemented by all common programming languages such as C, C++, Python or Matlab. The two most common floating-point numbers are called *single precision* and *double precision*, they are defined as:

| name | common name | base | sign | significand | exponent |
|------|-------------|------|------|-------------|----------|
| binary32 | single precision | 2 | 1 bit | 24 bits* | 8 bits |
| binary64 | double precision | 2 | 1 bit | 53 bits* | 11 bits |

*Inluding the implicit bit, see footnote.

Theoretically, some numbers have several possible representations, for example $34 = 3.4 \cdot 10^1 = 0.34 \cdot 10^2$ in the decimal system. This ambiguity has been overcome in the binary system by defining that the significand always has the form 1.f, where f has 52 digits[1]. Translated to decimal numbers, the standard defines that the exponents range

---

[1] Since the leading one is always present in the binary representation, it does not have to be stored explicitly. This is called the *leading* or *implicit bit convention*, and it is the reason why the significand is listed with 24 (53) bits even though only 23 (52) bits are really used.

from -126 to +127 for binary32 and from -1022 to +1023 for binary64, and that the number of decimal digits of the significand are roughly 7 for single and 16 for double. Furthermore, it implies that there is a largest representable number which has 1 in every bit of the significand and the largest possible value for the exponent. This is called the *overflow level* (OFL), for double precision it is given by[2]:

$$\mathrm{OFL}_{\mathrm{double}} = (2 - 2^{-52}) \cdot 2^{1013} \approx 1.8 \cdot 10^{308} \tag{3.2}$$

As soon as the correct result of a calculation is larger than the OFL, depending on the concrete programming language, the generated result will be either completely wrong or "infinity". Wherever possible, handling such large numbers should be avoided, for example by renormalizing the quantities used in the equation. In some situations, only an intermediate result is larger than the OFL, while the final result would be smaller. Then, changing the order of operations can sometimes circumvent the overflow, which we will exploit in section 3.3. In section 3.2, fractions with very large terms in both numerator and denominator appear. There, we avoid the overflow problem by identifying relevant contributions and then cancelling "by hand".

The finite precision of floating point numbers leads to rounding errors, which are generally small for a single operation (recall that the decimal precision is roughly 16 digits for binary64). However, even a single operation can lead to a large relative error when combining numbers of very different magnitude. An example is the so-called catastrophic cancellation:

$$\left[(1)_{\mathrm{binary64}} + (10^{-17})_{\mathrm{binary64}}\right] - (1)_{\mathrm{binary64}} = (0)_{\mathrm{binary64}}. \tag{3.3}$$

But even less extreme rounding errors can accumulate if many operations have to be performed on the same data set. Therefore, when comparing different realizations of the same quantity (e.g. an analytical result and a PW decomposition), we consider them "equal within the numerical accuracy" if the relative error is smaller than or on the order of $10^{-14}$. The accumulation of errors can have even worse effects when using recurrence relations to determine functions of high orders. Therefore, for the determination of the associated Legendre polynomials, we chose relations that are described as *stable recurrences* in the literature (see section 3.3). Finally, the number of numerical operations can be reduced if as many steps as possible are evaulated analytically. To this end, choosing a suitable coordinate system, basis function or integration surface is vital. This choice is demonstrated for the example of the total energy flux of a dipole emitter in a stratified system in section 3.4.

Besides the theoretical modeling of dipole emitters in various situations, a substantial part of this work consisted of the evaluation of experimental data. There, another task was encountered which can only be performed on a computer system: the fitting of theoretical models to measured fluorescence lifetime data. A description of the fit procedures used in this work in section 3.5 concludes the chapter.

---

[2] The binary numer $(0.1)_2$ is equal to $2^{-1}$. Correspondingly, $(0.\underbrace{00\ldots00}_{51 \text{ zeroes}}1)_2 = 2^{-52}$. Thus, while $(1.1)_2 = 1 + 2^{-1} = 2 - 2^{-1}$, we also have $(1.\underbrace{11\ldots11}_{52 \text{ ones}})_2 = 1 + 2^{-1} + 2^{-2} + \ldots + 2^{-51} + 2^{-52} = 2 - 2^{-52}$. This is the significand when all entries are one.

## 3.1 Convergence behaviour

We found in section 2.1.4 that the conversion of a plane wave to vector spherical harmonics requires an infinite sum, while the inverse transformation is based on an integral on a bounded (for VSH with spherical Bessel function) or unbounded (for VSH with spherical Hankel function) interval. Where analytical expressions exist for such an integral, using them is never less and often more precise than solving the integral numerically[3]. Thus, we use the analytical $\varphi'$ integration of equations (2.53)-(2.55) when changing from VSH to plane waves, leading to the already known relations

$$\boldsymbol{M}^j_{\ell m}(\rho, \varphi, z) = \frac{1}{4i^{\ell-1}} \int_0^\pi \mathrm{d}\theta' \mathrm{e}^{ik' \cos\theta' z} \left[\pi_{\ell m}(\theta')\boldsymbol{p}_m(\theta', \rho, \varphi) + i\tau_{\ell m}(\theta')\boldsymbol{s}_m(\theta', \rho, \varphi)\right], (2.56)$$

$$\boldsymbol{N}^j_{\ell m}(\rho, \varphi, z) = \frac{1}{4i^{\ell-1}} \int_0^\pi \mathrm{d}\theta' \mathrm{e}^{ik' \cos\theta' z} \left[\tau_{\ell m}(\theta')\boldsymbol{p}_m(\theta', \rho, \varphi) + i\pi_{\ell m}(\theta')\boldsymbol{s}_m(\theta', \rho, \varphi)\right], (2.57)$$

$$\boldsymbol{M}^h_{\ell m}(\rho, \varphi, z) = \frac{1}{2i^{\ell-1}} \int_\mathcal{C} \mathrm{d}\theta' \mathrm{e}^{ik' \cos\theta' z} \left[\pi_{\ell m}(\theta')\boldsymbol{p}_m(\theta', \rho, \varphi) + i\tau_{\ell m}(\theta')\boldsymbol{s}_m(\theta', \rho, \varphi)\right], (2.60)$$

$$\boldsymbol{N}^h_{\ell m}(\rho, \varphi, z) = \frac{1}{2i^{\ell-1}} \int_\mathcal{C} \mathrm{d}\theta' \mathrm{e}^{ik' \cos\theta' z} \left[\tau_{\ell m}(\theta')\boldsymbol{p}_m(\theta', \rho, \varphi) + i\pi_{\ell m}(\theta')\boldsymbol{s}_m(\theta', \rho, \varphi)\right], (2.61)$$

$$\text{with the path} \quad \mathcal{C} = \begin{cases} 0 \to \pi/2 \to \pi/2 - i\infty & \text{for } z > 0 \\ \pi/2 + i\infty \to \pi/2 \to \pi & \text{for } z < 0 \end{cases}.$$

When such an integral has to be evaluated numerically, it has to be discretized into a finite sum, where both the density of integration points and the cutoff (for an unbounded interval) have to be chosen carefully. A large number of numerical integration methods exist [101], which can generally be expressed in the form

$$F(a, b) := \int_a^b f(x) \, \mathrm{d}x \approx (b - a) \cdot \sum_{i=1}^N w_i \cdot f(x_i) \tag{3.4}$$

and which differ in the choice of weights $w_i$ and integration points $x_i$. The simplest technique is called the *midpoint* or *rectangle rule*, it approximates the integrand $f(x)$ as piecewise constant between neighbouring, equally spaced integration points with equal weights $w_i = 1/N$:

$$F_{\mathrm{MP}}(a, b) = \frac{b - a}{N} \cdot \sum_{i=1}^N f(x_i), \qquad x_i = a + \frac{b - a}{N} \cdot \left(i - \frac{1}{2}\right). \tag{3.5}$$

If the interpolating function between two neighbouring integration points is a parabola instead, one finds *Simpson's rule*, which has a smaller error than the midpoint rule with the same integration points:

$$F_{\mathrm{Simps}}(a, b) = \frac{b - a}{3N} \cdot \sum_{i=1}^{N/2} \left[f(x_{2j-2} + 4f(x_{2j-1}) + f(2x_j)\right]. \tag{3.6}$$

---

[3] Provided that the functions making up the analytical expression are implemented with high accuracy in the employed programs, which we assume here.

In this work, the focus was more on the physical phenomena and their mathematical description than on the optimization of the numerical implementation. Therefore, the midpoint rule was used everywhere in this work, with one exception: The calculation of the total energy flux of a dipole, $S_{\text{tot}}$, was performed with Simpson's rule since we optimized this calculation for the evaluation of our MIET experiments and also published the code in the form of a graphical user interface (see section 4.1.1). Nevertheless, the general convergence behaviour when increasing $N$ or changing the cutoff of unbounded integrals is similar for both methods.

We illustrate this behaviour for an example that was already shown in section 2.1.4, $\boldsymbol{M}_{4,1}^h$. Figure 3.1 investigates the effect of the number $N$ of integration points and the upper limit $T$ of $\mathfrak{Im}(\theta')$. It becomes clear that increasing $T$ generally improves the results (top row). However, if $N$ is kept constant while $T$ is increased, this leads to a coarser spacing of the integration points, ultimately decreasing the quality of the result (compare the top right figure with $N = 200$, $T = 10$ and the image in the center of the figure with $N = 200$, $T = 20$). On the other hand, increasing $N$ also leads to a better result (bottom row). Thus, when a good approximation of the field is desired, both $T$ and $N$ have to be chosen adequately. Since both computation time and required memory scale (only) linearly with $T$ and with $N$, using high values is generally possible. However, one must take into account the fact that rounding errors can lead to wrong results if the spacing is too fine. Consider the discretization of the integral

$$\int_a^b c \, \mathrm{d}x \to \sum_{n=1}^{N} c \cdot \frac{b-a}{N}. \tag{3.7}$$

In this case, since the integrand is constant, $N = 1$ would already give the correct result. For very large $N$, the integration measure $(b-a)/N$ becomes very small, and thus many small terms have to be added to yield the final result. This can lead to the accumulation of rounding errors. In an integral such as those in the above PW decompositions, the integrand is generally not a constant. Thus, a length scale has to be found on which the approximation of a piecewise-constant integrand is accurate enough and the accumulation of rounding errors not yet too bad. In practice, we used values on the order $10^2 - 10^3$ for $N$ and $10 - 20$ for $T$.

The inverse problem of transforming plane waves into VSH is treated next. Here, the formula presented in the theory section was:

$$(E_p \hat{\boldsymbol{e}}_p + E_s \hat{\boldsymbol{e}}_s) \, \mathrm{e}^{i\boldsymbol{k}' \cdot \boldsymbol{r}} = 4\pi \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \frac{i^{\ell-1}}{\ell(\ell+1)} \mathrm{e}^{-im\phi'} \quad .$$

$$\left( [E_p \pi_{\ell m}(\theta') - iE_s \tau_{\ell m}(\theta')] \, \boldsymbol{M}_{\ell m}^j(\boldsymbol{r}) + [E_p \tau_{\ell m}(\theta') - iE_s \pi_{\ell m}(\theta')] \, \boldsymbol{N}_{\ell m}^j(\boldsymbol{r}) \right). \tag{2.62}$$

Again, since it is impossible to sum over infinitely many terms numerically, a cutoff has to be chosen. In practice, we restrict $\ell$ to values between 1 and $L$. Figure 3.2 shows some exemplary results of a Matlab implementation of this expression for $L = 10$. An $s$-wave propagating in the direction $\boldsymbol{k}' = (1, 0, 1)/\sqrt{2}$ is depicted in subfigure 3.2a, an evanescent $p$-wave in subfigure 3.2b. The deviation between analytical result and VSH
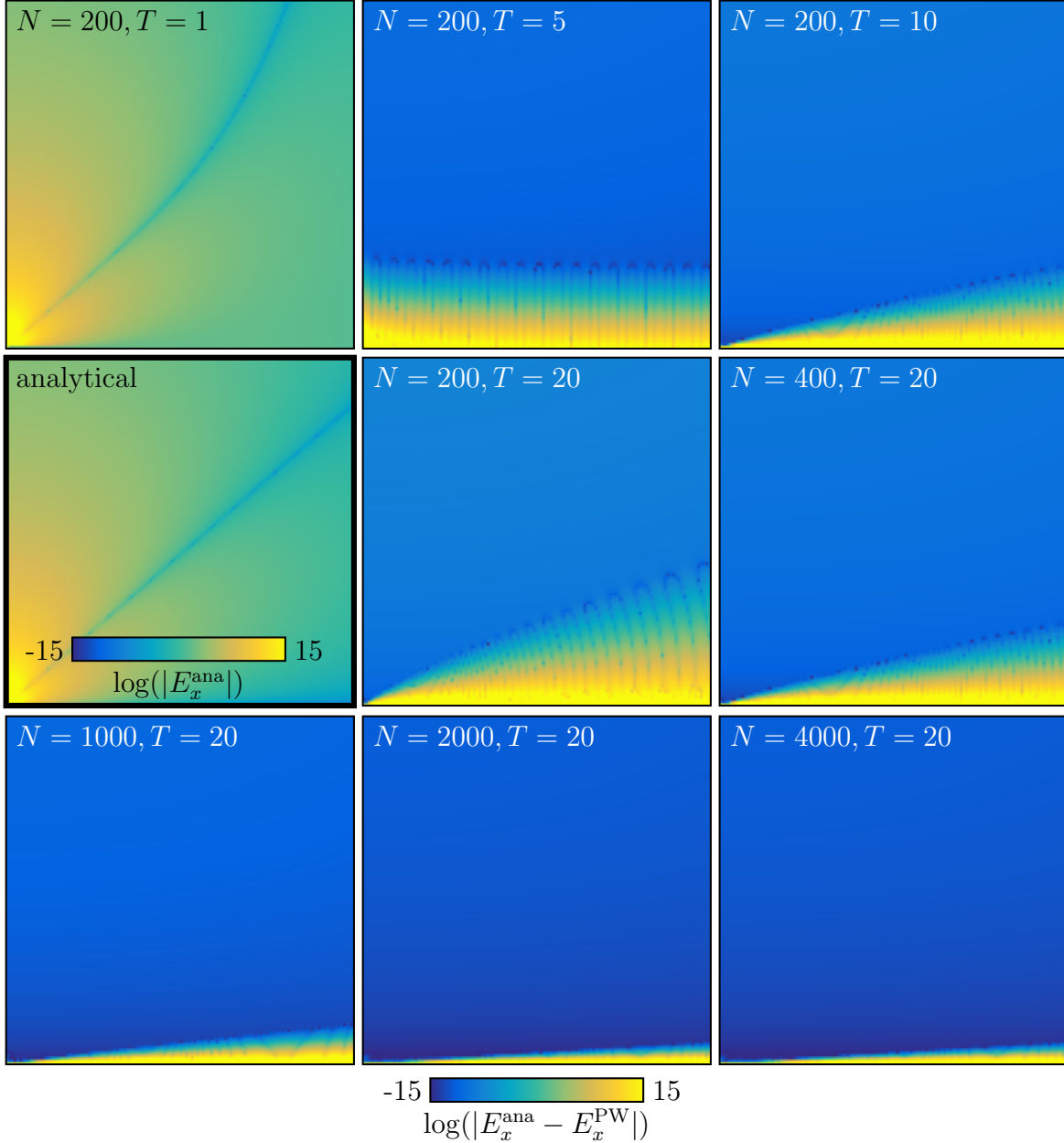
**Figure 3.1:** Convergence behaviour of $E_x := \mathfrak{Re}(M_{4,1}^h)_x$ in the first quadrant of the $x$-$z$-plane ($0 \leq x, z \leq \lambda/4$): The center left shows the analytical result of $\log(|E_x|)$ in order to give an idea of the relevant scales. All other plots depict the logarithmic difference of the analytical result and the PW-decomposition $\log(|E_x^{\mathrm{ana}} - E_x^{\mathrm{PW}}|)$ for varying number $N$ of evanescent waves and varying integration interval $[\pi/2, \pi/2 + iT]$ along the imaginary axis. The integration over propagating waves $\theta' \in [0, \pi/2]$ with 250 integration points is the same for all panels. The parameters of the central panel, $N = 200$ and $T = 20$, are the same as those used for figure 2.5.
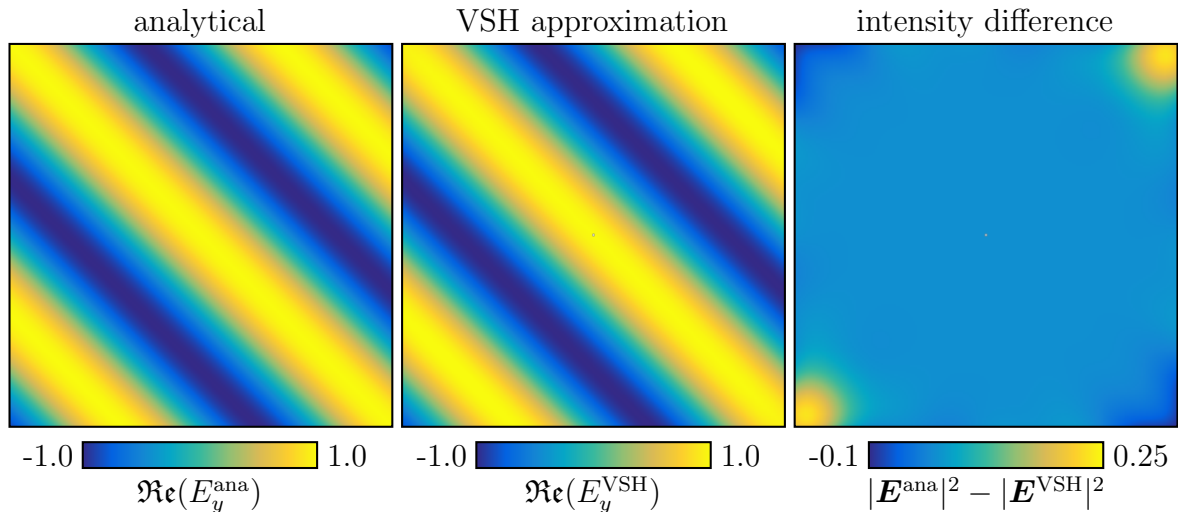
**(a)** VSH decomposition of a propagating s-wave $\boldsymbol{E} := \hat{\boldsymbol{e}}_s \mathrm{e}^{i\boldsymbol{k'} \cdot \boldsymbol{r}}$: Analytical (left) and VSH-approximated (center) version of $\mathfrak{Re}(E_y)$ in the $x$-$z$-plane ($-\lambda \leq x, z \leq \lambda$) with the direction of $\boldsymbol{k'}$ given by $\theta' = \pi/4$, $\phi' = 0$. The right panel depicts the difference between the intensities $|\boldsymbol{E}|^2$ of both forms.
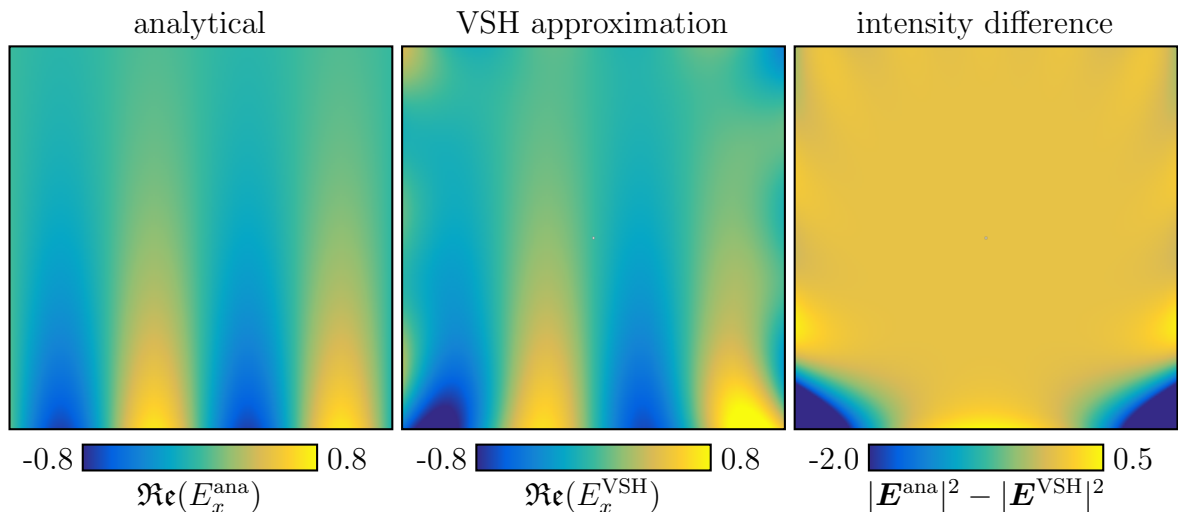


**(b)** VSH decomposition of an evanescent p-wave $\boldsymbol{E} := \hat{\boldsymbol{e}}_p \mathrm{e}^{i\boldsymbol{k'} \cdot \boldsymbol{r}}$: Analytical (left) and VSH-approximated (center) version of $\mathfrak{Re}(E_x)$ in the $x$-$z$-plane ($-\lambda \leq x, z \leq \lambda$) with the direction of $\boldsymbol{k'}$ given by $\theta' = \pi/2 - 0.2i$, $\phi' = 0$. The right panel depicts the difference between the intensities of both forms.

**Figure 3.2:** VSH decompositions of plane waves, summation truncated at $\ell = 10$.

decomposition increases with the distance from the origin. This becomes especially apparent when comparing the intensities, as in the rightmost column of both images. By varying $L$ it is possible to decrease the discrepancy between analytical result and VSH decomposition, as demonstrated in figure 3.3. We found that, for acceptable results up to the same distance from the origin, a larger $L$ is needed for quickly decaying evanescent waves. This is not shown systematically here, but can be anticipated by comparing figure 3.3 with $\theta' = \pi/2 - 2i$ with figure 3.2b with $\theta' = \pi/2 - 0.2i$ and $L = 10$.
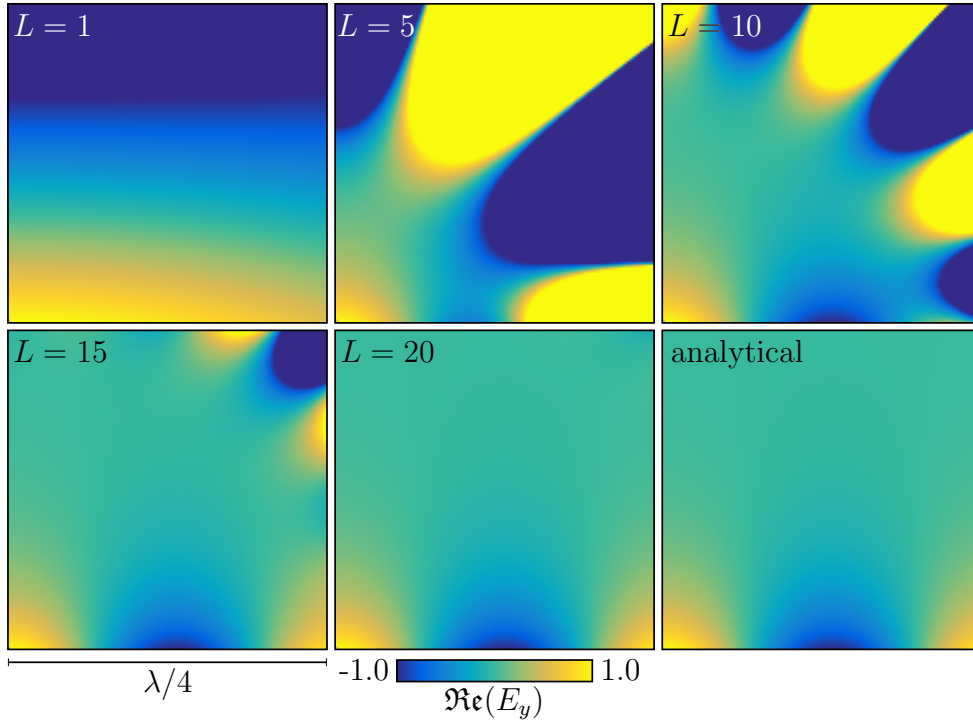
**Figure 3.3:** Convergence behaviour of the VSH decomposition of an evanescent s-wave with $\theta' = \pi/2 - 2i$. Shown is always $\mathfrak{Re}(E_y)$ in the first quadrant of the $x$-$z$-plane ($0 \leq x, z \leq \lambda/4$). The range of the colourmap was chosen to match the analytical solution, the artefacts actually have a much higher magnitude (between -50 and +60 for $L = 5$).

Finally, we turn to the decomposition of the field of an oscillating electric dipole into basis functions. We pointed out in section 2.2.2 that the field of a dipole situated in the origin and oriented parallel to one of the three cartesian axes is given by the linear combination of just one or two vector spherical harmonics. For a dipole with arbitrary position and orientation, the VSH decomposition was given as

$$\boldsymbol{E}_{\text{VSH}}(\boldsymbol{r}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^{D} \boldsymbol{M}_{\ell m}^{f}(k\boldsymbol{r}) + b_{\ell m}^{D} \boldsymbol{N}_{\ell m}^{f}(k\boldsymbol{r}) \qquad 2.80$$

with the expansion coefficients

$$a_{\ell m}^{D} = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{M}_{\ell m}^{g}(k\boldsymbol{r}_0)} \quad \text{and} \quad b_{\ell m}^{D} = \frac{4\pi i n k_v^3}{\ell(\ell+1)} \boldsymbol{p} \cdot \overline{\boldsymbol{N}_{\ell m}^{g}(k\boldsymbol{r}_0)},$$

where $f = h^{(1)}$ and $g = j$ for $r > r_0$ and $f = j$, $g = h^{(2)}$ for $r < r_0$. As in the VSH decomposition of plane waves, the summation over $\ell$ is restricted to values between 1 and $L$ in the concrete numerical implementation. Figure 3.4 shows the effect of $L$ on the VSH-decomposed field of a $z$-dipole that is placed $\lambda/5$ away from the origin on the $x$-axis. For small or large $r/r_0$, the sum quickly approaches the analytical value. However, on the circle with radius $r = r_0$, ringing-artefacts appear. They are reduced and the overall field is improved by choosing a larger $L$. However, since $-\ell \leq m \leq \ell$, this leads to a quadratic increase in memory usage and computation time.
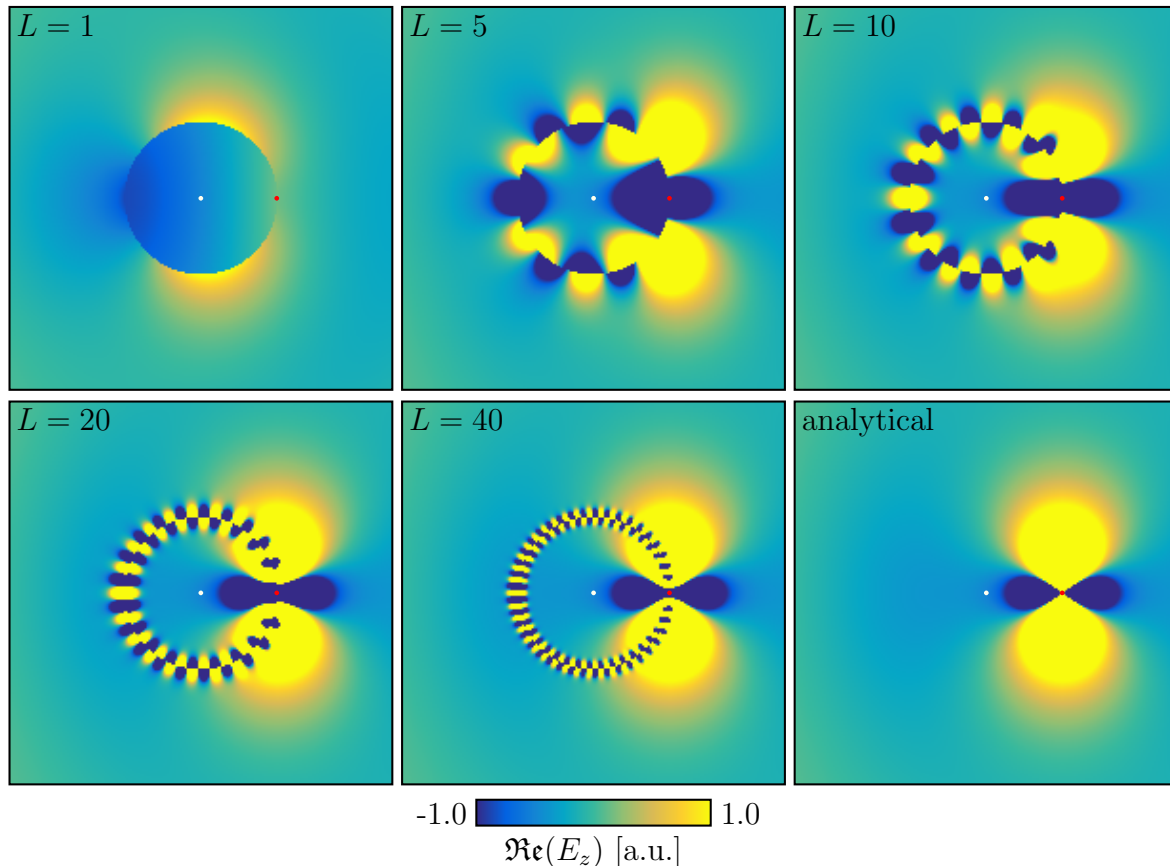
**Figure 3.4:** VSH-decomposition of the electric field of an oscillating $z$-dipole placed on the $x$-axis at $x = \lambda/5$ (red dot), with arbitrary amplitude. Shown is the real part of the $z$-component of the field in the $x$-$z$-plane in the interval $[-\lambda/2, \lambda/2]$. The origin has been marked with a white dot. Note the decrease in artefacts with increasing $L$.

The ringing-artefacts are similar to those encountered near $z = 0$ and $z = z_d$ in the plane wave decompositions of $\boldsymbol{M}_{\ell m}^{h}$, $\boldsymbol{N}_{\ell m}^{h}$ and the dipole field, respectively (compare with figures 2.5 and 2.8). In all three cases, the corresponding decompositions are defined piecewise, using different basis functions for different regions of space[4]. The artefacts always occur at the border between these regions, where they are very hard to reduce. However, they do not impact quantities that are determined far away from the dipole, such as the ADR. They would, however, lead to wrong results if the field directly at the dipole's position was needed. That is why we use the total energy flux according to Chance, Prock and Silbey (equation 2.222) rather than the reflected field at the position of the dipole (equation 2.221) to determine the lifetime of a dipole close to an interface. Unfortunately, this approach is not possible for all experimental situations. For example, if a dipole emitter is very close to a metal nanosphere, the calculated energy flux into the nanosphere will be wrong due to the artefacts in the VSH decomposition.

---

[4] For the plane wave decompositions, we changed the integration interval, and thus the subset of plane and evanescent waves used for the field, at $z = 0$. In the VSH decomposition, we switch between VSH with $R(r) = j(r)$ and $R(r) = h^{(1)}(r)$ at $r = r_0$.

# 3.2 Effective Fresnel Coefficients

After discussing the convergence behaviour of discretized expressions, we now turn to the second important aspect of the implementation of theoretical models on a computer system: the occurrence of numerical errors, and their avoidance through careful design of the used algorithms. A large part of this work is concerned with metal-induced energy transfer and its application to biological samples. From a physical point of view, these situations can be described as a dipole emitter in a stratified system. In order to calculate electric fields and energy fluxes in these systems with many planar interfaces, effective Fresnel reflection and transmission coefficients have to be determined. Their calculation is a good example for the appearance of numbers that are larger than the overflow level in intermediate steps, but not in the final result. In this section, we show how these large numbers can be handled to yield correct final results.

As described in section 2.1.1, the relationship between the amplitudes of incident, reflected and transmitted plane waves at an interface can be expressed through the transfer matrices

$$\underline{\underline{M}}^p_{j,j+1} = \frac{1}{2} \begin{pmatrix} w_{j,j+1}/n_{j,j+1} + n_{j,j+1} & -w_{j,j+1}/n_{j,j+1} + n_{j,j+1} \\ -w_{j,j+1}/n_{j,j+1} + n_{j,j+1} & w_{j,j+1}/n_{j,j+1} + n_{j,j+1} \end{pmatrix} \, , \tag{3.8}$$

$$\underline{\underline{M}}^s_{j,j+1} = \frac{1}{2} \begin{pmatrix} 1 + w_{j,j+1} & 1 - w_{j,j+1} \\ 1 - w_{j,j+1} & 1 + w_{j,j+1} \end{pmatrix} \quad \text{or more general}$$

$$\underline{\underline{M}}_{j,j+1} = \begin{pmatrix} a_{j,j+1} & b_{j,j+1} \\ b_{j,j+1} & a_{j,j+1} \end{pmatrix} \, . \tag{3.9}$$

with $n_{j,j+1} = n_{j+1}/n_j$ the ratio of the refractive indices and $w_{j,j+1} = w_{j+1}/w_j$ the ratio of the components of the wave vector that are perpendicular to the interface. Similarly, the phase difference between a plane wave at the top and bottom of a medium with thickness $d_j$ is described by

$$\underline{\underline{T}}_j = \begin{pmatrix} \exp(-iw_j d_j) & 0 \\ 0 & \exp(iw_j d_j) \end{pmatrix} \, . \tag{3.10}$$

Suppose there is a stack of $J$ layers with refractive indices $n_j$ and thicknesses $d_j$, where the bottommost ($j = 1$) and topmost ($j = J$) layers are infinitely thick. Then the amplitudes of the different waves at the top of the bottommost medium and the bottom of the topmost medium are related via

$$\begin{pmatrix} E^+_{(p,s)1,t} \\ E^-_{(p,s)1,t} \end{pmatrix} = \underline{\underline{M}}^{(p,s)}_{12} \prod_{j=2}^{J-1} \left[ \underline{\underline{T}}_j \underline{\underline{M}}^{(p,s)}_{j,j+1} \right] \begin{pmatrix} E^+_{(p,s)J,b} \\ E^-_{(p,s)J,b} \end{pmatrix}$$

$$:= \underline{\underline{A}}^{(p,s)} \begin{pmatrix} E^+_{(p,s)J,b} \\ E^-_{(p,s)J,b} \end{pmatrix} \, , \tag{3.11}$$

where the superscript denotes waves traveling in the positive (+) or negative (-) $z$-direction. We assume that there are an incident and a reflected wave in the bottommost medium ($E^\pm_{(p,s)1,t} \neq 0$) and a transmitted wave in the topmost medium ($E^+_{(p,s)J,b} \neq 0$,

$E^-_{(p,s)J,b} = 0$). Thus, the effective reflection and transmission coefficients are given by just two elements of the 2x2-matrix $\underline{A}$:

$$r_{p,s} = \frac{E^-_{(p,s)1,t}}{E^+_{(p,s)1,t}} = \frac{A^{(p,s)}_{21}}{A^{(p,s)}_{11}} \quad \text{and} \quad t_{p,s} = \frac{E^+_{(p,s)J,b}}{E^+_{(p,s)1,t}} = \frac{1}{A^{(p,s)}_{11}} \; . \tag{3.12}$$

Usually, these quantities can be easily obtained numerically. However, the exponential functions contained in $\underline{T}_j$ can quickly become larger than the OFL if $w_j$ has a large imaginary part. Consider the case $J = 3$:

$$\underline{\underline{M}}_{12}\underline{\underline{T}}_2\underline{\underline{M}}_{23} = \begin{pmatrix} a_{12}e^{-iw_2d_2}a_{23} + b_{12}e^{iw_2d_2}b_{23} & a_{12}e^{-iw_2d_2}b_{23} + b_{12}e^{iw_2d_2}a_{23} \\ b_{12}e^{-iw_2d_2}a_{23} + a_{12}e^{iw_2d_2}b_{23} & b_{12}e^{iw_2d_2}b_{23} + a_{12}e^{-iw_2d_2}a_{23} \end{pmatrix} \; . \tag{3.13}$$

If $w_2$ has a large positive imaginary part, $\exp(iw_2d_2)$ approaches zero while $\exp(-iw_2d_2)$ becomes very large. Thus, the effective reflection and transmission coeffincients converge to

$$\lim_{\mathfrak{Im}(w_2)\to\infty} r = \frac{b_{12}a_{23}}{a_{12}a_{23}} = \frac{b_{12}}{a_{12}} \quad \text{and} \quad \lim_{\mathfrak{Im}(w_2)\to\infty} t = \frac{\exp(-\mathfrak{Im}(w_2)d_2)}{a_{12}a_{23}} = 0 \; . \tag{3.14}$$

The large exponential functions in the reflection coefficient cancel, leading to the same result that would have been obtained if the second medium had been infinitely thick – the strongly evanescent wave is not influenced by the other layers because it decays before reaching them. Accordingly, the transmission coeffiecient is zero.

However, when the matrix multiplication and subsequent division of matrix elements is done numerically, $\exp(-iw_2d_2) >$ OFL leads to wrong results for the divisions. In order to identify and handle these exceptions, we test the size of $\exp(-iw_2d_2)$. If it is larger than a certain threshold (which is only slightly smaller than the OFL), both the transmission coeffiecient and all terms that have to be multiplied by $\exp(+iw_2d_2)$ are set to zero. Furthermore, the terms for the reflection coefficients are not multiplied with $\exp(-iw_2d_2)$ since it cancels anyway.

To prove that this approach still yields correct results in the general situation with more than three layers, we exploit the fact that only the left column of all matrices needs to be known: According to equation (3.12), $r$ and $t$ are defined by $A_{11}$ and $A_{12}$. But these can be calculated using only the left columns of all intermediate matrices, if the product in equation (3.11) is calculated with descending order of $j$ (i.e. from right to left):

$$\underline{\underline{M}}_{J-1,J} = \begin{pmatrix} a_{J-1} & \cdots \\ b_{J-1} & \cdots \end{pmatrix} \tag{3.15}$$

$$\underline{\underline{M}}_{J-2,J} = \underline{\underline{M}}_{J-2,J-1}\underline{\underline{T}}_{J-1}\underline{\underline{M}}_{J-1,J}$$

$$= \begin{pmatrix} a_{J-2}e^-_{J-1}a_{J-1} + b_{J-2}e^+_{J-1}b_{J-1} & \cdots \\ b_{J-2}e^-_{J-1}a_{J-1} + a_{J-2}e^+_{J-1}b_{J-1} & \cdots \end{pmatrix} \tag{3.16}$$

$$\underline{\underline{M}}_{J-3,J} = \underline{\underline{M}}_{J-3,J-2}\underline{\underline{T}}_{J-2}\underline{\underline{M}}_{J-2,J-1}\underline{\underline{T}}_{J-1}\underline{\underline{M}}_{J-1,J}$$

$$= \begin{pmatrix} a_{J-3}e^-_{J-2}[a_{J-2}e^-_{J-1}a_{J-1} + b_{J-2}e^+_{J-1}b_{J-1}] & \cdots \\ \quad + b_{J-3}e^+_{J-2}[b_{J-2}e^-_{J-1}a_{J-1} + a_{J-2}e^+_{J-1}b_{J-1}] & \\ b_{J-3}e^-_{J-2}[a_{J-2}e^-_{J-1}a_{J-1} + b_{J-2}e^+_{J-1}b_{J-1}] & \cdots \\ \quad + a_{J-3}e^+_{J-2}[b_{J-2}e^-_{J-1}a_{J-1} + a_{J-2}e^+_{J-1}b_{J-1}] & \end{pmatrix}, \qquad (3.17)$$

and so on. Here, we shortened the notation to $a_{j,j+1} := a_j$, $b_{j,j+1} := b_j$ and $\exp(\pm i w_j d_j) := e^\pm_j$. For small or intermediate values of $\mathfrak{Im}(w_j)$, all of the terms contribute to the final result, and the matrix multiplication can be carried out exactly. For large $w_j$, the same problem of finite data size arises, but now it is hard to say which of the many mixed exponential terms can be neglected. However, if $\mathfrak{Im}(w_1)$ is large enough, then

$$w_j = \sqrt{k_0^2 n_j^2 - k_0^2 n_1^2 - \mathfrak{Im}(w_1)^2} \to w_1 , \qquad (3.18)$$

i.e. all $w_j$ become imaginary and very large. Then, the only terms that contribute significantly to the matrix elements are those that contain only $e^-_j$, all other terms can be ignored. But this just corresponds to the result for three layers:

$$r = \frac{b_{12}}{a_{12}} \quad \text{and} \quad t = 0. \qquad (3.19)$$

The explanation is the same as before, a strongly evanescent wave only "explores" the first interface, and is therefore not affected by all further layers. In our numerical implementation, all refractive indices and thicknesses as well as $w_1$ are given. The matrix multiplication (3.11) is carried out numerically with descending $j$ (from right to left). If for any $j$ the term $\exp(-i w_j d_j)$ is larger than the overflow level, the multiplication is stopped and the result (3.19) is used instead.

## 3.3 Associated Legendre polynomials

Besides the overflow of floating-point numbers treated in the previous section, the accumulation of rounding errors is another important aspect to keep in mind when designing an algorithm. In particular, it influences the stability of recursive expressions, such as the definition of associated Legendre polynomials via recurrance relations.

Associated Legendre polynomials are needed for all calculations involving vector spherical harmonics. As already pointed out in [101], chapter 6.7, "There are many bad ways to evaluate associated Legendre polynomials numerically." Press *et al.* give an explicit expression for $P^m_l(x)$, which is a sum of terms with alternating signs that lead to "delicate cancellations between successive terms". Since these summands quickly grow large for increasing $\ell$, the accuracy of the result breaks down already for $\ell \sim 15 - 18$ [101]. Therefore, a recursive definition of $P^m_\ell(x)$ has to be chosen when $\ell > 15$ is needed. Mathematically, a large number of different recurrence formulas for associated Legendre polynomials exist, see e.g. [102], chapter 8. However, according to [101], "most of the recurrences involving $m$ [are] unstable, and so are dangerous for numerical work".

Therefore, we started out by using the only expression that is described as stable by the authors of [101]:

$$P_0^0(x) = 1,$$
$$P_\ell^\ell(x) = -(2\ell - 1)\sqrt{1 - x^2}P_{\ell-1}^{\ell-1}(x) \qquad\qquad \forall \ell \geq 1,$$
$$P_{\ell+1}^\ell(x) = (2\ell + 1)\ x\ P_\ell^\ell(x) \qquad\qquad \forall \ell \geq 1,$$
$$P_\ell^m(x) = \frac{2\ell - 1}{\ell - m} \cdot x \cdot P_{\ell-1}^m - \frac{\ell + m - 1}{\ell - m}\ P_{\ell-2}^m \qquad \forall m \geq 0, \ell \geq m + 2, \quad (3.20)$$
$$\frac{\mathrm{d}P_0^0(x)}{\mathrm{d}x} = 0,$$
$$\frac{\mathrm{d}P_\ell^0(x)}{\mathrm{d}x} = \frac{\ell}{1 - x^2} \cdot \left[P_{\ell-1}^0(x) - x \cdot P_\ell^0(x)\right] \qquad\qquad \forall \ell \geq 1,$$
$$\frac{\mathrm{d}P_\ell^m(x)}{\mathrm{d}x} = \frac{mx}{1 - x^2}P_\ell^m(x) + \frac{(\ell + m)(\ell - m + 1)}{\sqrt{1 - x^2}}P_\ell^{m-1}(x) \qquad \forall m \geq 1, \ell \geq m. \quad (3.21)$$

However, since the accociated Legendre polynomials themselves grow quite rapidly with $\ell$ and $m$, even this formula quickly reaches values that are larger than the overflow level. This is demonstrated exemplarily for $\ell = m$ and $x = 0.1$, $x = 40i$ in figure 3.5. Contrarily, the normalization factor $c_{\ell m}$ of the vector spherical harmonics becomes very small for large $\ell, m$, as also shown in the figure. Since we predominantly need $P_\ell^m$ for use in VSH, it is convenient to include this factor in the recurrence relations and thus keep the terms manageable. This was already suggested in [101] under the name *renormalized associated Legendre function*. But we do not only need $c_{\ell m}P_\ell^m(x)$, we are also interested in the derivative $c_{\ell m}\mathrm{d}P_\ell^m(x)/\mathrm{d}x$, mostly multiplied with additional $x$-dependent terms. The latter have been included in recurrence relations presented in [49], but without $c_{\ell m}$. Therefore, our second approach combines [101] with [49]. We define three auxiliary functions $\pi_{m\ell}(x)$, $\tau_{m\ell}(x)$ and $\gamma_{m\ell}(x)$ that are useful when evaluating VSH,

$$\pi_{m\ell}(x) := \sqrt{\frac{(2\ell + 1)}{4\pi}\frac{(\ell - m)!}{(\ell + m)!}}\frac{m}{\sqrt{1 - x^2}}P_\ell^m(x) \qquad = c_{\ell m}\frac{m}{\sqrt{1 - x^2}}P_\ell^m(x)\,, \qquad (3.22)$$

$$\tau_{m\ell}(x) := -\sqrt{\frac{(2\ell + 1)}{4\pi}\frac{(\ell - m)!}{(\ell + m)!}}\sqrt{1 - x^2}\frac{\mathrm{d}P_\ell^m(x)}{\mathrm{d}x} \qquad = -c_{\ell m}\sqrt{1 - x^2}\frac{\mathrm{d}P_\ell^m(x)}{\mathrm{d}x}\,, \quad (3.23)$$

$$\gamma_{m\ell}(x) := \sqrt{\frac{(2\ell + 1)}{4\pi}\frac{(\ell - m)!}{(\ell + m)!}}\ell(\ell + 1)P_\ell^m(x) \qquad = c_{\ell m}\ell(\ell + 1)P_\ell^m(x)\,. \qquad (3.24)$$

These auxiliary functions are calculated according to

$$\pi_{0,\ell}(x) = 0 \qquad\qquad \forall \ell \geq 1,$$
$$\pi_{1,1}(x) = -\sqrt{\frac{3}{8\pi}}\,,$$
$$\pi_{\ell,\ell}(x) = -\sqrt{\frac{\ell(2\ell + 1)}{2}} \cdot \frac{\sqrt{1 - x^2}}{\ell - 1} \cdot \pi_{\ell-1,\ell-1}(x) \qquad\qquad \forall \ell \geq 2,$$
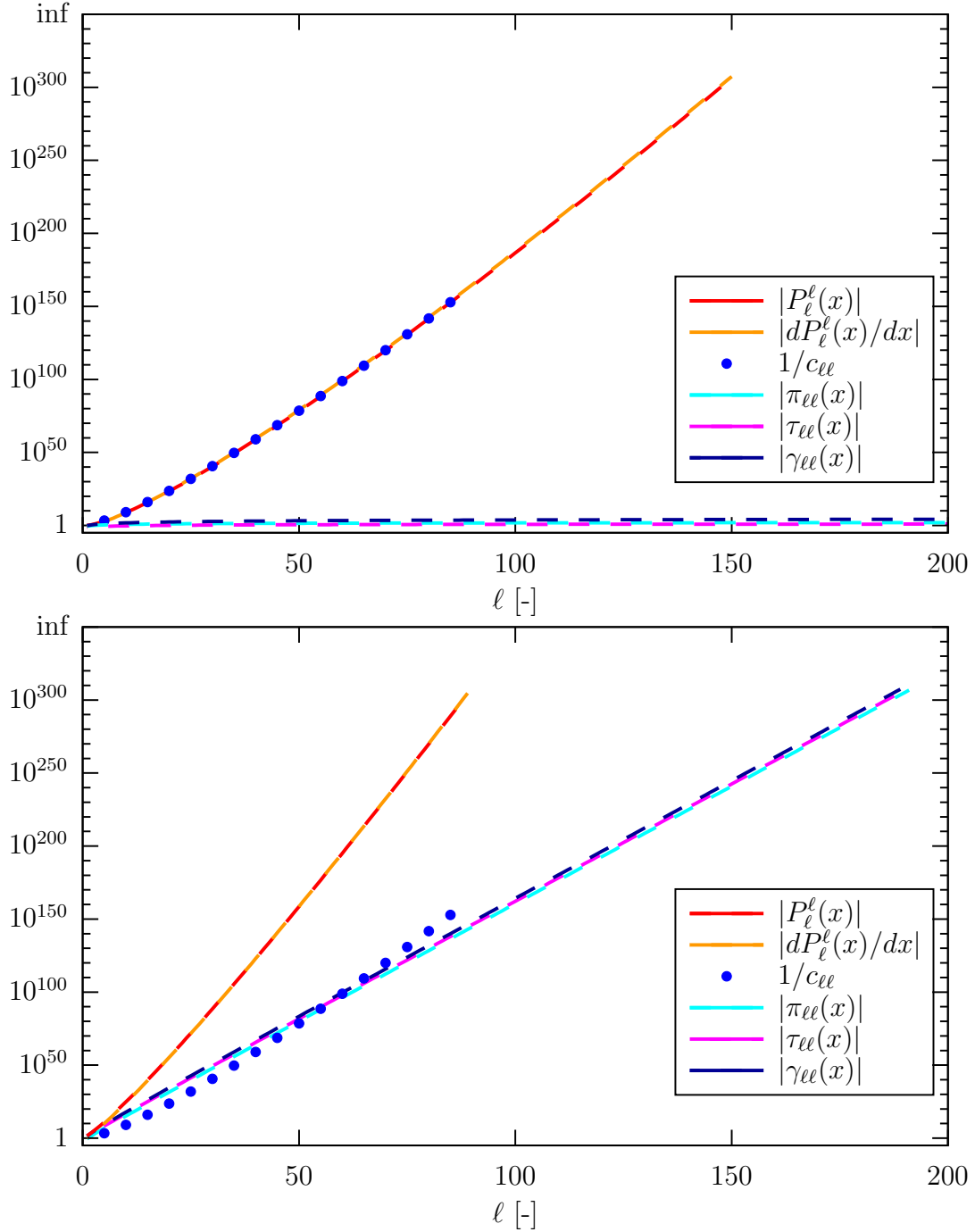
**Figure 3.5:** Behaviour of the associated Legendre polynomials $P_\ell^m(x)$ and some modified functions for large $\ell, m$; shown for the special case $\ell = m$ and two different arguments $x$. **Top:** $x = 0.1$: $|P_\ell^\ell(x)|$ and $|\mathrm{d}P_\ell^\ell(x)/\mathrm{d}x|$ can be calculated up to $\ell \approx 150$, then they are too large for a double precision floating point number. The normalization coefficient $c_{\ell\ell}$ becomes so small that Matlab sets it to zero at $\ell \approx 90$. In contrast, the functions $\pi_{\ell\ell}(x)$, $\tau_{\ell\ell}(x)$ and $\gamma_{\ell\ell}(x)$ only grow slowly. **Bottom:** $x = 40i$: $|P_\ell^\ell(x)|$ and $|\mathrm{d}P_\ell^\ell(x)/\mathrm{d}x|$ grow much more quickly than $c_{\ell\ell}^{-1}$, leading to a strong increase of the normalized functions $\pi_{\ell\ell}(x)$, $\tau_{\ell\ell}(x)$ and $\gamma_{\ell\ell}(x)$ with $\ell$. The former two can be calculated up to $\ell \approx 90$, the latter three for up to twice as many orders.

$$\pi_{\ell,\ell+1}(x) = \sqrt{2\ell+3} \cdot x \cdot \pi_{\ell,\ell}(x) \qquad\qquad \forall \ell \geq 1,$$

$$\pi_{m,\ell}(x) = \sqrt{\frac{(2\ell+1)(2\ell-1)}{(\ell+m)(\ell-m)}} \cdot x \cdot \pi_{m,\ell-1}(x)$$

$$- \sqrt{\frac{(2\ell+1)(\ell+m-1)(\ell-m-1)}{(2\ell-3)(\ell+m)(\ell-m)}} \pi_{m,\ell-2}(x) \qquad \forall m \geq 1, l \geq m+2,$$

$$(3.25)$$

$$\tau_{0,1}(x) = -\sqrt{\frac{3}{4\pi}}\sqrt{1-x^2} \ ,$$

$$\tau_{0,2}(x) = -\sqrt{\frac{5}{4\pi}}\sqrt{1-x^2}3x \ ,$$

$$\tau_{1,1}(x) = -\sqrt{\frac{3}{8\pi}}x \ ,$$

$$\tau_{0,\ell}(x) = \frac{\sqrt{(2\ell+1)(2\ell-1)}}{\ell-1}x\tau_{0,\ell-1} - \sqrt{\frac{(2\ell+1)}{(2\ell-3)}}\frac{\ell}{\ell-1}\tau_{0,\ell-2} \qquad\qquad \forall \ell \geq 3,$$

$$\tau_{m\ell}(x) = \frac{\ell}{m}x\pi_{m\ell}(x) - \frac{1}{m}\sqrt{\frac{(2\ell+1)(\ell-m)(\ell+m)}{(2\ell-1)}}\pi_{m,\ell-1}(x) \qquad \forall m \geq 1, \ell \geq m,$$

$$(3.26)$$

$$\gamma_{0,1}(x) = \sqrt{\frac{3}{\pi}}x \ ,$$

$$\gamma_{0,2}(x) = \sqrt{\frac{5}{\pi}}\frac{3}{2}(3x^2-1) \ ,$$

$$\gamma_{0,\ell}(x) = \frac{(\ell+1)}{\ell(\ell-1)}\sqrt{(2\ell+1)(2\ell-1)}x\gamma_{0,\ell-1}(x)$$

$$- \frac{(\ell+1)}{(\ell-2)}\sqrt{\frac{2\ell+1}{2\ell-3}}\gamma_{0,\ell-2}(x) \qquad\qquad \forall \ell \geq 3,$$

$$\gamma_{m\ell}(x) = \frac{\ell(\ell+1)}{m}\sqrt{1-x^2}\pi_{m\ell}(x) \qquad\qquad \forall m \geq 1, \ell \geq m.$$

$$(3.27)$$

These three functions are shown in the same situations as the pure associated Legendre polynomials in figure 3.5. For real $x$, these functions grow very slowly, with $\pi_{200,200}(0.1) \approx 83$, $\tau_{200,200}(0.1) \approx 8.3$ and $\gamma_{200,200}(0.1) \approx 1.7 \cdot 10^4$. For imaginary $x$ – needed e.g. in the decomposition of evanescent waves – the three functions grow more quickly. However, they can still be evaluated up to much higher orders than $P_\ell^m(x)$ and $\mathrm{d}P_\ell^m(x)/\mathrm{d}x$. Thus, the three auxiliary functions $\pi_{\ell m}(x)$, $\tau_{\ell m}(x)$ and $\gamma_{\ell m}(x)$ allow to evaluate VSH up to higher orders than the usual recurrence relations for associated Legendre polynomials. Therefore, these three functions were used in our work.

## 3.4 Energy flux through an infinite plane

Numerical errors can be avoided if the number of operations that have to be performed numerically is reduced. This can be achieved by choosing a coordinate system and basis functions that are suited to the geometry of the problem, and by evaluating integrals analytically when possible. In this section, we present an example of such prior considerations that is vital for the later evaluation of MIET experiments, namely the total energy flux of a dipole emitter close to a metal layer.

In section 2.4.5, we determined the angular distribution of radiation $\mathrm{d}^2 S/\mathrm{d}\Omega^2(\theta, \phi)$ of a dipole in a stratified environment, i.e. in a system of several planar layers consisting of different materials. If at least one of these materials is absorbing (e.g. a metal), the energy emitted in the far-field $\int (\mathrm{d}^2 S/\mathrm{d}\Omega^2)\mathrm{d}\Omega^2$ is smaller than the total energy emitted by the dipole $S_{\mathrm{tot}}$. The latter can be obtained by integrating the time-averaged Poynting vector over a surface that encloses the dipole but no other energy sources or sinks. Since the interaction of the dipole field with planar interfaces is most easily described using plane waves, integrating over a spherical surface as we did for spherical geometries in section 2.4.6 seems ill-advised. Instead, we form a closed integration surface by integrating over two infinite planes, one above and one below the emitter, but both in the same medium as the emitter[5]. Provided that the medium containing the dipole has a real refractive index, which is a reasonable assumption for all our experiments, the sum of the two integrals then yields $S_{\mathrm{tot}}$.

We start by considering the simple case that the dipole is situated in a medium with refractive index $n_1$ which occupies the whole upper halfspace $z > 0$, while arbitrary layers make up the lower halfspace $z < 0$. Then, the interface between the dipole emitter's medium and the adjoining material at $z = 0$ is a good choice for the lower of the two integration surfaces. Denoting the $z$-position of the dipole as $z_0$, we saw in section 2.3.2 that the field at $z = 0$ is given by:

$$\boldsymbol{E}(x, y, 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \ \sin\theta_1 \exp\left(iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0\right)$$
$$\cdot \left[(\hat{\boldsymbol{e}}_{p1} + R_p \hat{\boldsymbol{e}}_{p1}^*)(\hat{\boldsymbol{e}}_{p1} \cdot \boldsymbol{p}) + (1 + R_s)\hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p})\right]$$

with the path $\quad \mathcal{C} = \pi/2 + i\infty \to \pi/2 \to \pi \quad$ and with $\theta_1^* := \pi - \theta_1$. $\qquad$ (3.28)

The magnetic field can be obtained via $\boldsymbol{B} = \nabla \times \boldsymbol{E}/k_v$ using $\hat{\boldsymbol{e}}_k \times \hat{\boldsymbol{e}}_p = \hat{\boldsymbol{e}}_s$:

$$\boldsymbol{B}(x, y, 0) = \frac{in_1^2 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \ \sin\theta_1 \exp\left(iq\rho_d \cos(\phi - \phi_d) - ik_1 \cos\theta_1 z_0\right)$$
$$\cdot \left[(1 + R_p)\hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{p1} \cdot \boldsymbol{p}) - (\hat{\boldsymbol{e}}_{p1} + R_s \hat{\boldsymbol{e}}_{p1}^*)(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p})\right]. \quad (3.29)$$

These two fields cause a total flux through the plane $z = 0$ that is given by:

$$S_- = \frac{c}{8\pi}\mathfrak{Re}\left\{\int_{z=0} \mathrm{d}A \ \hat{\boldsymbol{e}}_z \cdot (\boldsymbol{E} \times \overline{\boldsymbol{B}})\right\}$$

---

[5] Mathematically speaking, two parallel planes intersect at infinity. From a physical point of view, we can simply say that all energy except the negligible fraction being emitted at *exactly* $\theta = \pi/2$ will cross one of the two planes sooner or later.

# Numerical Implementation

$$
= \frac{cn_1^3 k_v^6}{32\pi^3} \mathfrak{Re} \left\{ \int \mathrm{d}\theta_1 \int \mathrm{d}\phi \int \mathrm{d}\theta_1' \int \mathrm{d}\phi' \int \mathrm{d}^2\boldsymbol{\rho} \ \sin\theta_1 \sin\theta_1' \ \hat{\boldsymbol{e}}_z \right.
$$

$$
\cdot \quad \exp\left[i(\boldsymbol{q}-\boldsymbol{q}')(\boldsymbol{\rho}-\boldsymbol{\rho}_0) - ik_1(\cos\theta_1 - \overline{\cos\theta_1'})z_0\right]
$$

$$
\cdot \quad \left[(\hat{\boldsymbol{e}}_{p1} + R_p\hat{\boldsymbol{e}}_{p1}^*)(\hat{\boldsymbol{e}}_{p1}\cdot\boldsymbol{p}) + (1+R_s)\hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{s1}\cdot\boldsymbol{p})\right]
$$

$$
\times \quad \overline{\left[(1+R_p')\hat{\boldsymbol{e}}_{s1}'(\hat{\boldsymbol{e}}_{p1}'\cdot\boldsymbol{p}) - (\hat{\boldsymbol{e}}_{p1}' + R_s'\hat{\boldsymbol{e}}_{p1}'^*)(\hat{\boldsymbol{e}}_{s1}'\cdot\boldsymbol{p})\right]} \right\}. \tag{3.30}
$$

The naive implementation of this equation would require the numerical evaluation of five integrals, entailing a high risk of numerical errors. However, in this formula, we recognize the definition of Dirac's delta-distribution:

$$
\delta^2(\boldsymbol{q}-\boldsymbol{q}') = \frac{1}{(2\pi)^2} \int \mathrm{d}^2\boldsymbol{\rho} \exp[i\boldsymbol{\rho}\cdot(\boldsymbol{q}-\boldsymbol{q}')]. \tag{3.31}
$$

Taking the integral over $\boldsymbol{\rho}$ thus produces the term $\delta^2(\boldsymbol{q}-\boldsymbol{q}')$, which, when integrating over the whole $\boldsymbol{q}'$-plane, is zero everywhere except at $\boldsymbol{q}' = \boldsymbol{q}$. Taking into account that $\boldsymbol{q}' = k_1\sin\theta_1'(\cos\phi'\hat{\boldsymbol{e}}_x + \sin\phi'\hat{\boldsymbol{e}}_y)$, the integral over $\boldsymbol{q}'$-space is given by $q'\mathrm{d}q'\mathrm{d}\phi' = k_1^2\sin\theta_1'\cos\theta_1'\mathrm{d}\theta_1'\mathrm{d}\phi'$. We can thus take the integrals over $\boldsymbol{\rho}$, $\theta_1'$ and $\phi'$ analytically:

$$
S_- = \frac{cn_1 k_v^4}{8\pi} \mathfrak{Re} \left\{ \int \mathrm{d}\theta_1 \int \mathrm{d}\phi \ \frac{\sin\theta_1}{\cos\theta_1} \exp\left[2k_1\mathfrak{Im}(\cos\theta_1)z_0\right] \ \hat{\boldsymbol{e}}_z \right.
$$

$$
\cdot \quad \left. \left[(\hat{\boldsymbol{e}}_{k1} + R_p\hat{\boldsymbol{e}}_{k1}^*)\overline{(1+R_p)}|\hat{\boldsymbol{e}}_{p1}\cdot\boldsymbol{p}|^2 + (1+R_s)\overline{(\hat{\boldsymbol{e}}_{k1} + R_s\hat{\boldsymbol{e}}_{k1}^*)}|\hat{\boldsymbol{e}}_{s1}\cdot\boldsymbol{p}|^2\right] \right\}
$$

$$
= \frac{cn_1 k_v^4}{8\pi} \mathfrak{Re} \left\{ \int \mathrm{d}\theta_1 \int \mathrm{d}\phi \ \sin\theta_1 \exp\left[2k_1\mathfrak{Im}(\cos\theta_1)z_0\right] \right.
$$

$$
\cdot \quad \left. \left[(1-R_p)\overline{(1+R_p)}|\hat{\boldsymbol{e}}_{p1}\cdot\boldsymbol{p}|^2 + (1+R_s)\overline{(1-R_s)}|\hat{\boldsymbol{e}}_{s1}\cdot\boldsymbol{p}|^2\right] \right\}
$$

$$
= \frac{cn_1 k_v^4}{8\pi} \int_{\pi/2}^{\pi} \mathrm{d}\theta_1 \int \mathrm{d}\phi \ \sin\theta_1 \left[\left(1-|R_p|^2\right)|\hat{\boldsymbol{e}}_{p1}\cdot\boldsymbol{p}|^2 + \left(1-|R_s|^2\right)|\hat{\boldsymbol{e}}_{s1}\cdot\boldsymbol{p}|^2\right]
$$

$$
+ \frac{cn_1 k_v^4}{8\pi} \int_{\pi/2+i\infty}^{\pi/2} \mathrm{d}\theta_1 \int \mathrm{d}\phi \ \sin\theta_1 \exp\left[2k_1\mathfrak{Im}(\cos\theta_1)z_0\right]
$$

$$
\cdot \quad \left[-2i\mathfrak{Im}(R_p)|\hat{\boldsymbol{e}}_{p1}\cdot\boldsymbol{p}|^2 + 2i\mathfrak{Im}(R_s)|\hat{\boldsymbol{e}}_{s1}\cdot\boldsymbol{p}|^2\right] \tag{3.32}
$$

where we used $\hat{\boldsymbol{e}}_z \cdot (\hat{\boldsymbol{e}}_{p1} \times \hat{\boldsymbol{e}}_{s1}) = \hat{\boldsymbol{e}}_z \cdot \hat{\boldsymbol{e}}_{k1} = \cos\theta_1$, and equivalently $\hat{\boldsymbol{e}}_z \cdot (\hat{\boldsymbol{e}}_{p1}^* \times \hat{\boldsymbol{e}}_s) = \cos\theta_1^* = -\cos\theta_1$. Furthermore, we know that $\sin(\theta_1)$ is real for both $\theta_1 \in \mathbb{R}$ and $\theta_1 = \pi/2 + i\theta''$, while $\cos(\theta_1)$ can be either real ($\theta_1 \in \mathbb{R}$) or imaginary ($\theta_1 = \pi/2 + i\theta''$). Note that, since $\theta'' \geq 0$, $\mathfrak{Im}(\cos\theta_1) \leq 0$ – the exponential function has a damping effect on the amplitude, as expected for evanescent waves. We split the integral into propagating and evanescent waves in order to take the real part because the integration measure $\mathrm{d}\theta_1$ is real for real $\theta_1$ and imaginary for the path $\theta_1 = \pi/2 + i\infty \to \pi/2$. Since we assumed that no other materials are present above the dipole, there is also no energy sink, and we can choose an arbitrary $z > z_0$ for the upper integration surface. It then

(a) geometry       (c) phase accumulated due to reflections
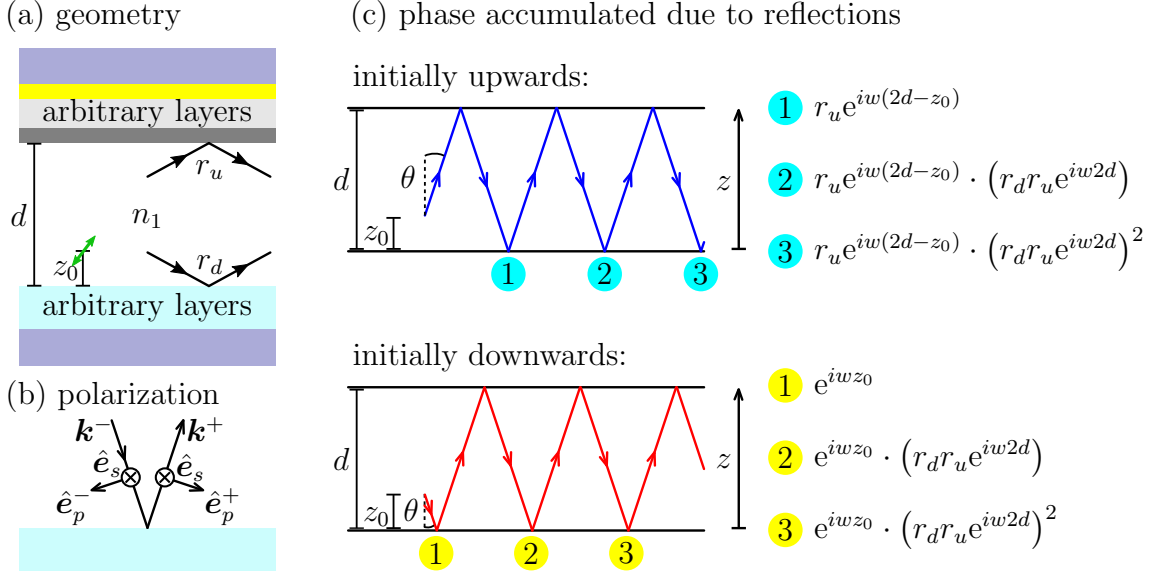


**Figure 3.6:** Multiple reflections in a stratified sample. (a) Sample geometry: The dipole (green arrow) is situated at $z = z_0$ in a medium with refractive index $n_1$ that extends from $z = 0$ to $z = d$. The interaction with the layers above and below is described by the effective reflection coefficients $r_u$ and $r_d$, which were calculated in section 3.2. (b) Unit vectors of the polarization of $p$- and $s$-waves for downwards (-) and upwards (+) traveling waves. (c) Due to multiple reflections between the top and bottom interface of the dipole's medium, a plane wave changes its amplitude and accumulates a phase shift. When we sum these waves, we do so at fixed $\boldsymbol{r}$, thus the phase is only acquired due to propagation along $z$. The shift along the perpendicular direction is only shown for clear labeling of the waves. The expressions on the right give the phase and amplitude for downwards traveling waves at the marked positions, assuming that the amplitude was one and the phase zero at the position of the dipole.

makes sense to simply use the far-field ADR from equation (2.170) and integrate it over $\theta_1 \in [0, \pi/2]$.

The situation gets slightly more complex if other materials are present above the dipole, too. Then, a plane wave that hits the lower interface gets partially reflected, reaches the upper interface, gets partially reflected again and so on. In order to take this into account, we consider the geometry in figure 3.6(a): The dipole's medium with refractive index $n_1$ fills the space $0 \leq z \leq d$. Above and below, there are arbitrary other materials. The effective Fresnel reflection coefficients for the upper and lower interfaces at $z = d$ and $z = 0$ are $r_{u,p/s}$ and $r_{d,p/s}$, respectively, calculated according to section 3.2. With these multiple reflections, it is easier to define $\theta_1$ as in figure 3.6(c)[6] and to take the different propagation directions into account via the sign of $w_1 = k_1 \cos\theta_1$.

The expressions on the right side of figure 3.6(c) correspond to the amplitude and phase that accumulate due to the multiple reflections. The electric field at $z = 0$ is then composed of four different types of waves, those that have been originally emitted into the upper or lower halfspace by the dipole (blue or red in the figure), which are each either travelling in positive or negative $z$-direction at $z = 0$. For $p$-waves, the former

---

[6] That is, with $\theta_1 \in [0, \pi/2]$ for propagating waves and $\theta_1 \in \pi/2 + i[0, \infty]$ for evanescent waves.

category determines the initial amplitude according to $\hat{\boldsymbol{e}}_{p1}^{\pm} \cdot \boldsymbol{p}$, while the latter category gives the polarization as $\hat{\boldsymbol{e}}_{p1}^{\pm}$ (see figure 3.6(b)). For $s$-waves, all four types of waves have the same initial amplitude and polarization. The phase and final amplitude of each wave type at $z = 0$ is found by summing up all single plane waves belonging to the same category:

$$\text{initially down, now down:} \qquad \frac{e^{iw_1 z_0}}{1 - r_u r_d\, e^{2iw_1 d}} =: c_{dd} \;,$$

$$\text{initially down, now up:} \qquad \frac{e^{iw_1 z_0}\; \cdot\; r_d}{1 - r_u r_d\, e^{2iw_1 d}} =: c_{du} \;,$$

$$\text{initially up, now down:} \qquad \frac{e^{iw_1 (2d - z_0)} \cdot r_u}{1 - r_u r_d\, e^{2iw_1 d}} =: c_{ud} \;,$$

$$\text{initially up, now up:} \qquad \frac{e^{iw_1 (2d - z_0)} \cdot r_u r_d}{1 - r_u r_d\, e^{2iw_1 d}} =: c_{uu} \;, \qquad (3.33)$$

where we used the fact that each sum is a geometric series and thereby eliminated another source of numerical errors. Thus, when going back to expression (3.28) for the electric field at the interface, all we have to do is replace the single plane waves by these sums:

$$\boldsymbol{E}(x, y, 0) = \frac{in_1 k_v^3}{2\pi} \int_{\mathcal{C}} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi \; \sin\theta_1 \exp\left(iq\rho_d \cos(\phi - \phi_d)\right)$$

$$\cdot \left[ (c_{dd,p}\hat{\boldsymbol{e}}_{p1}^- + c_{du,p}\hat{\boldsymbol{e}}_{p1}^+)(\hat{\boldsymbol{e}}_{p1}^- \cdot \boldsymbol{p}) + (c_{ud,p}\hat{\boldsymbol{e}}_{p1}^- + c_{uu,p}\hat{\boldsymbol{e}}_{p1}^+)(\hat{\boldsymbol{e}}_{p1}^+ \cdot \boldsymbol{p}) \right.$$

$$\left. + (c_{dd,s} + c_{du,s} + c_{ud,s} + c_{uu,s})\hat{\boldsymbol{e}}_{s1}(\hat{\boldsymbol{e}}_{s1} \cdot \boldsymbol{p}) \right]$$

$$\text{with the path} \quad \mathcal{C} = 0 \rightarrow \pi/2 \rightarrow \pi/2 - i\infty. \qquad (3.34)$$

The total flux through the plane $z = 0$, $S_-$, can then be derived as before. Since the reflection coefficients do not depend on $\phi$, the $\phi$-integration can be carried out analytically.

We give the explicit result for two special cases, a vertical dipole with $\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_z$, and a horizontal dipole with $\boldsymbol{p} \parallel \hat{\boldsymbol{e}}_x$:

$$\boxed{\begin{aligned}
S_{-,\perp} &= \frac{cn_1 k_v^4 p^2}{4} \int_0^{\pi/2} \mathrm{d}\theta_1 \; \sin\theta_1^3 \; \left[ |c_{dd,p} + c_{ud,p}|^2 - |c_{du,p} + c_{uu,p}|^2 \right] \\
&+ \frac{cn_1 k_v^4 p^2}{2} \int_{\pi/2}^{\pi/2 - i\infty} \mathrm{d}\theta_1 \; \sin\theta_1^3 \; i\Im\mathfrak{m}\left[ \overline{(c_{dd,p} + c_{ud,p})}(c_{du,p} + c_{uu,p}) \right] , \qquad (3.35) \\
S_{-,\parallel} &= \frac{cn_1 k_v^4 p^2}{8} \int_0^{\pi/2} \mathrm{d}\theta_1 \; \sin\theta_1 \; \left\{ \cos\theta_1^2 \left[ |c_{dd,p} - c_{ud,p}|^2 - |c_{du,p} - c_{uu,p}|^2 \right] \right. \\
&\qquad\qquad \left. + |c_{dd,s} + c_{ud,s}|^2 - |c_{du,s} + c_{uu,s}|^2 \right\} \\
&+ \frac{cn_1 k_v^4 p^2}{4} \int_{\pi/2}^{\pi/2 - i\infty} \mathrm{d}\theta_1 \sin\theta_1 i\Im\mathfrak{m}\left[ |\cos\theta_1|^2 \overline{(c_{dd,p} - c_{ud,p})}(c_{du,p} - c_{uu,p}) \right] \\
&\qquad\qquad + i\Im\mathfrak{m}\left[ \overline{(c_{dd,s} + c_{ud,s})}(c_{du,s} + c_{uu,s}) \right] . \qquad (3.36)
\end{aligned}}$$

These formulas were used, for example, to generate the curves in figure 2.18b that show $S_{\text{tot}}$ for a dipole in water close to a silver-coated glass cover slip. All calculations of MIET curves were done with these equations, too, exploiting the fact that for a dipole at an angle $\theta$ relative to the $z$-axis (see the end of section 2.5.3),

$$S_-(\theta) = \cos^2\theta \cdot S_{-,\perp} + \sin^2\theta \cdot S_{-,\parallel} \; . \tag{3.37}$$

As a quick consistency check, also consider the free space case. Then, all reflection coefficients are zero, and out of the four wave types only $c_{dd} \neq 0$. This leads to $S_- = cn_1 k_v^4 p^2/6 = S_0/2$, which is the correct result.
Due to the symmetry of the situation, a reflection of the whole geometry converts the flux $S_+$ towards the upper layers into $S_-$. Thus, $S_+$ can be obtained by exchanging all material parameters in the formulas above.

## 3.5 Lifetime fitting

In order to determine fluorescence lifetimes, we have to fit TCSPC histograms with appropriate model functions. As briefly mentioned in section 2.5.2, the detection system divides the arrival time of photons after an excitation pulse into intervals of equal width, called time bins, histogram bins or simply bins, and counts how many photon arrival times fall into each bin during one measurement. The resulting data set of bins (usually defined by the time points $t_b$ of their centers) and count numbers is called the TCSPC histogram $I_{\text{TCSPC}}(t_b)$. When introducing fluorescence in section 2.5, we stated that the
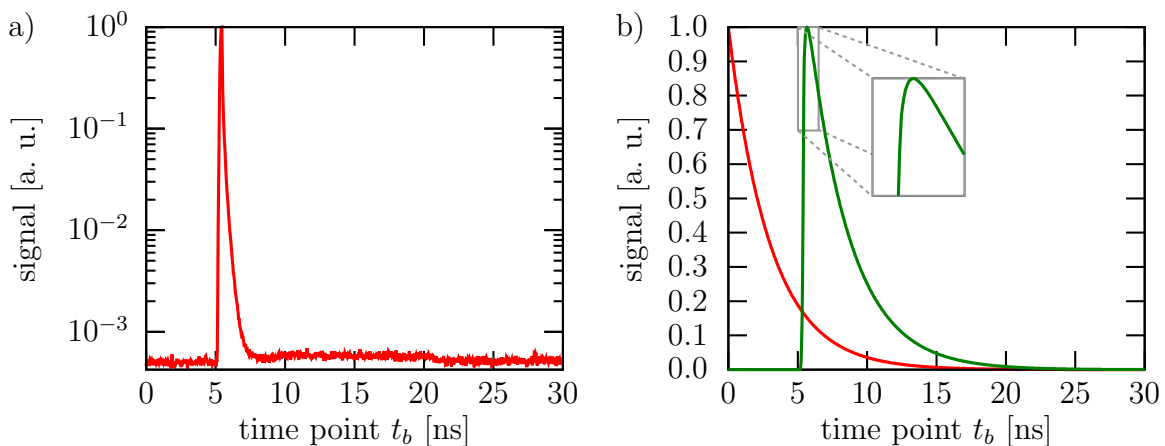


**Figure 3.7:** (a) A typical IRF measured for a scattering sample on a confocal FLIM setup. Note the constant background signal, which has to be subtracted before using the IRF for IRF-fitting. Data provided by Anna Chizhik. (b) A monoexponential decay with lifetime 3 ns, initial function (red) and convolved with the background-corrected IRF (green). The latter corresponds to the ideal TCSPC curve (no noise, i.e. infinite number of photons). The inset shows the peak of the TCSPC curve, which only differs from a shifted version of the initial function in a small number of time points.

intensity measured at the time $t$ after the excitation of an ensemble of fluorophors has the form

$$I(t) \propto \mathrm{e}^{-t/\tau} \tag{3.38}$$

if all fluorophores decay with the same rate $\tau^{-1}$ (monoexponential decay). This is not completely true. The first reason was already mentioned in section 2.5.2: The data is inherently noisy, the measured intensity at time $t$ actually follows a Poisson distribution with expectation value $I(t)$. The second deviation from a perfect exponential function is inevitably caused by the experimental setup itself. Even if all photons would be emitted at exactly the same time $t$ after their excitation, the TCSPC histogram would show a peak with a finite width. This is due to several factors, most notably the pulse width of the excitation laser (which leads to slightly different excitation times of the fluorophores) and the timing response of the detector, which is not exactly the same for each detection event. Furthermore, the time that the light needs to travel from the sample to the detector, and which the electronic signals need to travel through the cabel from the detector to the counting electronics, leads to a shift of the whole curve with respect to the time points of the excitation laser pulses. The measured intensity curve for a sample whose components should emit at exactly the same time $t = 0$ is called the *instrument response function* (IRF). It is often obtained by measuring a scattering sample because scattering is almost instantaneous. Figure 3.7(a) shows a typical IRF measured on a confocal FLIM setup like the one described in section 2.5.2. When a fluorescent sample is measured on the same setup, the recorded TCSPC curve is the convolution of the IRF and the exponential decay of the fluorescence, as demonstrated in figure 3.7(b).

Since a convolution in real space corresponds to a multiplication in Fourier space, it seems straightforward to calculate the Fourier transform of the measured signal, divide by the Fourier transform of the IRF, and perform an inverse Fourier transform to obtain the underlaying exponential decay. However, this approach entails two sources of significant errors. Firstly, if the Fourier transform of the IRF has zeros, information about the initial function has been lost during the convolution and cannot be restored during the deconvolution[7]. Secondly, the measured signal is not just the convolution of the decay and the IRF, but also has an additional noise component. This noise is also present in the Fourier transform of the signal, and can be amplified in unpredictable ways by the division by the Fourier transform of the IRF. For these reasons, a direct deconvolution is usually not performed on microscopy data. Instead, a number of more sophisticated deconvolution techniques exist, both for general and more specific situations. An example of the former is the well-known iterative Richardson-Lucy deconvolution scheme [103, 104].

When fitting TCSPC-curves, one can exploit the *a priori* knowledge that the underlying function is a mono- or multiexponential decay, thus rendering a geneal deconvolution algorithm unnecessary. For our purposes, good results can be obtained with an interative fitting procedure, which we call *IRF-fitting*. This approach is based on measuring the IRF independently, allowing to construct a model function $m(t_b, \boldsymbol{P})$ by numerically

---

[7] Also, from a purely practical point of view, division by zero leads infinite or undefined results in the corresponding pixels, which have to be handled somehow.

convolving a mono- or multiexponential decay with the IRF. By adjusting the decay rates and relative amplitudes of the different decay components, one then tries to find the "best fit" of the model to the data. The main difficulty of this approach lies in obtaining an IRF of high quality, since noise in the measured IRF will be reflected as a noisy model function that does not represent the true expectation value of the intensity curve. Furthermore, the constant background that is always present in a TCSPC measurement (see figure 3.7(a)) has to be subtracted from the IRF, which can be challenging if the background is so small that its magnitude cannot be determined precisely[8].

Alternatively, one can avoid the need for a deconvolution by emploing *tail-fitting* procedures. These rely on the assumption that the IRF is a rather sharp peak, and that the convolution of an exponential function with the IRF only differs from the pure exponential function up to the first few histogram bins after the maximum (see inset of figure 3.7(b)). Then, one can simply discard all bins that are still influenced by the IRF, and only fit the tail of the TCSPC curve. In this case, we benefit from a characteristic of the exponential function, namely that the average arrival time after a certain time point is independent of this time point:

$$
\begin{aligned}
\langle t \rangle_{[0,\infty]} &= \int_0^\infty \mathrm{d}t \; t \cdot \frac{1}{\tau} \cdot \mathrm{e}^{-t/\tau} \\
&= \tau \cdot \frac{\mathrm{d}}{\mathrm{d}\tau} \int_0^\infty \mathrm{d}t \; \mathrm{e}^{-t/\tau} \\
&= \tau \cdot \frac{\mathrm{d}}{\mathrm{d}\tau} \left[ 0 + \tau \right] \\
&= \tau, \\
\langle t - t_0 \rangle_{[t_0,\infty]} &= \int_{t_0}^\infty \mathrm{d}t \; (t - t_0) \cdot \frac{1}{\tau} \cdot \mathrm{e}^{-(t-t_0)/\tau} \\
&= \int_0^\infty \mathrm{d}z \; z \cdot \frac{1}{\tau} \cdot \mathrm{e}^{-z/\tau} \\
&= \tau.
\end{aligned}
\tag{3.39}
$$

Thus, in principle, one could simply choose a cutoff $t_0$, re-define it as time zero, and calculate the average arrival time relative to $t_0$ for all photons in histogram bins after the cutoff. However, here the same caveat as for the IRF-fit applies: When calculating the average arrival time, a constant background would shift this average towards the middle of the measurement interval, and thus has to be subtracted beforehand. But again, estimating the background e.g. from pixels without sample can be too imprecise, especially when the background is spatially varying. Therefore, the most common procedure is to define a model function $m(t_b; \boldsymbol{P})$, usually a mono- or multiexponential decay, and to vary the parameters $\boldsymbol{P}$ (decay constants, relative amplitudes, amplitude of the constant background) until the "best fit" of the model to the data is reached.

---

[8] Keep in mind that any TCSPC curve is a histogram whose values in different bins are integers. Thus, a low background results in bins with zeros or ones at later time points (after the decay of the fluorescence), who have to be averaged over a number of bins to yield an estimate of the true background level.

The main difference to the IRF-fitting is that the model function now only describes histogram bins $t_b$ after the cutoff $t_0$, and can therefore be a simple linear combination of exponential functions and the constant background.

For both IRF-fitting and tail-fitting, we defined the goal of finding the "best fit" between model function and measured data by varying the model parameters $\boldsymbol{P}$. This objective is also the purpose of regression analysis, a branch of statistical modeling. In regression analysis, several definitions of what might be a "best fit" exist. One possibility is to minimize the sum of squared residuals (SSR), where a residual is the difference between one data point and the corresponding value of the model. This procedure is also called a *least squares* fit. When modeling a TCSPC curve, the SSR corresponds to:

$$\text{SSR}(\boldsymbol{P}; I_{\text{TCSPC}}) = \sum_{b=1}^{B} \left[ I_{\text{TCSPC}}(t_b) - m(t_b; \boldsymbol{P}) \right]^2, \tag{3.40}$$

where the sum is taken over all histogram bins $t_b \in [0, T]$ for an IRF-fit and only over the bins $t_b \in [t_0, T]$ for the tail-fit. The estimate $\hat{\boldsymbol{P}}$ of the true underlying parameter vector $\boldsymbol{P}$ is then given as

$$\hat{\boldsymbol{P}} = \arg\min_{\boldsymbol{P}} \left\{ \text{SSR}(\boldsymbol{P}; I_{\text{TCSPC}}) \right\}. \tag{3.41}$$

The least squares fit implicitly assumes that all data points have the same uncertainty, and therefore "punishes" the deviation of any data point from the model quadratically. Thus, outliers can strongly change the resulting fit parameters. However, if it is known that some data points have a higher uncertainty than others, their deviation from the model function should have less impact on the parameter choice than the deviation of very accurate data points from the model. This is taken into account in weighted least squares formulations, which include the variance $\sigma^2(t_b)$ of each data point:

$$\text{WSSR}(\boldsymbol{P}; I_{\text{TCSPC}}) = \sum_{b=1}^{B} \frac{\left[ I_{\text{TCSPC}}(t_b) - m(t_b; \boldsymbol{P}) \right]^2}{\sigma^2(t_b)}. \tag{3.42}$$

In order to estimate the variance of the count numbers in each TCSPC bin, we consider the underlying measurement process. The histogram is acquired by summing the photon arrival times of many excitation cycles. Within each excitation cycle, the probability to detect a photon is proportional to the model function $m(t_b; \boldsymbol{P})$, for example a monoexponential decay. When considering the complete experiment with its hundreds of thousands of excitation cycles, we can reasonably assume that for each time bin $t_b$, the average rate of photons being sorted into this bin is constant, and proportional to $m(t_b; \boldsymbol{P})$. Furthermore, since the photon emission events are independent of each other, the probability of subsequent photons being sorted into bin $t_b$ is independent of the number of photons already sorted into this bin. Thus, the total number of photons $I_{\text{TCSPC}}(t_b)$ sorted into bin $t_b$ during the experiment is Poisson distributed with mean

$m(t_b)$. Since for a Poisson-distributed random variable the variance equals the mean, we find:

$$\text{WSSR}(\boldsymbol{P}; I_{\text{TCSPC}}) = \sum_{b=1}^{B} \frac{[I_{\text{TCSPC}}(t_b) - m(t_b; \boldsymbol{P})]^2}{m(t_b; \boldsymbol{P})}. \tag{3.43}$$

Another approach to find the "best fit" of the model curve to the measured data is to find the model function that maximizes the probability of measuring exactly the data that was found in the experiment. This probability is called the *likelihood* $\mathcal{L}(\boldsymbol{P}; I_{\text{TCSPC}})$ of the parameter set $\boldsymbol{P}$, and the thus obtained parameter vector a *maximum likelihood estimator*. The likelihood is determined as follows: Since we assume that the number of photons sorted into bin $t_b$ follows the Poisson distribution with expectation value $m(t_b)$,

$$p(I_{\text{TCSPC}}(t_b); \boldsymbol{P}) = \mathrm{e}^{-m(t_p; \boldsymbol{P})} \cdot \frac{m(t_b; \boldsymbol{P})^{I_{\text{TCSPC}}(t_b)}}{I_{\text{TCSPC}}(t_b)}, \tag{3.44}$$

the probability to obtain a specific TCSPC histogram for a given total number of photons is

$$\mathcal{L}(\boldsymbol{P}; I_{\text{TCSPC}}) = \prod_{b=1}^{B} p(I_{\text{TCSPC}}(t_b); \boldsymbol{P}). \tag{3.45}$$

Instead of directly maximizing $\mathcal{L}(\boldsymbol{P}; I_{\text{TCSPC}})$ with respect to $\boldsymbol{P}$, it is numerically easier to minimize $-\log \mathcal{L}$, called the *negative log-likelihood*, to find an estimate $\hat{\boldsymbol{P}}$ of the parameter vector $\boldsymbol{P}$:

$$\hat{\boldsymbol{P}} = \underset{\boldsymbol{P}}{\arg\min} \left\{ -\log \mathcal{L}(\boldsymbol{P}; I_{\text{TCSPC}}) \right\}$$

$$= \underset{\boldsymbol{P}}{\arg\min} \left\{ \sum_{b=1}^{B} m(t_b; \boldsymbol{P}) - \sum_{b=1}^{B} I_{\text{TCSPC}}(t_b) \cdot \log[m(t_b; \boldsymbol{P})] \right\}. \tag{3.46}$$

If the number of photons in one bin is very large, since this number is Poisson distributed, the signal-to-noise ratio (SNR) of this bin is also large:

$$\text{SNR}(t_b) = \frac{\text{signal}}{\text{standard deviation}} = \frac{I_{\text{TCSPC}}(t_b)}{\sqrt{I_{\text{TCSPC}}(t_b)}} = \sqrt{I_{\text{TCSPC}}(t_b)}. \tag{3.47}$$

Thus, with an increasing total number of detected photons, the measured TCSPC curve is less noisy and approaches the ideal model function more closely. For this case, it is known that the parameter vector estimates $\hat{\boldsymbol{P}}$ obtained by least squares or maximum likelihood are asymptotically the same (see e.g. [105]). For smaller sample sizes, maximum likelihood estimators have been shown to be more accurate [105]. However, least squares solvers are implemented in many commercial programs such as Matab [106], and are still often used.

When evaluating fluorescence lifetimes for MIET experiments, the choice of the fitting routine depends on the sample. If the expected lifetimes are very short, most of the fluorescence photons arrive in the first few bins. Therefore, ignoring the beginning of

the curve results in a significant loss of signal, which is why an IRF-fit is preferred over a tail-fit. The same is true when the IRF is very broad, such that a large cutoff $t_0$ would have to be chosen. The experimental setup used for the applications presented in this work has a very sharp IRF, and fluorescence lifetimes were usually longer than one nanosecond. Therefore, we avoided the potential complications arising from measuring, smoothing and background-correcting an IRF by only employing tail-fits.

If it is known that the fluorescence signal follows a monoexponential decay (or else has exactly $N$ lifetime components), a maximum likelihood fit with the corresponding model function is assumed to give the best result, as mentioned above. However, for most MIET measurements, several fluorophores at different $z$-positions are contained within the focal volume. Since the fluorescence lifetime is $z$-dependent, the measured TCSPC curve is a linear combination of an unknown number of exponential decays. In this situation, our group is routinely using a method that we call a *distributed tail-fit* (DTF), and which has proven to be quite robust in extracting the average lifetime from this type of curve. The distributed tail-fit is based on a model function $m_{\mathrm{DTF}}(t_b)$ that is a sum of a large number (typically $N = 200$) of monoexponential decay curves and a constant background, where the decay rates $\tau_j^{-1}$ are fixed and only the relative amplitudes $a_j$ are fit parameters ($B$ = number of TCSPC bins):

$$m_{\mathrm{DTF}}(t_b) = \frac{a_0}{B} + \sum_{j=1}^{N} a_j \frac{1}{\tau_j} e^{-t_b/\tau_j}, \quad \boldsymbol{P} = \{a_j | j = 0, \cdots, N; \ a_j \geq 0 \ \forall j\}. \tag{3.48}$$

This model function is then used in a least-squares approximation. The important point to notice here is that all amplitudes are either zero or positive but cannot become negative, reducing the risk of overfitting. Thus, this approach can be interpreted as decomposing the measured signal into the relevant basis functions, which correspond to different fluorescence decay components.

The optimization problem can be written in matrix notation using the vector $\boldsymbol{t} = \{t_b\}$ of the time points of the TCSPC bins, the data vector $\boldsymbol{d} = \{I_{\mathrm{TCSPC}}(t_b)\}$, the vector $\boldsymbol{a} = \{a_j\}$ of the amplitudes and defining the model matrix $\underline{\underline{M}}$

$$\underline{\underline{M}} = \left( \begin{array}{ccccc} \frac{\mathbf{1}}{B} & \frac{\exp(-\boldsymbol{t}/\tau_1)}{\tau_1} & \frac{\exp(-\boldsymbol{t}/\tau_2)}{\tau_2} & \cdots & \frac{\exp(-\boldsymbol{t}/\tau_N)}{\tau_N} \end{array} \right). \tag{3.49}$$

Then, the estimator $\hat{\boldsymbol{a}}$ of the amplitudes is given as

$$\hat{\boldsymbol{a}}^{\mathsf{T}} = \underset{\boldsymbol{a}^{\mathsf{T}}}{\arg\min} \, ||\underline{\underline{M}} \cdot \boldsymbol{a}^{\mathsf{T}} - \boldsymbol{d}||_2^2 \quad \text{where } a_j \geq 0 \ \forall j. \tag{3.50}$$

Exactly this fitting problem is implemented in many numerical computing environments, e.g. as the function $\boldsymbol{a} = \texttt{lsqnonneg}(\underline{\underline{M}}, \boldsymbol{d})$ in Matlab, which we use in all lifetime fits presented in this work. Once an estimator $\hat{\boldsymbol{a}}$ of the different amplitudes has been found, this can be utilized to estimate the average lifetime $\langle \tau \rangle$ of all fluorophores contributing to the TCSPC curve:

$$\langle \hat{\tau} \rangle = \frac{\sum_{j=1}^{N} \hat{a}_j \cdot \tau_j}{\sum_{j=1}^{N} \hat{a}_j}. \tag{3.51}$$

With this last step, we acknowledge that the individual fluorescence lifetimes of all fluorophores within the focal volume generally cannot be distinguished by this (or any other) method. Instead, we assume that when the average lifetime $\langle \tau \rangle$ is converted to a height using a MIET lifetime-versus-height curve, the resulting height corresponds approximately to the average height of the fluorophores. Some discrepancy is expected because the brightness of a fluorophore is also distance-dependent and thus influences the amplitude $a_j$ of the corresponding lifetime component $\tau_j$, however, the deviation is small, as we will now demonstrate for an example.

We start by simulating a TCSPC curve generated by a single fluorophore whose decay curve is monoexponential with a lifetime of 1 ns or 3 ns, where we take into account the Poisson-distribution of the count numbers of the single TCSPC bins. As figure 3.8(a, b) shows, the lifetime estimated using the distributed tail-fit converges to the correct value as the total number of photons is increased, accompanied by a decrease of the variance of the lifetime estimate. There is a small bias towards larger lifetimes at small photon numbers. For 3000 photons, the minimum number required for our MIET experiments (pixels with less counts are ignored in the evaluation), the bias is less than 3 %. In the next step, we simulate the TCSPC curve generated by two fluorophores of equal brightness and with fluorescence lifetimes $\tau_1 = 1$ ns and $\tau_2 = 3$ ns, see figure 3.8(c). Again, the estimated average lifetime approaches the correct value $\langle \tau \rangle = 2$ ns for large photon numbers.

Now, we move on to a more realistic situation. We model the experiment conducted in [38], where a microtubule (diameter 25 nm) was uniformly labeled with the fluorescent dye Alexa Fluor 488 (emission maximum $\lambda = 525$ nm, quantum yield $\Phi = 0.8$, free space lifetime $\tau_0 = 3.5$ ns), and suspended over a gold-coated glass cover slip (1 nm titanium as an adhesion layer, 15 nm gold) using various spacer molecules. We consider the case that the microtubules are covered by neutravidin and immobilized on the surface by avidin, a situation for which the average distance between the lowest point of the microtubule and the top of the gold layer was found to be $13.3 \pm 1.5$ nm. As depicted in figure 3.9(a), we determine the height $h_i$ of 40 dye molecules evenly spaced on the surface of the microtubule. Using the MIET calibration curve determined for the sample parameters, we convert these height values into 40 lifetime values $\tau_i$ (b). The corresponding TCSPC-curves (c) are then added up (d), where we neglect the height-dependent brightness of the fluorophores, attributing the same total number of photons to each curve. The resulting total TCSPC histogram is then passed to our distributed tail-fit, resulting in an estimate of the average lifetime $\langle \hat{\tau} \rangle$. When directly averaging the 40 "true" lifetime values $\tau_i$, we find $\langle \tau \rangle = 0.8629$ ns, which is identical to the estime $\langle \hat{\tau} \rangle$ obtained for $10^7$ photons. The mean and standard deviation of 200 simulations per photon number are depicted in figure 3.9(e), here the bias for small photon numbers is less than 1 %. The average height estimated using this procedure shows a similar convergence behaviour (not shown), approaching the value corresponding to the average lifetime $\langle \tau \rangle$, $\langle \hat{h} \rangle = 26.9$ nm. The true average height of the fluorophores is $(13.3 + 25/2)$ nm $= 25.8$ nm. The (small) discrepancy between these values can be explained by the fact that the MIET curve is not a perfect straight line. For this reason, Berndt *et al.* – knowing the diameter of the microtubules – modeled the complete lifetime distribution to find the distance between the bottom of the microtubules and the top of the gold surface [38]. In situations where the labeled structure is not as
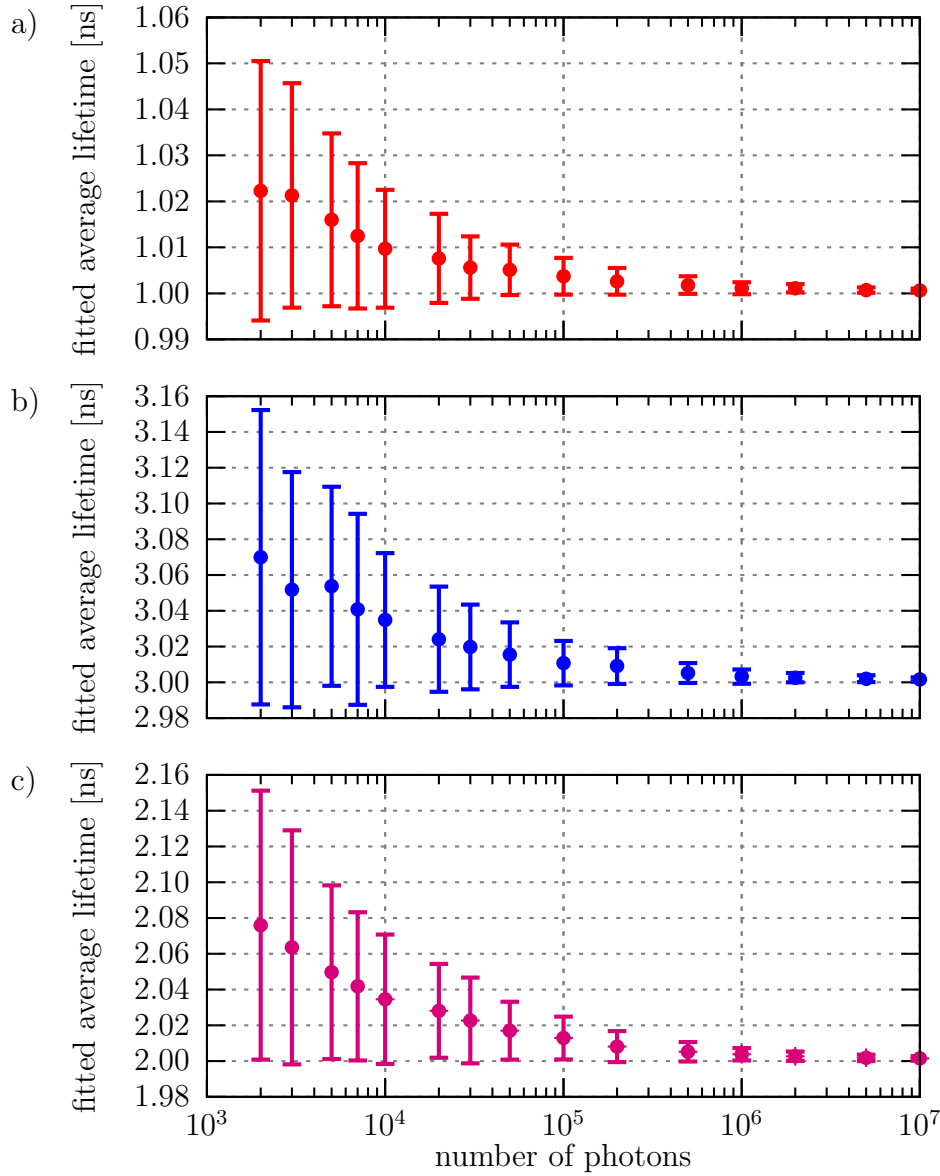
**Figure 3.8:** Application of the distributed tail-fit (DTF) to simulated mono- (a, b) or biexponential (c) noisy data. The ideal model function $m(t_b)$ for TCSPC time bins of width 32 ps and a total length of the TCSPC curve of 50 ns (corresponding to a pulse repetition rate of 20 MHz) is a monoexponential decay with lifetime $\tau = 1$ ns (a) or $\tau = 3$ ns (b), or a linear superposition of both with relative amplitudes 0.5 (c). By adjusting the total amplitude of the model function in the simulation, the number of detected photons for the whole curve (horizontal axis) could be chosen. For each time bin $t_b$, the ideal model $m(t_b)$ was then taken as the expectation value of a Poisson distribution to generate a noisy curve. Subsequent evaluation of the thus generated data by the DTF yielded an estimate of the average lifetime $\langle \tau \rangle$. The mean of 200 repetitions of the data generation and evaluation is shown as filled circles in the graph, the errorbars denote the standard deviation. Note that the estimated average lifetime approaches the true average lifetime for large photon numbers, with a small bias towards larger values at small photon numbers.
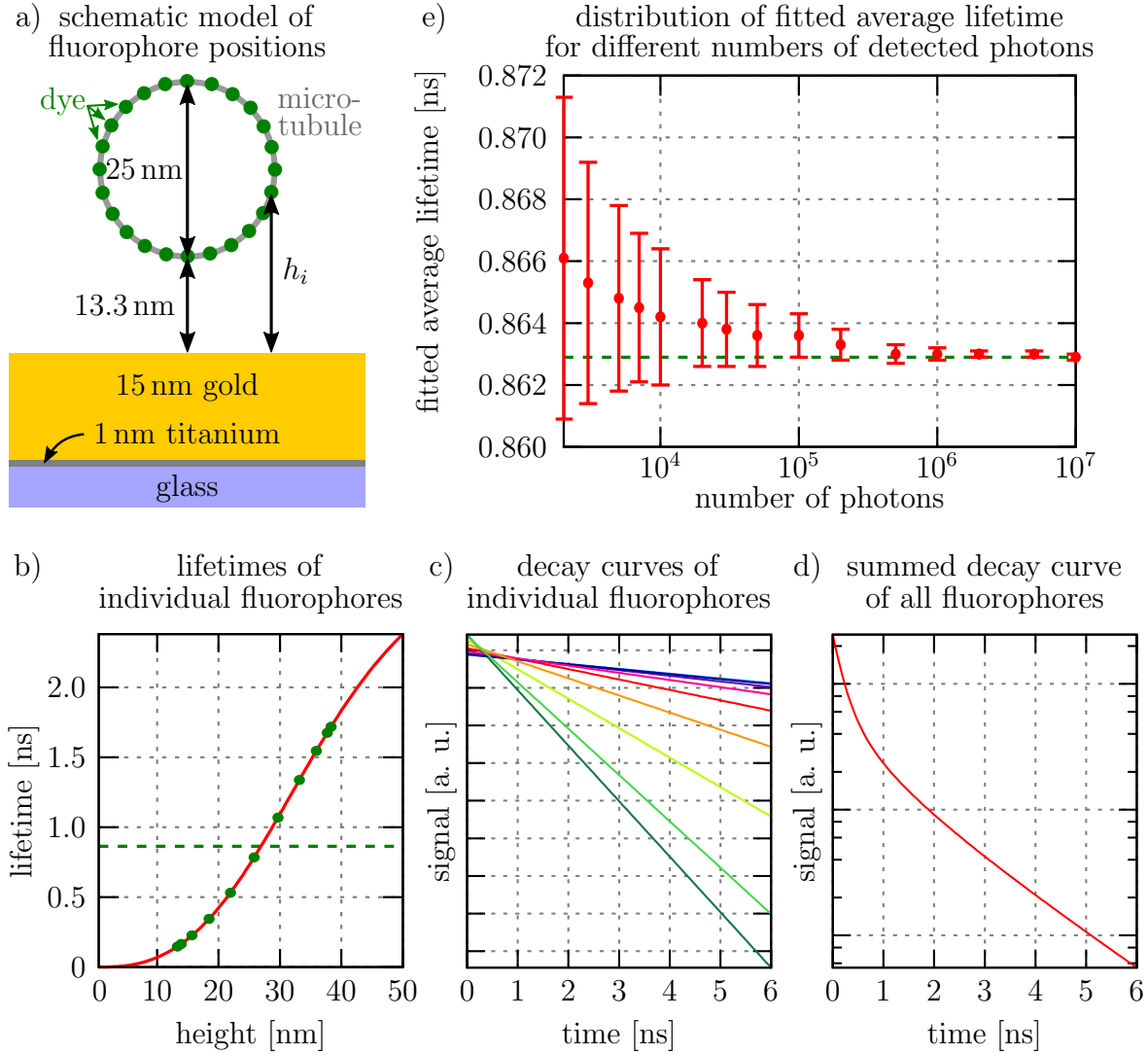
**Figure 3.9:** Application of the distributed tail-fit (DTF) to simulated realistic MIET data. (a) Schematic of the modeled situation: A microtubule (grey circle, ⌀25 nm) is homogeneously labeled with the dye Alexa Fluor 488 (green dots). The bottom of the microtubule is elevated 13.3 nm above the MIET substrate by avidin and neutravidin. (b) MIET curve calculated using the optical parameters of the MIET substrate, the buffer surrounding the microtubule and the dye. The height-lifetime pairs $(h_j, \tau_j)$ of some of the dye molecules are indicated as green dots, the dotted green line corresponds to the average lifetime of all dye molecules, 0.863 ns. (c) Semilogarithmic plot of the monoexponential decay curves of some of the individual fluorophores. (d) Semilogarithmic plot of the sum of all the decay curves, corresponding to the ideal TCSPC curve one would measure in this situation if there was no noise (i.e. infinite number of detected photons). (e) When evaluating noisy TCSPC curves – generated from the curve in (d) by using only a finite number of photons – with the DTF, an estimator of the average lifetime $\langle \tau \rangle$ is obtained. The mean and standard deviation of 200 repetitions of this procedure are shown as circles and errorbars, respectively. The dotted green line is again the true average lifetime, 0.863 ns. Note that here, the bias at 2000 photons is 3 ‰ (or 2 ‰ at 3000 photons).

well-defined, the average lifetime and the derived height value offer good descriptions of the sample. In particular, when relatively thin structures such as lipid membranes are labeled, the different lifetime components are very similar, and thus the averaging leads to a height value that approximates the true average height even more closely.

# 4 Applications

Due to the properties of MIET discussed in the previous section, the method offers an opportunity to study systems where the positions of one or several structures within the first 200 nm from the surface are of interest. In this chapter, three such examples are presented: focal adhesion complexes in human mesenchymal stem cells, the nuclear membrane of HeLa cells, and the behaviour of the basal cell membrane in the epithelial-to-mesenchymal transition of NMuMG cells. In all three cases, the sample preparation and all measurements were performed by my colleagues as indicated in the respective sections; my contributions were the extraction of lifetime values from the raw data, their conversion to height maps, as well as a variety of statistical evaluations. The data evaluation was greatly facilitated by a flexible software package which was developed as part of this work as described in section 4.1.1.

Besides demonstrating the versatility of MIET in sections 4.1.2 to 4.1.4, we also present an investigation of its accuracy and the parameters influencing it in section 4.2. These findings provide a valuable basis when appraising height values obtained from MIET measurements.

Based on the profound knowledge of the interaction of fluorophores with metal layers presented in sections 2.3 to 2.5, the challenging determination of absolute quantum yield values has been accomplished. When measuring lifetimes close to metal layers, out of the five parameters free space lifetime $\tau_0$, free space quantum yield $\Phi_0$, orientation $\theta$, lifetime $\tau$ and height $z$, four have to be known in order to determine the fifth. In MIET experiments, we presume the knowledge of $\tau_0$, $\Phi_0$ and $\theta$, such that a measurement of $\tau$ yields $z$. Conversely, if $z$ is known – for example when molecules are deposited on a silica spacer of known thickness – and $\tau$ is measured, the quantum yield could theoretically be determined by a single measurement. In practice, however, every lifetime measurement suffers from noise. Therefore, the accuracy of the determined value for $\Phi$ can be improved by doing several such measurements at different spacer thicknesses and fitting for $\Phi$. As a drawback, this approach is very time-intensive. Furthermore, it relies on the fact that the molecules are located directly on the silica/air interface, and is thus not suited for measurements of quantum yields in solution. A more sophisticated technique using a nanocavity was developed by our group in 2013 [107] and further improved in this work. It is presented in section 4.3.

Due to the large scope of this work, not all projects and publications can be presented in the main text. Some further examples of MIET measurements, as well as an application of the theory describing the interaction of light with spherical particles, are therefore summarized in section 6.3 in the appendix.

# 4.1 Application of MIET to biological samples

Following the proof-of-principle study [38] described at the end of section 2.5.4, MIET was applied to living cells in 2013/2014 [39]. Cell-substrate distance maps were obtained for several cell lines, showing different average distances for benign and more motile, cancerous cells. Furthermore, the three phases of cell spreading of one cell type could be observed in a MIET time series. This publication proved MIET microscopy to be a versatile tool to answer relevant biological questions.

However, at this stage, the MIET data analysis was still performed using a bundle of custom-written software which required a lot of manual adjustment and a deep understanding of the underlying physical concepts. In order to make the technique more accessible for general users, an easy-to-adopt software package was compiled as part of this work, providing a graphical user interface (GUI) that allows to evaluate MIET data using Matlab without any knowledge of programming. This program is described in section 4.1.1. The software was used to evaluate several different experiments, and was in turn constantly improved. The application of MIET imaging to structures within cells posed new challenges, which will be adressed in the context of these experiments in sections 4.1.2 to 4.1.4.

## 4.1.1 MIET analysis software

After the first applications of MIET microscopy had been published, the further development and establishment of the technique quickly required a well-maintained software package implementing the latest analysis methods. Thus, a major part of this work was to improve, extend and re-implement previous proof-of-principle methods based on the theoretical concepts introduced above into a reliable and consistent toolchain. To simplify and document the data analysis steps for general users, the software was equipped with a graphical user interface (GUI). Since this is the part of the program which primarily determines the user experience, we call the whole software package *MIET GUI*. It has been described in detail in [7], and a free version is available online from the institute's homepage.

In general, the data analysis of a MIET experiment consists of three steps. Firstly, lifetime values are extracted from the raw data. Secondly, a MIET lifetime-versus-height calibration curve is calculated based on the sample parameters. Thirdly, the lifetimes are converted to height values based on the MIET curve. The implementation of the first step varies strongly based on the FLIM setup used to acquire the data, starting from the fundamental differences between frequency- or time-domain measurements, but also including the very specific data formats used by different manufacturers. Commercial FLIM setups usually include a proprietary software which performs this analysis step, while users who tailor a FLIM setup to a specific purpose often like to have full control over the employed fitting algorithms. In an internal version of the software package, we included the full lifetime fitting process: Starting with the extraction of photon arrival times from the binary data produced by the counting electronics, TCSPC histograms are compiled and dead-time corrected as described in section 6.3.2 in the appendix, and finally fitted using IRF- or tail-fits as described in section 3.5. However, this module was not included in the published version for the aforementioned reasons.
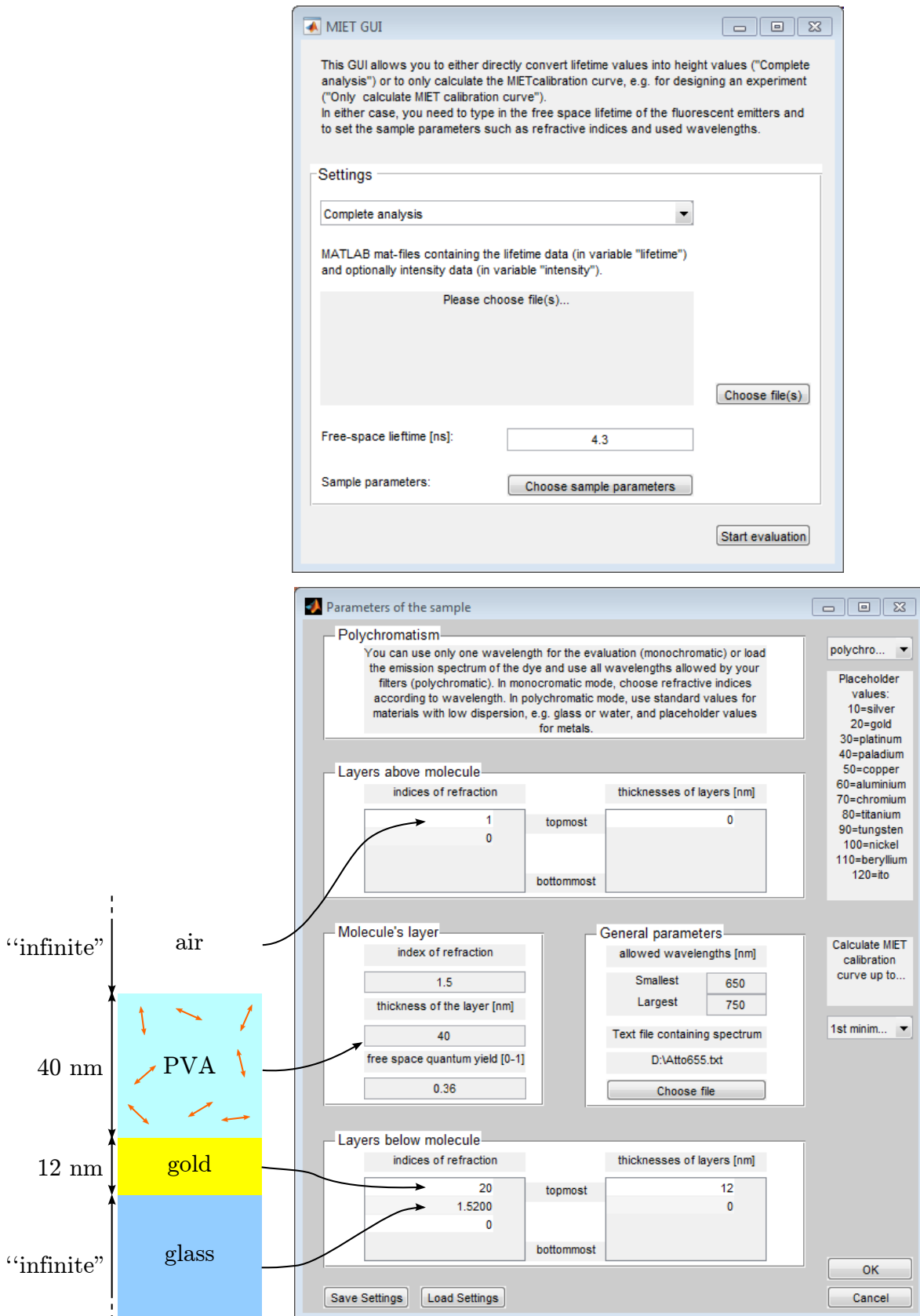
**Figure 4.1:** Screenshots of the MIET graphical user interface. **Top:** Main window, choice between conversion of lifetimes to height values (*Complete analysis*) or only calculation of MIET curve; specification of lifetime data. **Bottom:** Parameters window, opened by button *Choose sample parameters*, defines stratified system and optical properties of fluorophore. In the schematic, orange arrows represent fluorophores.

# Applications

A screenshot of the published version of the MIET GUI is shown in figure 4.1. It allows the user to specify the two important sets of sample parameters – the refractive indices and thicknesses of the materials making up the stratified system, as well as the optical properties of the fluorophore – which are needed to calculate a MIET calibration curve, and to specify the lifetime values that are to be converted into heights.

The software simplifies the workflow in several ways. To name just two examples, the GUI offers the option to only output the MIET curve[1], or to process an arbitrary number of data sets from the same sample type (i.e. with the same MIET curve) in bulk. The representation of the stratified system as three stacks of layers – one layer containing the fluorophore, with one stack above and one below it, see lower half of figure 4.1 – is capable of mapping virtually any sample that has a planar geometry. This is in accordance with our aim of creating a general-purpose MIET analysis software that is not restricted to a specific sample type. The only assumption made by the algorithm is that the bottommost medium of the lower stack (usually the glass cover slip) and the uppermost medium of the upper stack (usually buffer in a cell sample, or air when molecules are deposited on a silica spacer) are assumed to be infinitely thick.

In the simplest case, a MIET curve is calculated for one single wavelength $\lambda$, usually the maximum emission wavelength of the fluorophore. In reality, however, emitters fluoresce over a range of wavelengths, which all have slightly different MIET curves. This is taken into account if the free space fluorescence spectrum $F(\lambda)$ is specified:

$$\langle \tau(z) \rangle_\lambda = \int_{\lambda_{\min}}^{\lambda_{\max}} \tau(z, \lambda) \cdot F(\lambda) \, \mathrm{d}\lambda \; \Big/ \int_{\lambda_{\min}}^{\lambda_{\max}} F(\lambda) \, \mathrm{d}\lambda. \tag{4.1}$$

When the actual evaluation is initiated, the program first calculates the MIET curve. In principle, the total energy $S_{\text{tot}}$ emitted per time by one emitter depends not only on the $z$-position of the emitter but also on its orientation $\theta$ with respect to the optical axis. When denoting the values for fluorophores with a vertical or a horizontal dipole moment as $S_{\text{tot},\perp}(z)$ and $S_{\text{tot},\|}(z)$, respectively, the linearity of the relevant equations leads to the simple relation:

$$S_{\text{tot}}(z, \theta) = S_{\text{tot},\perp}(z) \cdot \cos^2 \theta + S_{\text{tot},\|}(z) \cdot \sin^2 \theta. \tag{4.2}$$

In our evaluations, we assume an ensemble of randomly oriented fluorophores, and thus set $S_{\text{tot}}(z) = [S_{\text{tot},\perp}(z) + 2 S_{\text{tot},\|}(z)]/3$[2]. This value is then used to determine the MIET curve according to equation (2.222): $\tau(z) = \tau_0 \cdot S_{\text{tot}}(z)/S_0$. Since this function oscillates with decreasing amplitude at large $z$, an unambiguous relation between $\tau$ and $z$ can only be found up to the $z$-value of the first minimum, $z_{\min}$. If some prior knowledge makes it unlikely that fluorophores are at high $z$-values, the user can alternatively choose to use the curve up to the $z$-value of the first maximum, $z_{\max}$. All lifetimes higher than $\tau(z_{\min/\max})$ are marked as "not a number" when transforming lifetime values into heights.

Finally, the resulting height images are automatically plotted and the raw data is saved.

---

[1] Which can be used to estimate the impact of certain sample parameters on the obtainable height resolution when designing an experiment.

[2] If it is known that the fluorophores have a specific orientation, the code can be adapted accordingly.

## 4.1.2 Cell-substrate dynamics of the epithelial-to-mesenchymal transition

The first of three MIET projects presented in this work (which all employed the software package presented above) concerned cell-substrate measurements. The results, and all figures shown in this section, were published in:

[1] T. Baronsky, **D. Ruhlandt**, B. R. Brückner, J. Schäfer, N. Karedla, S. Isbaner, D. Hähnel, I. Gregor, J. Enderlein, A. Janshoff, and A. I. Chizhik, "Cell-substrate dynamics of the epithelial-to-mesenchymal transition," *Nano Letters*, vol. 17, no. 5, pp. 3320–3326, 2017. T. B., **D. R.** and B. R. B. contributed equally to this work. **D. R.** provided analysis tools and analysed the data.

**Motivation**

Epithelial tissues line the outer surfaces of blood vessels and organs throughout the body, where one of their main tasks is to provide a barrier between different compartments of the body and to protect against physical trauma. This is achieved by very tight cell-cell adhesions (also called *anchoring junctions*), where actin fibers – relatively stiff structures that are a part of the cytoskeleton, which gives the cell its shape and resistance to mechanical transformation – are anchored. These, in turn, transmit stresses across the interior of each cell, distributing them over larger areas of the tissue. More details can be found in textbooks such as [108] (chapter 19), or directly in our publication [1]. Under certain conditions, for example as a component of wound healing but also in cancer progression, epithelial cells can transform to more mobile mesenchymal cells. This process, called epithelial-to-mesenchymal transition (EMT), can be triggered externally by adding certain growth factors to epithelial cells. During the EMT, cells undergo a variety of changes, including a loss of cell-cell adhesions, actin reoganization, and an altered cell-substrate interaction, that allow them greater mobility. Further details on the mechanism, as well as its biological relevance, are given in [1]. It was our aim to observe one of the mentioned effects, the altered cell-substrate interaction, using MIET microscopy, and to draw conclusions regarding possible stages of the EMT from these measurements.

The temporal progression of the EMT had already been studied in a previous work [109] by Schneider *et al.* employing electric cell-substrate impedance sensing (ECIS). In ECIS, a confluent cell layer is grown on an electrode and the electrical impedance measured for different alternating current frequencies. By fitting these frequency sweeps with appropriate models, conclusions about cell-cell contacts, cell shape and the impedance in the cleft between the cells and the substrate (which is linked to the cell-substrate distance) can be drawn. Furthermore, an analysis of the temporal fluctuations of the impedance provides information about collective vertical motility of the cells. A repetition of these experiments at several time points after addition of TGF-$\beta$1 revealed increased impedance fluctuations and thus a higher micromotility that started about one to three hours after induction of the EMT and lasted for roughly five to seven hours. This time span was denoted *transitional state I*, the elevated variance of the fluctuations was attributed to membrane ruffling. About eight hours after addition of

TGF-$\beta$1, the impedance fluctuations reduced to the value of untreated cells. However, fits of the frequency sweeps indicated a gradual decrease of cell-cell contacts after ten to thirty hours, accompanied by the formation of some actin stress fibers and stronger adhesion of the cells to the substrate. This interval was given the name *transitional state II*. Afterwards, a *mesenchymal-like state* was reached where NMuMG cells further changed their morphology and developed more stress fibers. Forty-eight hours after addition of TGF-$\beta$1, the EMT was complete.

The study by Schneider *et al.* formed the basis for our MIET measurements. While they already provided a comprehensive image of the EMT, some aspects still required further investigation. Firstly, ECIS is an ensemble measurement: In [109], for each impedance measurement 100,000 NMuMG cells were seeded on a circular gold electrode with a diameter of 250 µm. Thus, a precise picture of the different stages on the cellular or even subcellular level was lacking due to the limited spatial resolution. Secondly, fits of ECIS frequency sweeps only provide a quantity that depends both on the cell-substrate distance $z$ and the resistivity of the medium between cells and substrate. The latter, however, may change due to altered ion concentrations in the different stages of the EMT. In contrast, MIET is *de facto* independent of the ion concentration. Thus, accurate height values could not be obtained in [109]. It was therefore our goal to investigate the evolution of cell-substrate distances during the different EMT stages both on a single-cell level and regarding average values.

## Methods

As a first step, a protocol to induce EMT in the mouse mammary gland epithelial cell line NMuMG was established. Figure 4.2 demonstrates that upon addition of the transforming growth factor-$\beta$1 (TGF-$\beta$1) to NMuMG cells, the cells exhibit several hallmarks of an EMT. In particular, the phase contrast images in figure 4.2(a) and (b) show a transition of the general morphology from compact cells forming a tight epithelial layer to individual, elongated, mesenchymal-like cells. Fluorescence images of F-actin filaments labeled with the fluorescent dye Alexa Fluor 546 (Life Technologies, Darmstadt, Germany) in (c) and (d) exhibit a remodeling of the F-actin network to well-defined stress fibers, which traverse the entire cell. Finally, another typical property of the EMT is the downregulation of E-cadherin, a protein that is essential for junctions between neighbouring cells. This process can be observed in the fluorescence images in figure 4.2(e) and (f), where E-cadherin was labeled with Alexa Fluor 488. These three observations, together with some further checks presented in the supplementary material of [1], proved a successful induction of the EMT in NMuMG cells.

With this, we moved on to conduct lifetime measurements on NMuMG cells in the epithelial state or during and after the EMT. The setup we used is shown schematically in figure 4.3, it is basically the same as already presented in figure 2.25. The only specialty is an incubation chamber which keeps the cells at 37 °C, allowing measurements of living cells for up to 1.5 hours. Details on the specifications of the used optical components can be found in the supporting information of [1].

The plasma membrane of the cells was stained using the fluorophore CellMask Deep Red (Life Technologies), which has its emission maximum around 680 nm. We checked that effectively no dye was internalized by the cell by acquiring confocal images of a
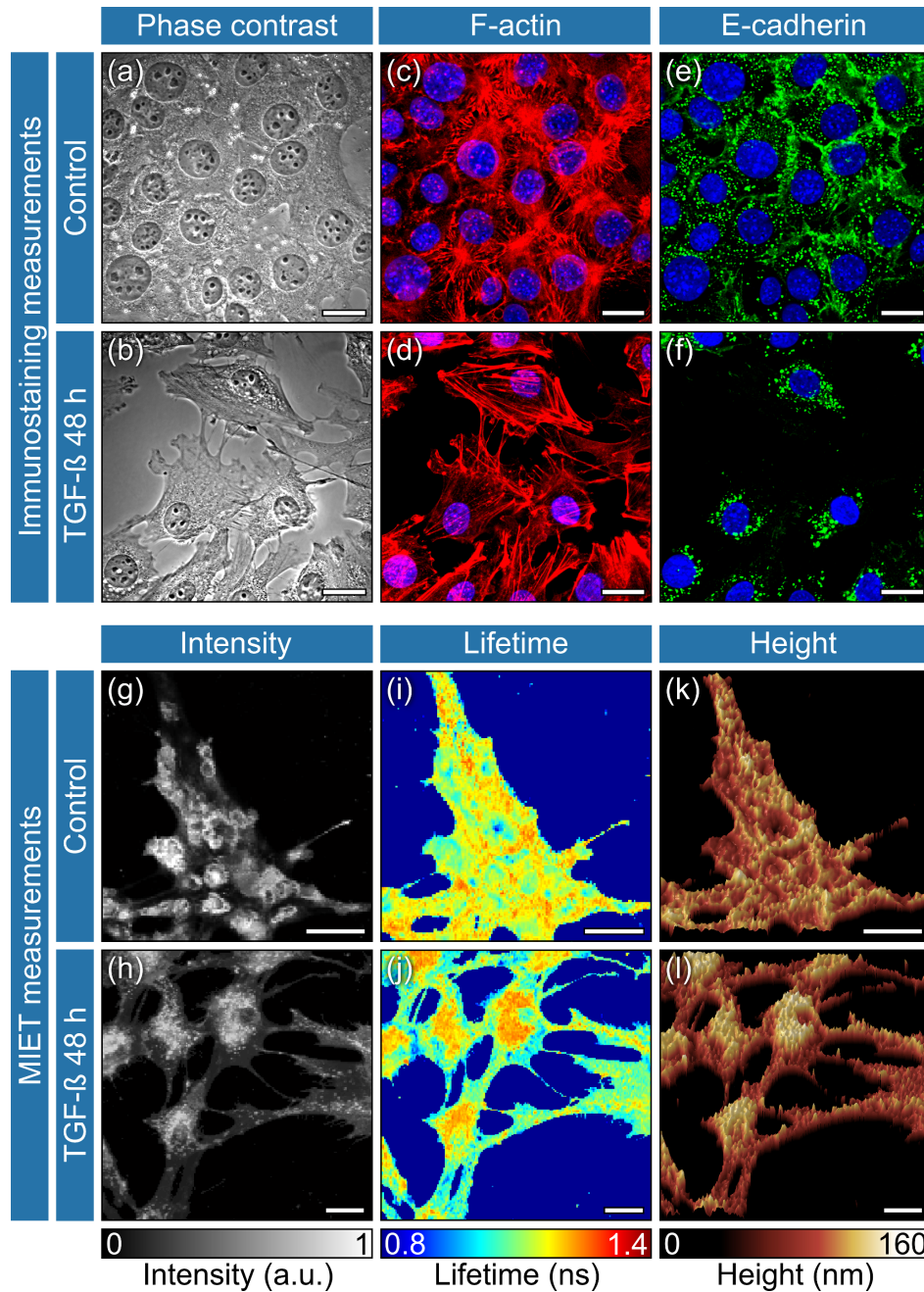
**Figure 4.2:** Morphological changes of NMuMG cells during EMT. Comparison of cells without ("Control") and after 48 hours of exposure to the transforming growth factor TGF-$\beta$1 ("TGF-$\beta$ 48 h"). (a, b) Shape changes of NMuMG cells as observed by phase contrast microscopy. (c, d) Reorganization of the F-actin cytoskeleton (red, Alexa Fluor 546-phalloidin) visualized by confocal fluorescence miscrocopy. Nuclei in blue labeled with DAPI. (e, f) Dowregulation of E-cadherin (green, Alexa Fluor 488) as observed by confocal fluorescence microscopy. (g, h, i, j) Simultaneously acquired fluorescence intensity and lifetime images of the basal membrane of living NMuMG cells grown on a gold-covered glass cover slip. (k, l) Three-dimensional height profiles of the basal cell membrane calculated from the lifetime images. Each row of images shows the same cells. Scale bars correspond to 20 μm. Immunostaining measurements performed by Bastian Brückner, MIET measurements performed by Alexey Chizhik. This image was published in [1].
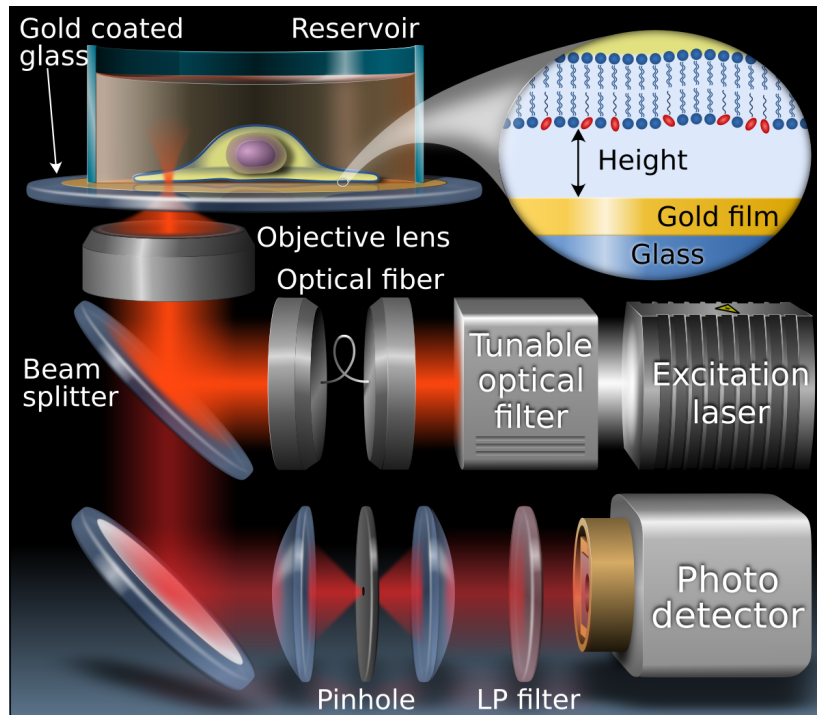
**Figure 4.3:** FLIM setup for MIET measurements of NMuMG cells during EMT. The basic setup is the same as already presented in figure 2.25. The main new feature is an incubator that keepts the cells at 37 °C; LP = longpass; tunable optical filter set to an excitation wavelength of 645 nm. This image was created by Alexey Chizhik and published in [1].
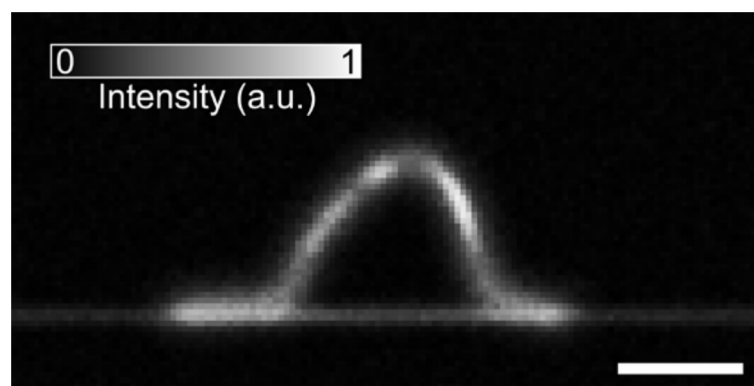


**Figure 4.4:** Confocal $z$-stack of an NMuMG cell whose plasma membrane has been stained with CellMask Deep Red, and which is attached to a glass cover slip. The scale bar corresponds to 20 µm. Note that basically no dye has been internalized by the cell, only the plasma membrane fluoresces. Image recorded by Thilo Baronsky & Alexey Chizhik and published in the supplementary information of [1].

labeled cell grown on a normal glass cover slip at different focus positions (commonly called *confocal z-stack*), see figure 4.4. Thus, measured lifetimes belong to dye molecules situated at the basal cell membrane. Employing a technique explained in detail in section 4.3, we determined the free space lifetime and quantum yield of CellMask Deep Red in water, $\tau_{0,\,\text{water}} = 2.4\,\text{ns}$ and $\Phi_{\text{water}} = 0.80$. The refractive index of the cell culture medium was the same as that of water ($n = 1.33$), thus we did not expect a change of the radiative emission rate of the fluorophore when introducing it in our sample. However, a different chemical environment can change the non-radiative emission rate. Thus, we measured the lifetime of CellMask Deep Red labeled cells on untreated glass substrates to determine "free space" values for calculation of the MIET curves, and found $\tau_{0,\,\text{cell}} = 1.5\,\text{ns}$. Assuming that this lifetime reduction is only due to a changed non-radiative emission rate, we calculated the quantum yield of the dye in the cell sample:

$$\Phi_{\text{cell}} = \frac{k_{\text{r}}}{k_{\text{r}} + k_{\text{nr, cell}}} = \frac{k_{\text{r}}}{k_{\text{r}} + k_{\text{nr, water}}} \cdot \frac{k_{\text{r}} + k_{\text{nr, water}}}{k_{\text{r}} + k_{\text{nr, cell}}} = \Phi_{\text{water}} \cdot \frac{\tau_{0,\,\text{cell}}}{\tau_{0,\,\text{water}}} = 0.50 \quad (4.3)$$

MIET was induced by a 20 nm thick gold film, which was attached to a glass cover slip (thickness 170 μm) via a 3 nm thick titanium layer. The metal films were prepared by vapor deposition using an electron beam source (Univex 350, Leybold, Cologne, Germany) under high-vacuum conditions at $\sim 10^{-6}$ mbar. During vapor deposition, the film thickness was monitored using an oscillating quartz unit, and afterwards verified by atomic force microscopy. The refractive indices of the metals were taken from [110]. Knowing all refractive indices, the thicknesses of the metal layers and the optical properties of CellMask Deep Red, we were able to calculate the lifetime-versus-height MIET curve, implemented in the MIET software package according to the procedure explained in sections 2.5.3 and 3.4. With the help of this curve, measured lifetime values could be converted to cell-substrate distances, in the following called "height values" for short.

**Results**

Exemplary fluorescence lifetime and intensity images of both untreated and TGF-$\beta$1 treated NMuMG cells are shown in figure 4.2(g)-(j). They confirm the basic conclusions drawn from the phase contrast images, namely that NMuMG cells undergo morphological changes from a compact epithelial type towards a more spindle shaped mesenchymal type upon TGF-$\beta$1 treatment. However, they offer even more information, namely how the cell-substrate distance changes during the EMT. Transforming the lifetimes into height values by means of the MIET curve, we could reconstruct a three-dimensional map of the basal membrane of the cells, as shown in figure 4.2 (k) and (l). This type of height map was then used to gain insights into the temporal progression of the EMT. To this end, MIET measurements as described above were conducted at different time points after addition of TGF-$\beta$1 to NMuMG cells. Figure 4.5(b)-(f) shows some exemplary height maps acquired 0 h (b), 12 h (c), 24 h (d), 48 h (e) and 72 h (f) after addition of the growth factor. As expected from the previous observations, the cells change their phenotype from a compact epithelial layer (b) to larger, separated cells (e, f). On a single-cell level, it was found that cells displayed a more homogeneous height
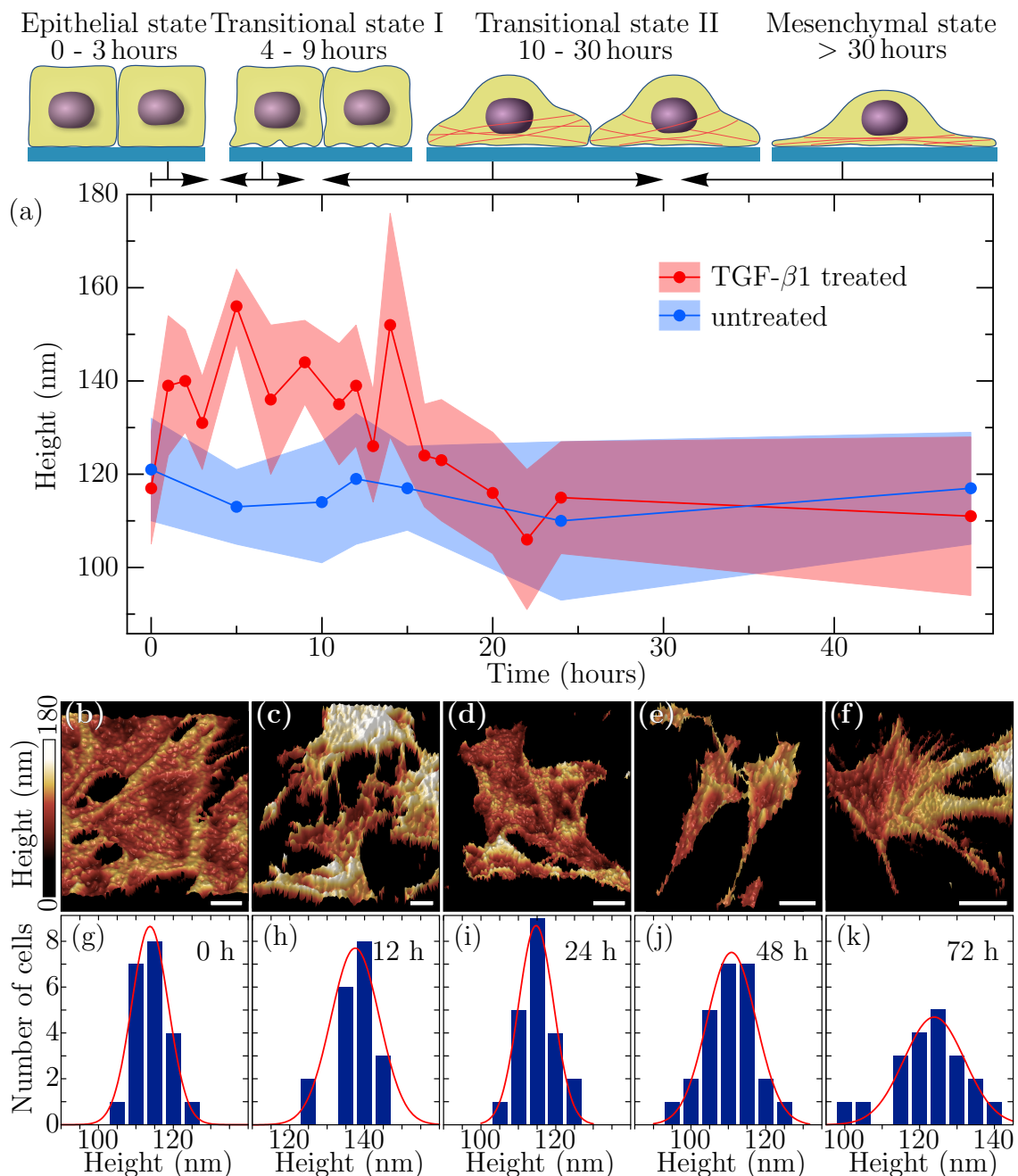
**Figure 4.5:** Cell-substrate distances during different stages of the EMT. Height maps were reconstructed from lifetime data at different time points, as shown exemplarily in (b)-(f) for cells examined 0 h, 12 h, 24 h, 48 h and 72 h after TGF-$\beta$1 addition (scale bar 20 μm), allowing to draw conclusions on a single-cell level. Distributions of average cell-substrate distances for the different time points were found as explained in figure 4.6; histograms and Gaussian fits are shown in (g)-(k) for the same time points. The mean values of these distributions are plotted as red dots in (a), where the shaded region corresponds to the standard error of the mean. See the table in the main text for numbers of evaluated cells. As a comparison, the results for untreated cells are shown in blue. Above the graph, cartoons depict the compact, ordered epithelial state, the transitional state I with dissolution of cell-cell contacts and increased membrane ruffling, the transitional state II where cells change their shape and start to form actin stress fibers (red lines) and finally the mesenchymal-like state, where cells spread further and form more stress fibers. Data recorded and image created by Alexey Chizhik, published in [1].
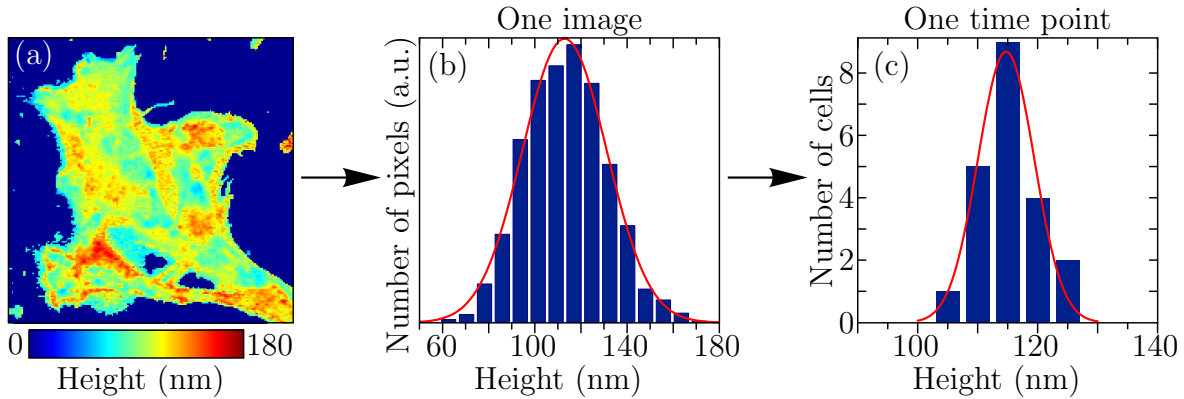
**Figure 4.6:** Illustration of the necessary steps to gain information about average cell-substrate distances at different time points after addition of TGF-$\beta$1. (a) For each single cell, the lifetime image was converted to a height map using the MIET curve. Some examples can be found in figure 4.5(b)-(f). (b) By employing an intensity threshold, the pixels belonging to the cell were identified and their height values histogrammed. (c) The mean height values of all cells belonging to the same time point were then collected in a common histogram, some of these are shown exemplarily in figure 4.5(g)-(k). The average cell-substrate distances of all time points as determined by Gaussian fits to the data (red line) are plotted as red dots in figure 4.5(a), together with the standard error of the mean. Image created by Alexey Chizhik and published in the supplementary information of [1].

after completion of the EMT (e, f) compared to cells at the beginning of the process (b, c). This is in agreement with the MIET study [39] mentioned at the beginning of this chapter, where the cell-substrate distances of three cell types were compared. There, it was found that the benign epithelial cell line displayed a rather wavy basal membrane, while the height map of a mesenchymal cancer cell line appeared much smoother.

In order to collect information about average cell-substrate distances, MIET images of a larger number of cells at different time points were acquired, see the following table:

| treatment time [h] | 0 | 1 | 2 | 3 | 5 | 7 | 9 | 11 | 12 | 13 | 14 | 16 | 17 | 20 | 22 | 24 | 48 | 72 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| number of cells | 21 | 21 | 10 | 17 | 7 | 11 | 18 | 17 | 19 | 6 | 6 | 14 | 10 | 16 | 20 | 21 | 25 | 20 |

The data evaluation process is illustrated in figure 4.6: For each single cell, a height map of the basal membrane was constructed as before. With the help of an intensity threshold, pixels belonging to the cell were identified and their height values histogrammed. The mean values of all cells belonging to the same time point were then collected in another histogram, shown for exemplary time points in figure 4.5(g)-(k). Gaussian fits of these latter histograms resulted in mean cell-substrate distances for 17 time points, which are shown as red dots in figure 4.5(a). The standard error of the mean is visualized as a red shaded area, identical measurements for untreated (control) cells are depicted in blue. For comparison, the different stages of the EMT identified in [109] are indicated with black arrows and schematic cartoons at the top of figure 4.5. While untreated cells maintain a constant average height of 110 nm-120 nm during the whole observation time, TGF-$\beta$1 treated cells significantly increase the cell-substrate distance by 20 nm to 30 nm in the initial stages of the EMT. Approximately ten to twenty hours after EMT

initiation, i.e. within the transitional state II, the average height of the basal membrane decreases, ultimately returning to the level of untreated cells. Cells were measured up to 72 h without noticeable further changes of the average cell-substrate distance (data not shown).

It remains to investigate what might be the biological mechanism of these height changes. One hypothesis is that the *extracellular matrix*[3] changes its thickness during the EMT, pushing the cell upwards. If this was the case, one would expect that all parts of the cell were elevated similarly. Another hypothesis is that the number of *focal adhesions* (the main subject of section 4.1.3, large protein complexes connecting the cytoskeleton to the extracellular matrix) decreases during the EMT, allowing the membrane in between to lift-off and be raised higher above the surface.

For a test of these hypotheses, we conducted time-resolved MIET measurements within the first six hours of the EMT, the time range during which cells display the initial and strongest height changes. Since even when observed in an incubator, cells gradually lost vitality, each single cell was only examined for 1.5 hours. Each coloured dot in figure 4.7(a) corresponds to the average height of one cell measured at one time point, where all data points belonging to the same cell have the same colour. Images were acquired every two minutes, providing a good time resolution with only moderate photobleaching of the dye. Data sets with exactly the same time points belong to experiments where it was possible to monitor two cells simultaneously. The general trend of the plot is in accordance with figure 4.5(a): The cell-substrate distance increases rapidly within the first hour of the EMT and then stays at a high level with large fluctuations. In contrast, untreated control cells (shown as black solid and dashed lines in figure 4.7(a)) display a constant height of the basal membrane above the substrate. This indicates that the repeated lifetime measurements did not have a detrimental effect on the cells within the duration of the experiment. Thus, we are now able to exploit the full potential of MIET microscopy by studying in detail the cell-substrate distance on a subcellular level. Figure 4.7(b) shows a rendered three-dimensional height map of the basal membrane of one single cell at four different time points within the first 1.5 hours of the EMT. Already from these images, it is apparent that the majority of the membrane area is elevated during this time, while there are also regions with close cell-substrate contact that remain largely unchanged. In the following, the latter areas are called *adhesion points* since their exact molecular nature was not studied in this work, which means that they could not be identified unequivocally as focal adhesions. Histograms of the height values of all pixels whose intensity was higher than the background are presented in figure 4.7(c). With the exception of the time $t = 0$ h, these distributions show a pronounced skewness toward larger height values and thus do not correspond to single Gaussian functions anymore. Instead, we decomposed them into two populations for $t > 0$ h by fitting two Gaussian distributions to the data (red solid lines). We interpret the areas closer to the substrate as adhesion points. Their average height is always close to 120 nm during the measurement, increasing by less than 10 nm. This increase might be attributed to the production of extracellular matrix proteins. The second peak at larger cell-substrate distances was assigned to the cell membrane between the adhesion points. The average distance of this population increases more strongly, from

---

[3] A three-dimensional macromolecular network comprised of water, proteins and polysaccharides that are secreted by the cells which forms a sheet-like structure underneath epithelial cells.
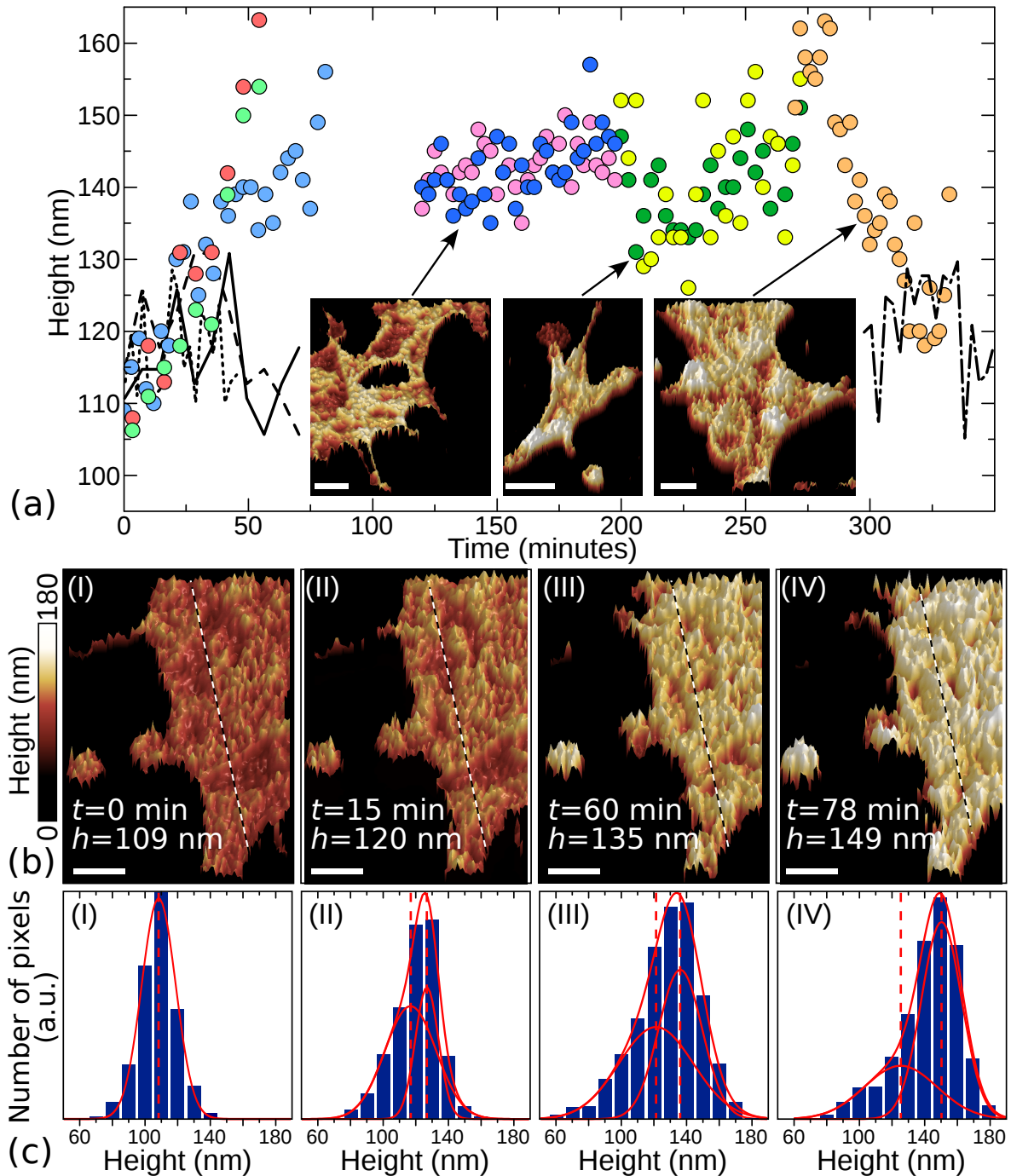
**Figure 4.7:** Time-elapsed MIET imaging of the cell-substrate distance during the first six hours of the EMT. (a) Coloured dots correspond to the average cell-substrate distance of one cell at one time point, where subsequent measurements of the same cell have the same colour. Exemplary height maps are shown as insets. (b) Height maps of the same cell acquired 0 min, 15 min, 60 min and 78 min after induction of the EMT. The average cell-substrate distance increases, however, some regions remain close to the substrate. Cross sections along the dashed lines are shown in figure 4.8. The scale bars correspond to 20 µm in all height maps. (c) Histograms of the height values calculated for single pixels of the height maps shown in (b). The histograms can be fitted with a sum to two Gaussian distributions (dashed lines, sum represented by solid line). Vertical dashed lines show the maxima of the Gaussian distributions, indicating that average cell-substrate distance of the population at larger height increases during the early EMT. Data recorded and image created by Alexey Chizhik, image published in [1].
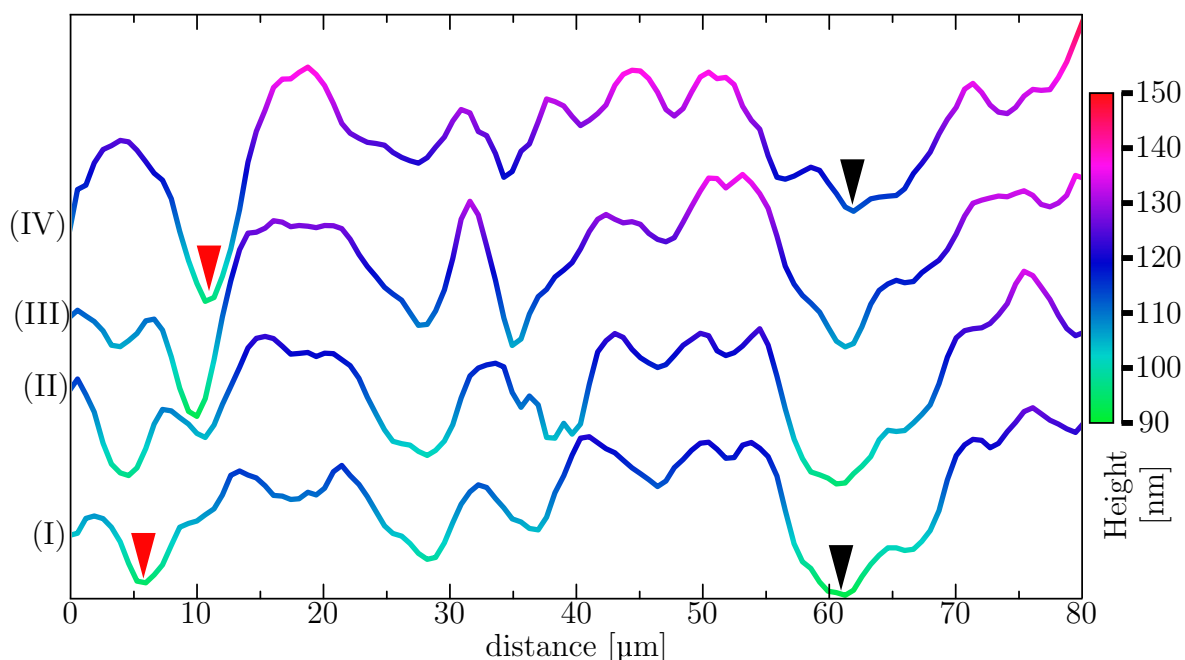
**Figure 4.8:** Time evolution of a cross section through the basal cell membrane at 0 min to 80 min after induction of the EMT (black dashed lines in the height maps shown in figure 4.7(b)). Colour encodes height, the numbers (I-IV) indicate the time points. Both the formation of new adhesion points (red arrows) and the dissolution of adhesion sites (black arrows) can be observed, while the membrane between these adhesion points overall displays an increasing height ("lift off"). Data recorded and image created by Alexey Chizhik, image published in [1].

approximately 110 nm to 150 nm within 80 minutes. Furthermore, the relative amount of membrane associated with adhesion points decreases steadily during this measurement:

| time after EMT induction [min] | 15 | 60 | 78 |
|---|---|---|---|
| fractional area of Gaussian fit 1 [%] | 65 | 49 | 33 |
| fractional area of Gaussian fit 2 [%] | 35 | 51 | 67 |

These findings motivated us to study the adhesion points in more detail. Figure 4.8 depicts four consecutively acquired cross sections (stacked) of the membrane-substrate distance taken along the dashed lines in figure 4.7. Both dissolution (black arrowheads) and formation (red arrowheads) of adhesion points could be observed. However, on average, more adhesion sites are lost than newly formed during this period (data not shown). The temporal resolution of 2 min is in line with the expected turnover time of focal adhesions [111]. The cross-sections also show that the height of the membrane in between focal points increases, as discussed above.

**Summary**

To sum up, the $z$-resolution of MIET of approximately 2 nm to 3 nm allowed us to study the cell-substrate distance of NMuMG cells during the epithelial-to-mesenchymal transition with high precision. We found that during the initial stages of the EMT, the average distance between the basal membrane of the cell and the substrate increases significantly, by 20 nm-30 nm. This might be partly due to an increased thickness of the extracellular matrix, but is mostly caused by a reduction of focal adhesion points and an elevation ("lift-off") of the membrane between these points. In the transitional state II, the cell-substrate distance slowly decreases, finally reaching the level of untreated cells. Furthermore, our results indicate that the general structure of the basal membrane changes from a rather wavy form to a smoother profile during the EMT, which could be explained by the formation of stress fibers and a higher tension of the final mesenchymal phenotype reported in [112].

Our findings are in agreement with previous works using complementary methods that each studied different aspects of the EMT and related fields: The ECIS measurements of [109] not only suggested the different stages of the EMT and reported increased vertical motion of the cell membrane at the onset of EMT, but also found evidence of active processes generating long memory fluctuations also explainable by the lift-off process found by us. MIET measurements of different cell lines in [39] indicated the same trend of wavy or smooth membranes for epithelial and mesenchymal cells as we saw. And finally, results concerning signalling cascades and the activation of certain proteins involved in remodeling of focal adhesions support the hypothesis of dynamical changes of the adhesion sites, see [1] for further details. Thus, we were able to successfully shine light on a complex cellular adhesion process with unprecedented axial resolution.

## 4.1.3 Dual-color metal-induced and Förster resonance energy transfer for cell nanoscopy

With the next application, we stay in the realm of cell-substrate interactions. Instead of studying the cell membrane, we now focus on focal adhesions, sites of strong interactions between the cell's cytoskeleton and the substrate. The results were published in:

[4] A. M. Chizhik, C. Wollnik, **D. Ruhlandt**, N. Karedla, A. I. Chizhik, L. Hauke, D. Hähnel, I. Gregor, J. Enderlein, and F. Rehfeldt, "Dual-color metal-induced and Förster resonance energy transfer for cell nanoscopy," *Molecular Biology of the Cell*, vol. 29, no. 7, pp. 846–851, 2018. **D. R.** provided analysis tools and analysed the MIET data.

**Motivation**

Depending on the situation, cells need to be able to preserve or change their shape, be tightly linked to their environment or move relative to it. A key player in these tasks is the *cytoskeleton*, an intracellular biopolymer scaffold encompassing three types of long, thin structures: actin filaments, intermediate filaments, and microtubules. This network is highly dynamic, filaments can polymerize and depolymerize, or interact with other cellular components such as motor proteins to enable movement and intracellular transport. In section 4.1.2, we already encountered one component of the cytoskeleton, so-called actomyosin *stress fibers*. These are contractile bundles of actin fibers, motor proteins and various crosslinking proteins. Due to the relative rigidity of the cytoskeleton, it also plays an important role in the transmission of forces both within a single cell and between neighbouring cells or cells and the extracellular matrix. The latter situation is mediated by *focal adhesions*, large complexes consisting of different proteins (among
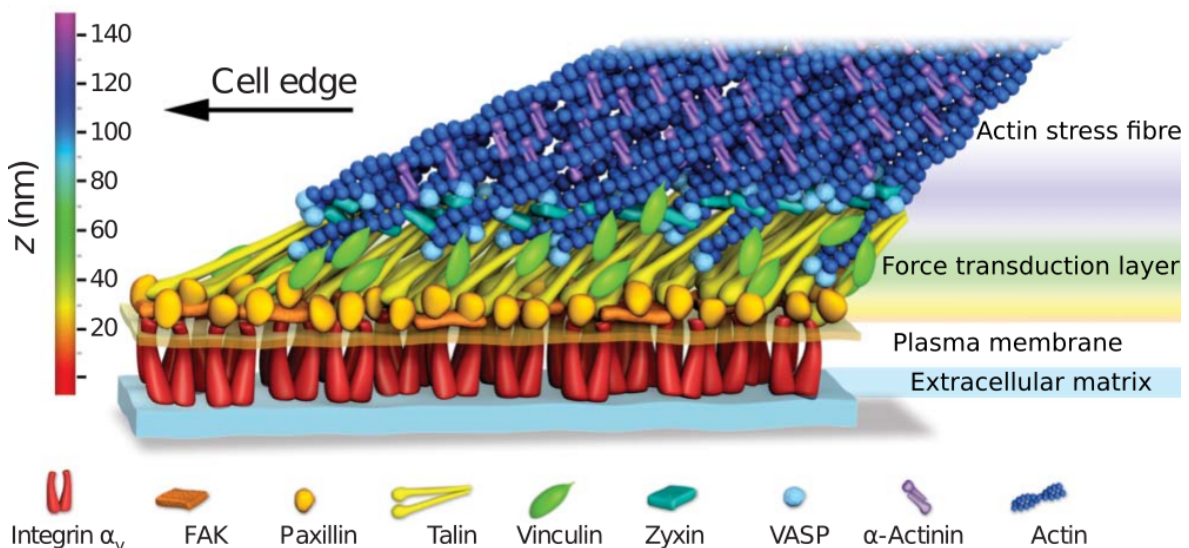


**Figure 4.9:** Architecture of focal adhesions as suggested by Kanchanawong *et al.* in [113]. Focal adhesions consist of many different proteins, among them vinculin, which are suggested to be organised in (partially overlapping) layers, from which actin stress fibers emerge. Image adapted from [113].

them *vinculin*, on which we focused in our experiments) which link stress fibers to the extracellular matrix. Focal adhesions are known to translate mechanical cues from the extracellular environment into biochemical signaling, influencing the behaviour of the cell. Further details on the cytoskeleton and focal adhesions can be found in biology textbooks such as [108] and also in our publication [4].

In order to understand the *mechano*-sensing of cells, it is important to deduce the forces acting at focal adhesions and the adjoining stress fibers. A prerequisite for this step is to know the exact architecture of the focal adhesion/stress fiber complexes. In 2010, Kanchanawong *et al.* used interferometric PALM to map the organization of nine different proteins in focal adhesions with an axial resolution of 10 nm-15 nm [113]. Their study resulted in the model of a stratified system with partly overlapping layers shown in figure 4.9. However, Kanchanawong *et al.* were mostly interested in the average height values of the different proteins, and did not directly examine the co-localization of proteins (i.e., they only labeled one type of protein per cell, conducting a separate experiment for each protein). Furthermore, iPALM requires specialized markers, namely photoactivatable fluorescent proteins fused to the protein of interest, and complex instrumentation, thus the work by Kanchanawong *et al.* is not easily reproducible. In contrast, MIET can be achieved with any fluorescent marker[4], in particular with commercially available fluorescently labeled antibodies, and using any FLIM setup. The only peculiarity is the need for metal-coated substrates, but those can be either purchased from commercial suppliers or be produced by vapor deposition using an electron beam source as existing in many research institutions. Therefore, we set out to confirm the suggested model by investigating the three-dimensional architecture of focal adhesions using MIET microscopy. Furthermore, we wanted to complement the previous work by detecting two types of proteins simultaneously, allowing us to directly observe their relative positions.

**Methods**

Since we wanted to elucidate the interplay of focal adhesions and stress fibers, we had to co-localize focal adhesions and stress fibers in the same cells. To this end, we extended MIET microscopy to a dual-colour technique, termed *dual-colour MIET* (dcMIET). This required an adjustment of the optical setup, the experimental procedure and also the data evaluation, as will be explained below. As a model system, we chose human mesenchymal stem cells (hMSCs) because it is known that focal adhesions and stress fibers are key players in their mechano-induced differentiation, see e. g. [114]. Stress fibers were stained using the red dye Atto 647N (ATTO-TEC GmbH, Siegen, Germany) and vinculin as an exemplary focal adhesion protein was labeled with the green dye Atto 488 (ATTO-TEC). MIET substrates were prepared as described in section 4.1.2, they consisted of glass cover slips covered with 3 nm titanium, 15 nm gold, 3 nm titanium and 20 nm silica ($SiO_2$), in this order. The silica spacer was introduced to prevent quenching of the dyes if the focal adhesions were very close to the substrate. The setup we used is shown schematically in figure 4.10, it is the same type of confocal FLIM microscope as introduced in section 2.5.2. The combination of a white-light

---

[4] Provided that the lifetime is not too short (free space value of at least $\sim 2$ ns), otherwise the lifetime fits can become unreliable.
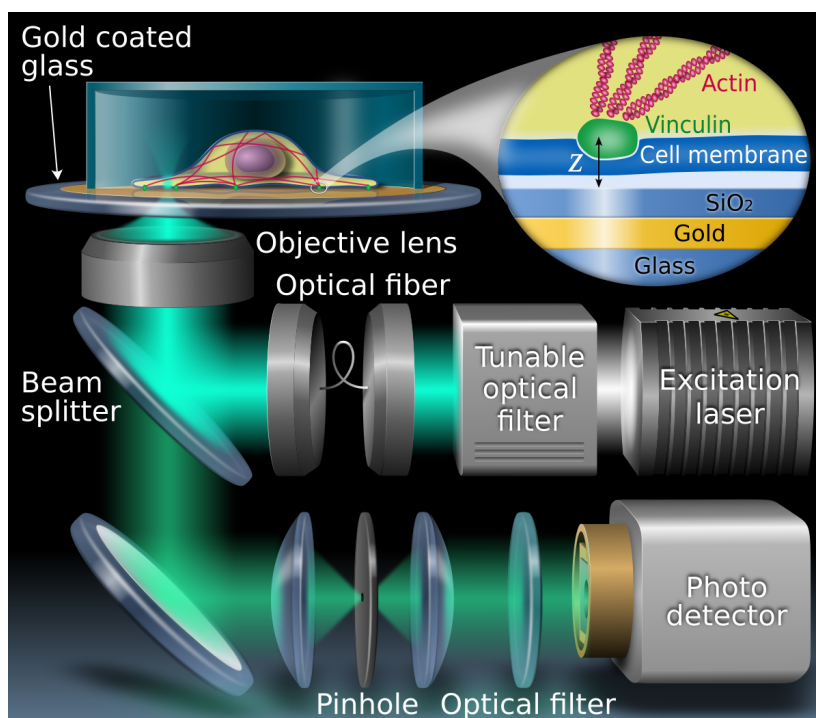
**Figure 4.10:** Dual-colour MIET setup used to study focal adhesions and stress fibers. The basic design is the same as in figures 2.25 and 4.3. Since we worked with chemically fixed cells, no incubator was needed. The main new feature is that two different excitation wavelengths, namely 488 nm and 635 nm, were generated consecutively by the white light laser and the acousto-optic tunable filter without any required hardware changes in the excitation pathway. In the detection pathway, different longpass and bandpass filters were utilized to block back-scattered excitation light. For further details on the manufacturers and the specifications of the optical components, see [4]. Image created by Alexey Chizhik and published in [4].

laser and an acousto-optical tunable filter allowed to switch easily between the two excitation wavelengths (488 nm and 635 nm) needed to efficiently excite the fluorescent dyes. Different longpass and bandpass filters were used to block residual excitation light for the red and green channels, and fluorescent lifetimes were calculated according to the procedure described in section 3.5. Additionally, a high-powered red laser (laser power $\sim 1\,\text{mW}$ after passing the objective lens) was used in some measurements to bleach Atto 647N. See the methods section of [4] for further details on the cell culture and staining protocols as well as the manufacturers and the specifications of the optical components.

When designing the experimental procedure, we had to take into account the probable occurrence of an additional process that can change the fluorescence lifetime, called *Förster resonance energy transfer* (FRET) after its discoverer Theodor Förster, see e.g. textbooks such as [57] chapter 13, or reviews such as [115]. FRET can be observed when two different types of fluorophores are used within the same sample, and the emission spectrum of one fluorophore (called the *donor*, Atto 488 in our experiments) overlaps with the absorption spectrum of the other fluorophore (called the *acceptor*, Atto 647N). If two such molecules come very close ($\sim 10\,\text{nm}$), the near field of the

donor can excite the acceptor, thus resulting in an energy transfer. The presence of this additional deexcitation pathway leads to a reduction of the donor's fluorescence lifetime.

This effect can be treated exactly using concepts introduced in chapter two. Briefly, the excitation rate of an acceptor molecule is given by the product of its wavelength-dependent absorption cross section $\sigma_A(\lambda)$ and the number of (virtual) photons per area and time reaching it from the donor. The latter is simply the Poynting vector of the donor's near field divided by the energy per photon $\lambda/hc$. Since the acceptor dipole excitation is proportional to the scalar product of its dipole moment and the exciting electric field orientation, both the relative orientation between donor and acceptor and their distance $r$ also influence the rate of energy transfer. The exact mathematical formulation of these considerations finally results in an energy transfer rate $k_{\mathrm{FRET}}$ that is given by

$$k_{\mathrm{FRET}} = \frac{1}{\tau_{\mathrm{D}}} \cdot \left( \frac{R_0}{r} \right)^6, \tag{4.4}$$

where $\tau_{\mathrm{D}}$ is the donor's lifetime in the absence of the acceptor and $R_0$ is the *Förster radius*. This quantity includes all parameters mentioned above:

$$R_0^6 = \frac{9}{128\pi^5 n^4} \; \kappa^2 \; \Phi_{\mathrm{D}} \int \mathrm{d}\lambda \; \lambda^4 F_{\mathrm{D}}(\lambda)\sigma_A(\lambda). \tag{4.5}$$

Here, $n$ is the refractive index of the surrounding medium, $\kappa^2$ is an orientation factor which has the value 2/3 for freely rotating donor and acceptor, $\Phi_{\mathrm{D}}$ is the quantum yield of the donor and $F_{\mathrm{D}}(\lambda)$ is the normalized emission spectrum of the donor ($\int \mathrm{d}\lambda \; F_{\mathrm{D}} = 1$). Since it is more common to measure the molar extinction coefficient $\varepsilon_A$ (in L·mol$^{-1}$·cm$^{-1}$) than the absorption cross section (in cm$^2$) of a fluorophore, one can also use the relation

$$\sigma = \frac{1000 \cdot \ln 10}{N_A} \; \varepsilon \tag{4.6}$$

to arrive at the standard textbook expression [57] for $R_0$:

$$R_0^6 = \frac{9000 \cdot \ln 10}{128\pi^5 n^4 N_A} \; \kappa^2 \; \Phi_{\mathrm{D}} \int \mathrm{d}\lambda \; \lambda^4 F_{\mathrm{D}}(\lambda)\varepsilon_A(\lambda). \tag{4.7}$$

For most donor-acceptor pairs, the Förster radius is between 3 nm and 6 nm [115]. For our sample, we calculated $R_0 = 5.0$ nm (see supporting information of [4] for details). Due to the sixth-power dependence of $k_{\mathrm{FRET}}$ on $r$, FRET is very sensitive to small distance changes for donor-acceptor distances in the order of 1 nm to 10 nm. At larger distances, $k_{\mathrm{FRET}}$ is negligible compared to the excited donor's radiative and non-radiative decay rates $k_{\mathrm{r}}$ and $k_{\mathrm{nr}}$. Because of this high distance sensitivity, FRET microscopy has become a routine tool in modern cell biology, for example in the field of cellular signalling phenomena [115]. Since $k_{\mathrm{FRET}}$ cannot be measured directly, it is
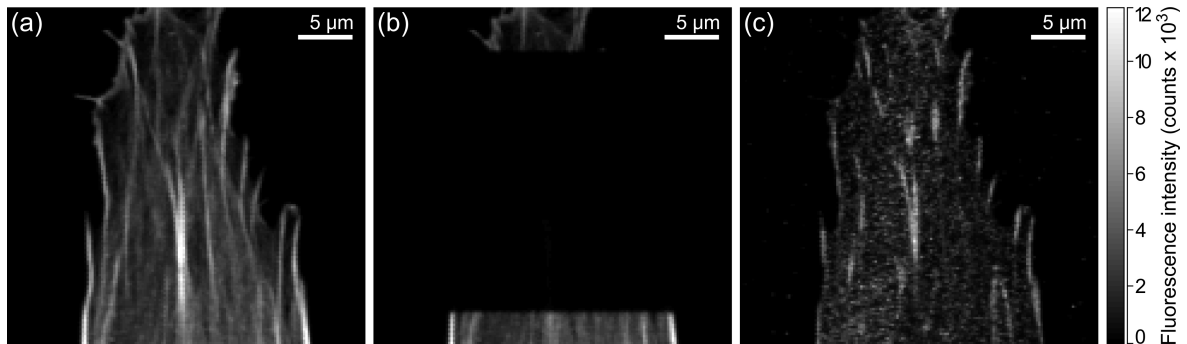
**Figure 4.11:** Demonstration of the photobleaching of the FRET acceptor. (a, b) Fluorescence intensity images of Atto 647N-actin before and after photobleaching with a red diode laser, respectively. (c) Fluorescence intensity image of Atto 488-vinculin after actin bleaching. The cell was fixed on glass 24 h after seeding. Image size is 32 μm × 32 μm (150 × 150 scan positions), acquisition time was 5 ms per pixel. Note that the potential FRET acceptor Atto 647 N is *de facto* bleached completely, while the FRET donor Atto 488 is not affected at all. Data recorded by Anna Chizhik, published in [4].

inferred from other quantities such as the donor's fluorescence intensity or lifetime in the absence (D) and presence (DA) of the acceptor. The lifetime changes according to:

$$\tau_{\mathrm{D}} = (k_{\mathrm{r}} + k_{\mathrm{nr}})^{-1}$$

$$\tau_{\mathrm{DA}} = (k_{\mathrm{r}} + k_{\mathrm{nr}} + k_{\mathrm{FRET}})^{-1} = \left(\frac{1}{\tau_{\mathrm{D}}} + k_{\mathrm{FRET}}\right)^{-1} = \tau_{\mathrm{D}} \cdot \left(1 + \frac{R_0^6}{r^6}\right)^{-1}. \qquad (4.8)$$

Thus, both FRET and MIET decrease the fluorescence lifetime of the donor, while the fluorescence lifetime of the acceptor is only influenced by MIET. This has to be taken into account when two types of fluorophores with overlapping spectra are used simultaneously in a MIET measurement. In our experiments, we combined FRET and MIET to make full use of the advantages of both methods. To this end, we modified an experimental procedure that has long been established [116] in FRET studies: In order to determine both $\tau_{\mathrm{DA}}$ and $\tau_{\mathrm{D}}$ at the same points of a sample, one can perform a first lifetime measurement of the donor ($\tau_{\mathrm{DA}}$), then completely bleach the acceptor with a strong excitation source that coincides only with the acceptor's and not the donor's absorption spectrum (and thus does not destroy the donor), and then acquire a second lifetime value for the donor ($\tau_{\mathrm{D}}$). In some of our measurements, we employed exactly this technique on cells seeded on normal glass cover slips to measure FRET between Atto 488-vinculin and Atto 647N-actin. In all MIET experiments on metal-coated substrates, we started by acquiring a lifetime image in the green channel (Atto 488-vinculin). The dye was then bleached as for the FRET experiments, enabling the subsequent collection of lifetime data in the red channel (Atto 647N-actin) that was only influenced by MIET, not FRET. Figure 4.11 demonstrates the effectiveness of this approach, by comparing an intensity image of Atto 647N-actin before (a) and after (b) ten minutes of photobleaching with light at $\lambda = 640$ nm provided by a diode laser (laser power $\sim 1$ mW after passing the objective lens). An intensity image of Atto 488-vinculin acquired after the photobleaching of Atto 647N does not show any

bleaching of Atto 488, see 4.11(c).

The main novelty in the data evaluation was the need for two MIET curves, one for each dye. The refractive index of the mounting medium Fluoroshield[5] (Sigma-Aldrich, St. Louis, MO) was determined to be 1.4, the wavelength-dependent refractive indices of the metals were taken from [110]. The manufacturer provided emisison spectra of the two dyes as well as the quantum yield and fluorescence lifetime of Atto 647N in phosphate-buffered saline (PBS), $\Phi_{647N, \text{PBS}} = 0.65$ and $\tau_{0, 647N, \text{PBS}} = 3.5\,\text{ns}$. For Atto 488, we measured the quantum yield and fluorescence lifetime in PBS with a metal nanocavity, a technique that will be explained in detail in section 4.3. We found $\Phi_{488, \text{PBS}} = 0.67$ and $\tau_{0, 488, \text{PBS}} = 2.9\,\text{ns}$. When a dye is conjugated to an antibody and introduced in a cell, the non-radiative decay rate of its excited state can be modified due to interactions with the surroundings, while the radiative rate can change due to a different refractive index of the environment. This leads to changes of both the lifetime and the quantum yield as compared to measurements of the pure dye suspended in a buffer such as PBS. We took these effects into account by acquiring FLIM images of labeled cells on uncoated glass cover slips[6], see figure 4.12, and found average lifetimes $\tau_{0, 488, \text{cell}} = (1.9 \pm 0.2)\,\text{ns}$ and $\tau_{0, 647N, \text{cell}} = (3.4 \pm 0.2)\,\text{ns}$. Since we saw in section 2.4.4 that the total energy flux $S_{\text{tot}}$ of an ideal electric dipole emitter is proportional to the refractive index, we assumed that the radiative rate changes according to[7]

$$\frac{k_{\text{r, cell}}}{k_{\text{r, PBS}}} = \frac{S_{\text{tot, cell}}}{S_{\text{tot, PBS}}} = \frac{n_{\text{cell}}}{n_{\text{PBS}}}. \tag{4.9}$$

This resulted in a modified quantum yield

$$\Phi_{\text{cell}} = \Phi_{\text{PBS}} \cdot \frac{\tau_{\text{cell}}}{\tau_{\text{PBS}}} \cdot \frac{n_{\text{cell}}}{n_{\text{PBS}}}. \tag{4.10}$$

With $n_{\text{PBS}} = 1.33$ and the already mentioned refractive index of the mounting medium $n_{\text{cell}} = 1.4$, we obtained $\Phi_{488, \text{cell}} = 0.46$ and $\Phi_{647N, \text{cell}} = 0.66$. Finally, we assumed an isotropic orientation of the dye molecules, which is justified considering that they were linked to the makromolecules recognizing actin or vinculin (the filamentous-actin-binding toxin phalloidin for Atto 647N-actin and one primary and one secondary antibodiy for Atto 488-vinculin) via a flexible linker. The two MIET curves calculated using these sample parameters are shown in figure 4.13 (solid lines), they provided a direct relation between measured lifetimes on metal-coated cover slips and height values above the silica spacer. As a reference, the free space lifetimes in the cells $\tau_{0, 488, \text{cell}}$ and $\tau_{0, 647N, \text{cell}}$ are given as dashed lines.

When evaluating MIET measurements, we only took pixels into account whose intensity was above a certain threshold (usually 3000 counts, compare with simulations in section

---

[5] Since the chemically fixed cells were permeabilized using Triton X-100 before being stained and mounted, we assume that the dye molecules are surrounded by the mounting medium.

[6] To exclude FRET, we firstly measured the lifetime of Atto 647N, then bleached this dye and measured the lifetime of Atto 488, as described above.

[7] After the submission of the work presented here, we explored this topic further, as will be explained in section 4.1.4. When using those improved formulas, we find $\Phi_{488, \text{cell}} = 0.49$ and $\Phi_{647N, \text{cell}} = 0.70$. However, the simulations presented in section 4.2 show that this only leads to a shift of determined height values in the order of $1\,\text{nm}$ to $3\,\text{nm}$ for dye molecules in focal adhesions.
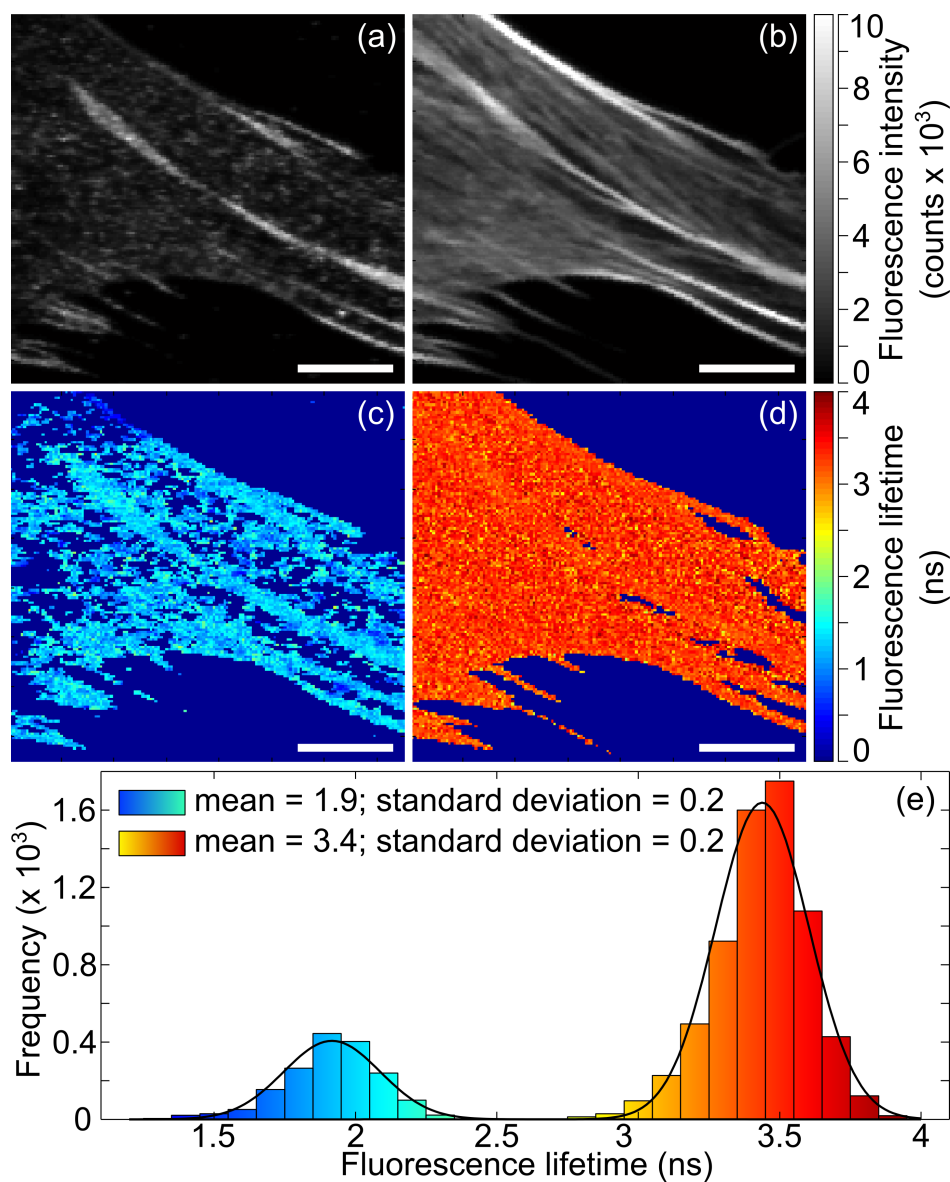
**Figure 4.12:** Free space lifetime measurements of Atto 488-vinculin and Atto 647N-actin. Simultaneously acquired fluorescence intensity (a, b) and lifetime (c, d) images of vinculin (a, c) and actin (b, d) filaments. The images have a total size of 20 μm × 20 μm (150 × 150 scan positions), the acquisition time was 5 ms. All scale bars correspond to 5 μm. (e) Depicts the lifetime distributions for vinculin (blue) and actin (orange). The fit with Gaussian distributions (black lines) resulted in $\tau_{0,\,488,\,\text{cell}} = (1.9 \pm 0.2)$ ns and $\tau_{0,\,647\text{N},\,\text{cell}} = (3.4 \pm 0.2)$ ns. Data recorded by Anna Chizhik, image published in the supplementary information of [4].
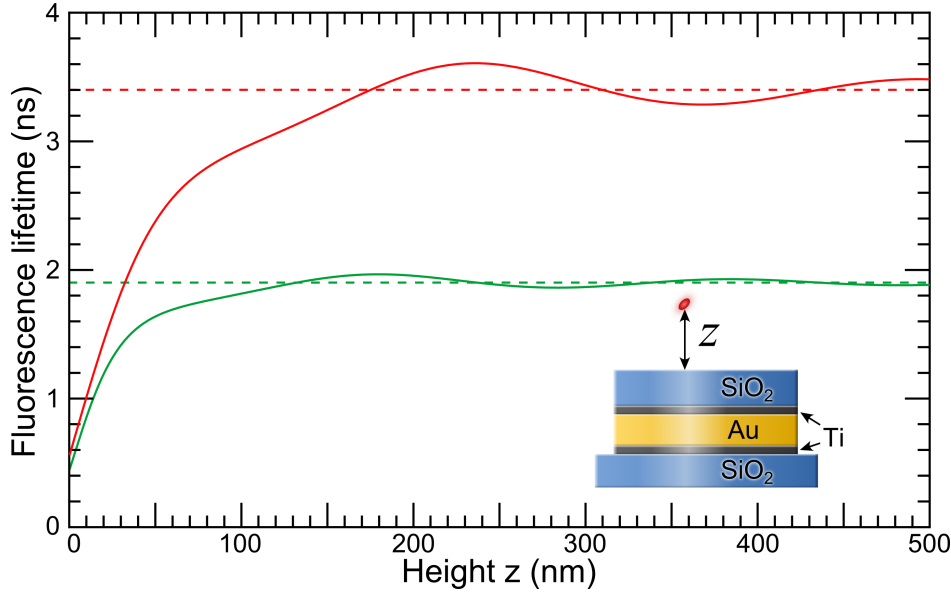
**Figure 4.13:** MIET lifetime-*versus*-height curves used to convert measured fluorescence lifetimes into distances between the fluorophores and the top of the substrate. The red solid curve was calculated for Atto 647N-actin, the green solid curve for Atto 488-vinculin. The dashed horizontal lines show the free space lifetimes $\tau_{0,\,647N,\,cell} = 3.4\,\text{ns}$ and $\tau_{0,\,488,\,cell} = 1.9\,\text{ns}$. Image published in the supplementary information of [4].

3.5). This ensured that only meaningful lifetime (and thus height) values were used to draw conclusions regarding the biological sample: Pixels with smaller intensities could either belong to regions without sample, where the detected fluorescence is just noise and the fitted lifetimes are random, or to true sample regions with a low intensity, where lifetime fits would be highly unreliable.

For the evaluation of the FRET measurements, the fluorescence lifetime of the donor Atto 488-vinculin acquired before (DA) and after (D) photobleaching of the acceptor Atto 647N-actin had to be compared. Since the calculated distance $r$ between donor and acceptor depends sensitively on the lifetime, we optimized our evaluation scheme to account for the accuracy of the lifetime fits: In a first step, we identified connected pixel clusters with a high fluorescence intensity in both images, which we interpreted as focal adhesions if they contained a minimum of 20 pixels. This largely prevented us from evaluating e.g. contaminations whose autofluorescence was in the same wavelength region as the fluorescence of Atto 488, or small amounts of vinculin in the cytoplasm or other areas of the cell. In a second step, the clusters were divided into sections of approximately 4×4 pixels (deviating from this number at the edges of the clusters), pooling all photons to increase the signal-to-noise ratio. For each such section, the fluorescent lifetimes $\tau_{DA}$ and $\tau_D$ were determined using a bootstrapping procedure: In order to be able to estimate the lifetime uncertainty, the detection events were divided into bunches of $10^4$ photons. From each such bunch, a TCSPC histogram was generated, and a lifetime value obtained by tail-fitting as explained in section 3.5. Subsequently, the mean $\langle\tau_{DA}\rangle$ (or $\langle\tau_D\rangle$) and standard deviation $\sigma_{DA}$ (or $\sigma_D$) of all lifetime values belonging to the same section were calculated. Only if the difference between $\langle\tau_{DA}\rangle$
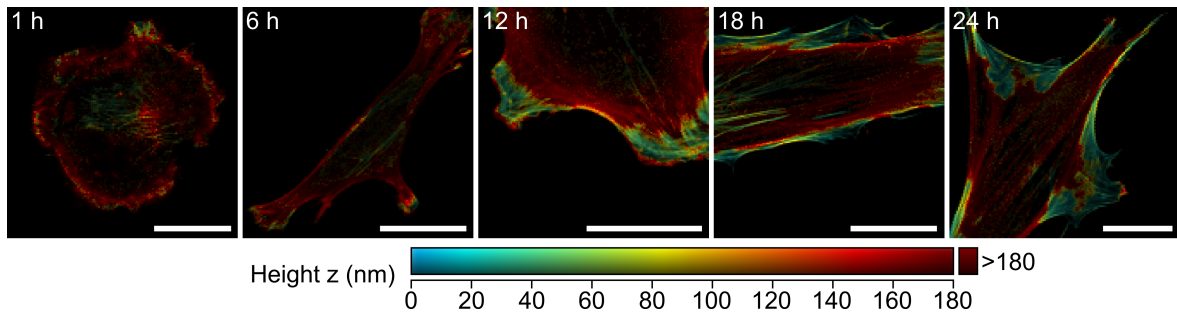
**Figure 4.14:** Time series of three-dimensional actin maps reconstructed from lifetime measurements. Human mesenchymal stem cells were fixed 1 h, 6 h, 12 h, 18 h or 24 h after seeding on metal-coated MIET substrates and actin filaments were labeld using Atto 647N. Hue encodes height according to the colour bar, while the saturation encodes fluorescence intensity. All scale bars correspond to 30 μm. Data recorded by Anna Chizhik, image published in [4].

and $\langle \tau_D \rangle$ was larger than the two standard deviations, a donor-acceptor distance was calculated according to

$$r = R_0 \left( \frac{\langle \tau_{DA} \rangle}{\langle \tau_D \rangle - \langle \tau_{DA} \rangle} \right)^{\frac{1}{6}}. \tag{4.11}$$

We are aware that the condition $\langle \tau_D \rangle - \langle \tau_{DA} \rangle > \sigma_{DA}, \sigma_D$ creates a bias towards smaller donor-acceptor distances $r$. However, as we will explain below, the aim of our FRET measurements was not to determine precise values $r$, but rather to provide an all-or-nothing check if FRET occurred. For this purpose, demanding a lifetime change that is larger than the uncertainty of the lifetime determination reduces the number of false positives.

**Results**

In order to study the developing actomyosin stress fiber structure and the architecture of focal adhesions, we fixed human mesenchymal stem cells (hMSCs) at different time points after seeding, followed by labeling of actin and vinculin and fluorescence imaging as described above. Figure 4.14 shows a time series of actin height maps acquired between one and 24 hours after seeding. While only few stress fibers are visible in the first image, they are more distant at later time points. In the early regime (1-6 h), stress fibers are mainly found at a distance of 150 nm or more from the silica spacer. Subsequently, they form tight connections to focal adhesions at low heights ($\sim 40$ nm) at the edges of the cells, spanning throughout the cell to heights above 180 nm with their other ends. The height maps are weighted by the fluorescence intensity to better visualize the positions of the actin filaments. These findings confirmed the success of our fixation, labeling and imaging techniques, and allowed us to now study focal adhesions in greater detail.

To this end, we acquired FLIM images of double labeled samples with reduced pixel size. Figure 4.15 focuses on areas at the edges of the cells, illustrating the temporal
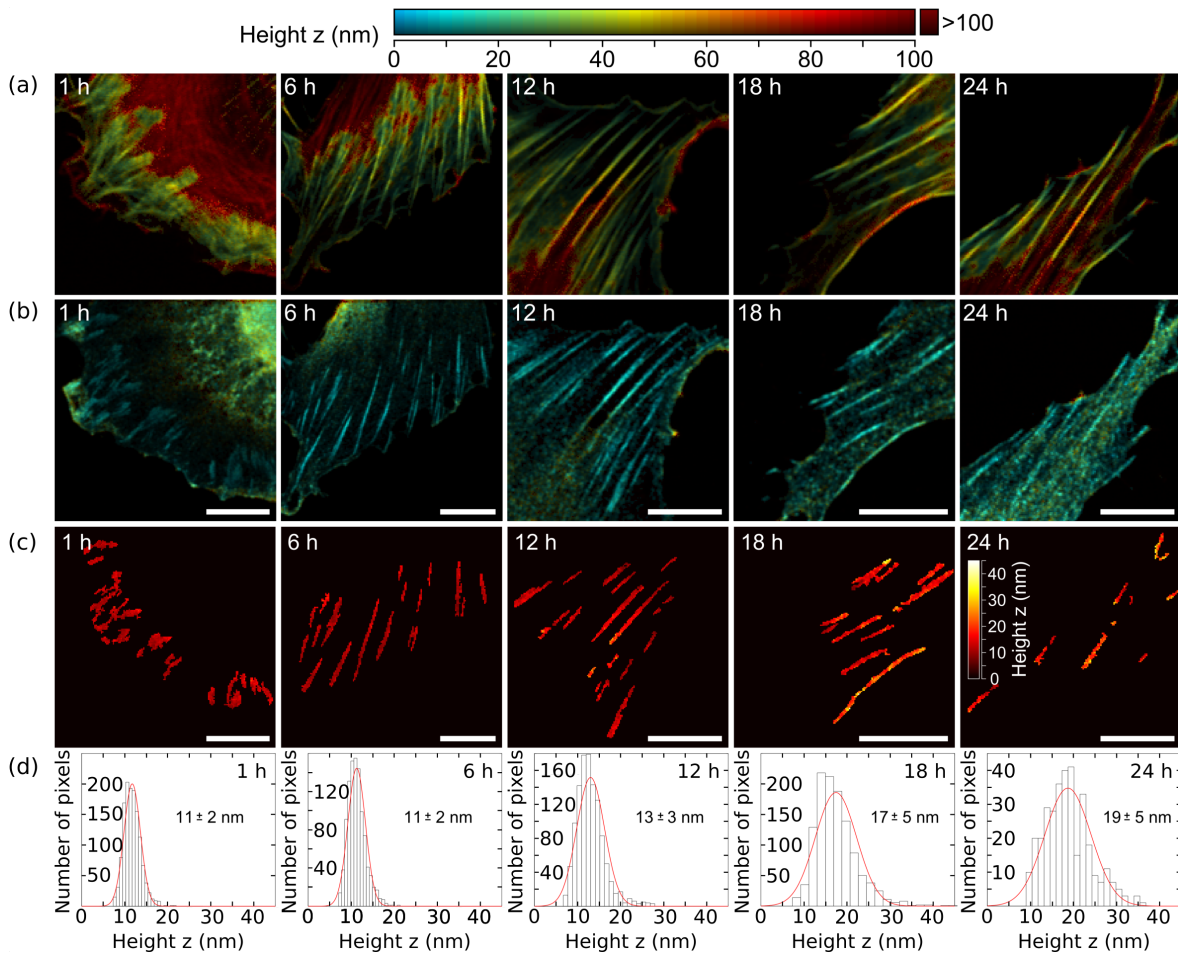
**Figure 4.15:** Detailed view of the temporal maturation of focal adhesions. (a, b) Intensity-weighted height maps of Atto 647N-actin and Atto 488-vinculin, respectively. Hue encodes height according to the colour bar at the top of the figure. (c) Identification of focal adhesions by detection of Atto 488-intensity clusters. Within these clusters, height values are colour coded according to the small colour bar in the rightmost panel. (d) Height distributions accumulated over all clusters detected within the cells from (c). In red, Gaussian distributions with the same mean, standard deviation and area are shown as a comparison. Cells were fixed on metal-coated cover slips at different time points after seeding. These are exemplary images, total numbers of cells imaged for the time points are: $1\,h - 4$, $6\,h - 3$, $12\,h - 4$, $18\,h - 4$, $24\,h - 7$. All scale bars correspond to $10\,\mu m$. Data recorded by Anna Chizhik, cluster identification performed by Narain Karedla. Image published in [4].

maturation of focal adhesions. Again, it becomes clear that actin stress fibers develop over time, moving close to the surface at the edge of the cell and reaching larger heights towards the center of the cell, see figure 4.15(a). While the Atto 488-vinculin signal is weak one hour after seeding, cells fixed at later time points show distinct regions with strong signal (b). Within these regions, vinculin is on average at a lower height than the actin ensembles, which is in agreement with the standard picture of stress fibers anchoring at focal adhesions [113]. By employing an intensity threshold and detecting clusters (as described for the FRET evaluation), vinculin height distributions were

obtained, see figure 4.15(c, d). These revealed a rather small spread of height values in the order of 2 nm (early stages) to 5 nm (24 h), as well as a slight increase of the average vinculin-substrate distance from 11 nm to 19 nm. We tested if these values were reproducible by determining the average height of vinculin in seven cells fixed 24 h after seeding, see figure 6.3 in the appendix. The result of $(20 \pm 8)$ nm agrees well with the value in figure 4.15(d). It is also within the range of previous works that found vinculin only slightly higher than the plasma membrane (90 nm-106 nm vs. $\sim 100$ nm [117] and $(54 \pm 6)$ nm vs. $(32 \pm 5)$ nm [113]).

When interpreting these results, we have to take into account that the obtained lifetime value for a single scan posision is always the average lifetime of all fluorophores present within the focal volume. Thus, if several vinculin molecules within a focal adhesion complex are fluorescently labeled, the reconstructed MIET height is actually the distance between the top of the silica spacer and the center of mass of the complex. Similarly, if a stress fiber is homogeneously labeled with Atto 647N, the reconstructed height corresponds to the center of the fiber[8]. In the first proof-of-principle study of MIET on microtubules [38], this relation was explicitly taken into account, and the lifetime values were directly related to the distance between the bottom of the microtubule and the top of the substrate. However, those were *in vitro* experiments under tightly controlled conditions. When labeling structures within a cell, it is conceivable that there might be unlabeled regions due to crowding – dye molecules, and especially the larger antibodies needed for staining, might simply be unable to access certain structures. Therefore, the center of mass of the labeled structure might not coincide exactly with the center of mass of the fluorophores. But in general, both should be very similar. Under this assumption, we suggest that the increased average height and also the larger spread of height values of vinculin during maturation of focal adhesions is caused by a slow growth of the vinculin aggregates.

For a detailed view of the relative positions of actin and vinculin, we generated height profiles along stress fibers. For hMCSs fixed 12 h or 24 h after seeding, we identified focal adhesions and anchoring stress fibers in the fluorescence intensity images, figure 4.16(a, b) and (d, e). By hand, we traced complete vinculin clusters (b, e) or actin stress fibers from the first significantly bright scan positions approximately up to the end of the focal adhesion (a, d). Resulting height profiles are depicted in figure 4.16(c) and (f). We derived the uncertainties $\Delta z$ of the single height values (shaded areas in the figure) by multiplying the inclination of the MIET height-versus-lifetime curves with the uncertainties $\Delta \tau$ of the lifetime values. The latter were estimated from the number $N$ of counts contained in the corresponding TCSPC curves according to $\Delta \tau \approx 4.8\tau/\sqrt{N}$. For our fitting procedures, this empirical formula was found to reproduce the spread of lifetime values of both simulations and bootstrapping of experimental data quite well. The height profiles confirm the general observations that vinculin is situated below actin, and that actin likely forms tighter bonds over time, resulting in lower height values at 24 h than at 12 h. Note the different scales of the horizontal and vertical axes in (c, f): Averaged over the three stress fibers per cell, we find very shallow inclination angles of $(0.15 \pm 0.01)°$and $(0.18 \pm 0.05)°$. Even though there is an apparent gap of at least 20 nm (12 h) or 10 nm (24 h) between actin and vinculin, it is likely that they are actually in

---

[8] Neglecting in both cases the height-dependent brightness of the molecules, which serves as weights during averaging.
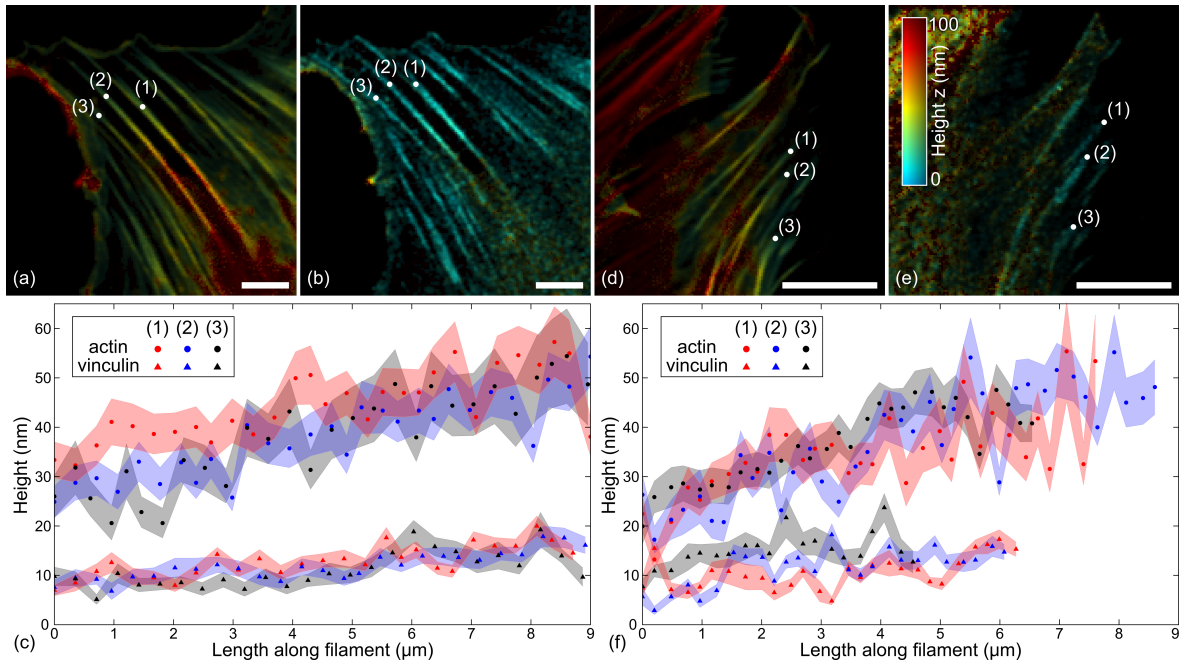
**Figure 4.16:** Height profiles of actin and vinculin in focal adhesions in hMSCs fixed 12 h (left) or 24 h (right) after seeding. FLIM images were converted to intensity-weighted height maps for Atto 647N-actin (a, d) and Atto 488-vinculin (b, e). Hue encodes height according to the colour bar shown in (e). In each cell, we chose three focal adhesions, which we traced by hand (starting points indicated by white dots). The height values along these traces are depicted in (c, f). Each colour represents one focal adhesion, circles mark actin heights, and triangles represent vinculin. Shaded regions indicate the height uncertainty $\Delta z$ determined as described in the main text. All scale bars correspond 10 µm. Data recorded by Anna Chizhik, image published in [4].

contact, keeping in mind that the reported heights are the positions of the centers of mass, as discussed above. In this case, many actin and vinculin molecules – and with them also Atto 647N and Atto 488 – should be in rather close contact, enabling FRET in unbleached samples.

We checked this assumption by measuring FRET[9] for cells fixed at different time points after seeding. Figure 4.17 shows lifetime measurements of the FRET donor Atto 488-vinculin before ($\tau_{\mathrm{DA}}$, top row) and after ($\tau_{\mathrm{D}}$, second row) bleaching of the FRET acceptor Atto 647N-actin. As explained before, donor-acceptor distances were only calculated for clusters with significant difference $\tau_{\mathrm{DA}} - \tau_{\mathrm{D}}$, they are presented as intensity-weighted distance maps in the third row. The distance distributions in the bottom row reveal no significant change in the calculated distances over time. This reveals the proximity between actin and vinculin in focal adhesions that is stable throughout the maturation process. Furthermore, FRET was not observed in all focal adhesions, or sometimes only in distinct small areas. While this might be partly due to our strict criterion of significant lifetime changes compared to the measurement uncertainty, we believe that it also illustrates the complexity and heterogeneity of these

---

[9] For these measurements, cells were seeded on uncoated glass cover slips. In this way, all lifetime changes could be assumed to originate from FRET, not MIET.
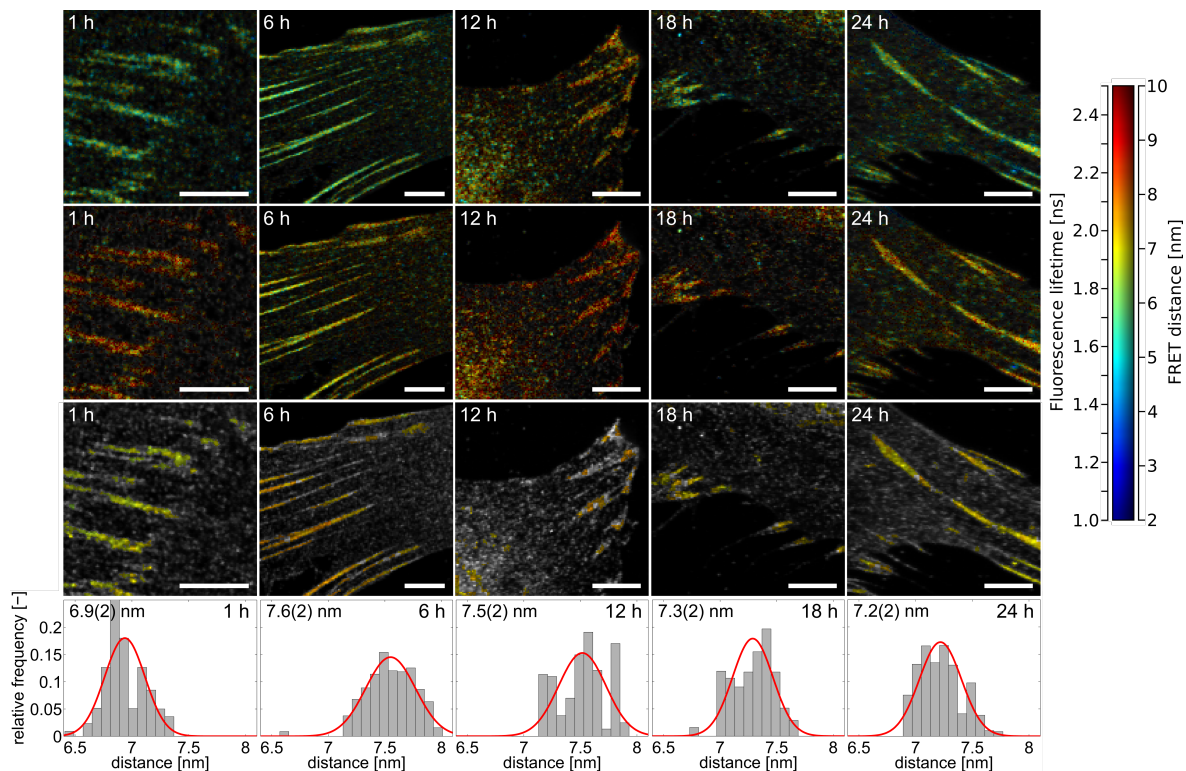
**Figure 4.17:** FRET measurements in focal adhesions. Lifetime images of the FRET donor Atto 488-vinculin before (top row) and after (second row) bleaching of the FRET donor Atto 647N-actin. For clusters chosen as explained in the text, donor-acceptor distances were calculating using the quantitative theory of FRET, assuming that each donor interacts with exactly one acceptor (third row). Distributions of the resulting distances are shown in the bottom row, together with Gaussian fits. Cells were fixed, stained and imaged 1 h, 6 h, 12 h, 18 h and 24 h after seeding. All scale bars correspond to 5 μm. Note that the assumption of one-on-one interactions is likely not true and these distances are thus a lower threshold, see text for a discussion. Data recorded by Anna Chizhik and analysed by Narain Karedla. Image published in [4].

supramolecular structures.

One has to be careful, though, when trying to draw conclusions about the exact distances between Atto 488 and Atto 647N from the results presented in figure 4.17. The theory that lead to equation (4.11) assumed an interaction of exactly one donor and one acceptor. Contrarily, in a focal adhesion with an anchoring stress fiber, there are many fluorescently labeled actin and vinculin molecules in close proximity. If a donor can interact with $N$ acceptors, the different decay rates simply add up:

$$\frac{1}{\tau_{\mathrm{DA}}} = k_{\mathrm{r}} + k_{\mathrm{nr}} + \sum_{j=1}^{N} k_{\mathrm{FRET,\,j}}(r_j), \tag{4.12}$$

where $r_j$ is the distance between the donor and the acceptor $j$. In the most general case without any prior knowledge, measuring just one data set $(\tau_{\mathrm{DA}}, \tau_{\mathrm{D}})$ does not allow to reconstruct the positions of the acceptors. However, we can estimate the effect of the
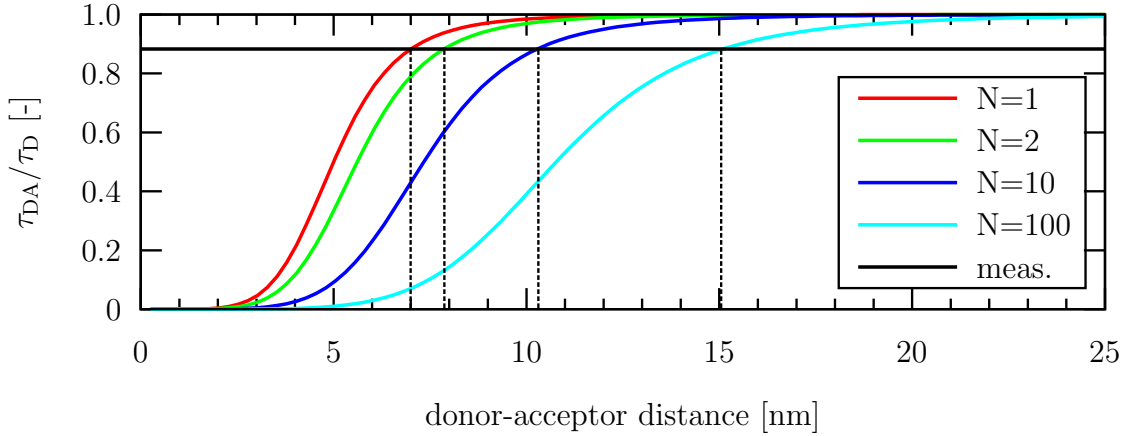
**Figure 4.18:** Influence of additional acceptors on FRET donor lifetime. The graph depicts the lifetime ratio $\tau_{DA}/\tau_D$ assuming the presence of $N$ acceptors at a distance $r$ from the donor. The black horizontal line shows the average measured lifetime ratio for hMSCs fixed at different time points. The dashed black lines indicate which donor-acceptor distance would result in this value for the given numer of acceptors. Even for $N = 100$, $r$ is still within approximately $15\,\mathrm{nm}$. Note that in a real situation, the acceptors likely all have different distances from the donor.

presence of multiple acceptors by assuming that they all have the same distance from the donor, $r_j = r \ \forall j$. Then, all FRET rates are equal. Inserting this assumption and the relation (4.12) in the calculations at the beginning of this section, equation (4.11) transforms to:

$$r_N = R_0 \cdot \left( \frac{\tau_{DA}}{N\left(\tau_D - \tau_{DA}\right)} \right)^{\frac{1}{6}}. \tag{4.13}$$

Thus, if the number of acceptors increases by a factor of $N$, the donor-acceptor distance calculated from the same data set decreases by the sixth power of $N$. Alternatively, one can study this effect from the perspective of the difference in the measured lifetime when increasing $N$ and keeping $r$ constant. Starting from the same considerations as before, we find:

$$\frac{\tau_{DA}}{\tau_D} = \frac{1}{1 + N\left(\frac{R_0}{r}\right)^6}. \tag{4.14}$$

This relation is illustrated in figure 4.18, where different colours correspond to various values of $N$. In black, we show the average ratio $\tau_{DA}/\tau_D \approx 0.88$ measured in the cells in figure 4.17. Dashed vertical lines indicate which donor-acceptor distance $r$ would result in this value for different numbers of acceptors if all acceptors had the same distance from the donor. Even for $N = 100$, $r$ is still in the order of $15\,\mathrm{nm}$. Thus, we interpret the value of $r \approx 7\,\mathrm{nm}$ from figure 4.17 as a lower bound, and repeat our previous conclusion that the occurrence of FRET proves the general proximity of at least some populations of actin and vinculin in stress fibers. This is in accord with

the results of Kanchanawong *et al.* who found that approximately half of the vinculin molecules overlap with actin [113].

## Summary

On a methodological level, we have established dual-colour MIET as a powerful technique for the axial co-localization of two fluorescently labeled species with nanometer precision. We have furthermore shown that the occurrence of FRET in this situation can be a benefit rather than a problem by proving the existence of close contacts within distributions of molecules at different heights.

On a biological level, we investigated the interaction of actin stress fibers and focal adhesion complexes (represented by vinculin). We were able to confirm the standard model of stress fibers anchoring at focal adhesions, and thus of actin being located at larger distances from the surface than vinculin. We could observe the maturation of focal adhesions, including the growth of vinculin aggregates and the gradual lowering of the actin filaments, probably indicative of the formation of tighter bonds. The occurrence of FRET suggests at least a partial overlap of the actin and vinculin populations. Finally, by making use of the nanometer resolution to reconstruct height profiles of focal adhesions, we found that actin emerges from the focal adhesions under a very shallow angle of below one degree.

## 4.1.4 Three-dimensional reconstruction of nuclear envelope architecture using dual-colour MIET imaging

After having studied the plasma membrane itself in section 4.1.2 and membrane-associated focal adhesions in section 4.1.3, we now move on to a structure that is located higher within the cell, and thus approaches the operating limit of MIET: the nuclear envelope. The results were published in:

[3] A. M. Chizhik, **D. Ruhlandt**, J. Pfaff, N. Karedla, A. I. Chizhik, I. Gregor, R. H. Kehlenbach, and J. Enderlein, "Three-dimensional reconstruction of nuclear envelope architecture using dual-color metal-induced energy transfer imaging," *ACS Nano*, vol. 11, no. 12, pp. 11839–11846, 2017. A. M. C., **D. R.** and J. P. contributed equally. **D. R.** provided analysis tools and analysed the MIET data.

**Motivation**

The cells of eukaryotes are characterized by having a cell nucleus that is enclosed by membranes, containing most of the cell's genetic material. The membrane system is called the *nuclear envelope*, it consists of the *inner nuclear membrane* (INM) and the *outer nuclear membrane* (ONM). Employing electron microscopy, it has been shown that the space between INM and ONM has a typical width of 30 nm - 50 nm [118]. The protein composition of both membranes is not identical – while some proteins can be found on either membrane, others localize preferentially at the INM or the ONM. This topic has been studied using a large variety of techniques, from immunoelectron microscopy to differential permeabilization of membranes, to name just two. Recently, a study by Mudumbi *et al.* investigated the concentration ratios between the INM and the ONM of several proteins using single-point fluorescence recovery after photobleaching [119]. The proteins they chose localized preferentially at the ONM, but were also present at the INM with a ratio of 1:10 to 1:2. However, for several hundred proteins, it is known that they are found preferentially at the INM [3]. One of the best-described members of this class is the lamina-associated polypeptide $2\beta$ (LAP2$\beta$), which we chose as a representative of the INM in this work.

An exchange of molecules between the nucleus and the surrounding cytoplasm requires a transport mechanism across the nuclear envelope. They key players for this process are *nuclear pore complexes* (NPCs), large protein complexes which consist of a core that is embedded between the INM and the ONM and nuclear extensions forming a basket-like structure (see figure 4.19(a) for a schematic drawing). The height of the core structure is approximately 80 nm, or 150 nm - 200 nm when the extensions are taken into account, while the diameter is in the order of 120 nm. Like the nuclear envelope, the NPC contains proteins which localize either on both sides, or are found preferentially on one of them. In this study, we concentrated on Nup358, which is known to be specific to the endoplasmatic side of the NPC (i.e. on the same side as the ONM). More details on NPCs and references to the relevant literature can be found in the introduction of [3]. NPCs have attracted much attention both due to their biological relevance, but also because their well-defined structure is an interesting model system to test super-resolution methods, e. g. demonstrating the high lateral resolution of dSTORM

of about 20 nm [120]. However, what has been missing so far is a study of the NPC focusing on the axial localization of different components using optical microscopy.

We wanted to close this gap by studying the three-dimensional architecture of the nuclear envelope not just at one point (as in [119]), but rather along the whole basal part of the nuclear envelope. To this end, we chose Lap2$\beta$ and Nup358 as proteins with distinct localization in the INM and ONM, respectively, and determined their axial positions with nanometer resolution using MIET. Furthermore, we seized the opportunity to study the influence of sample parameter uncertainties (such as intricate refractive index compositions) on the localization accuracy of MIET. The nuclear envelope is more suited for this endeavour than, for example, the basal cell membrane, since it is located deeper within the cell and thus in a more complex environment.

## Methods

We studied the nuclear envelope in HeLa cells. The proteins of interest were fluorescently labeled using primary and secondary antibodies, resulting in the labeling of Lap2$\beta$ and Nup358 with the green dye Alexa Fluor 488 (Molecular Probes by Invitrogen [now Thermo Fisher Scientific, Waltham, MA]) and the far red dye Alexa Fluor 633. For compactness of notation, we call the resulting labeling schemes Lap2$\beta$-A488 etc. For comparison, we measured samples with Lap2$\beta$-A633 and Nup358-A488 (*samply type I*) as well as samples with switched labels, i.e. Lap2$\beta$-A488 and Nup358-A633 (*sample type II*), see schematic drawings in figure 4.19(a) and 4.20(a). As already mentioned in the previous section, coupling a fluorescent dye to an antibody and introducing it into a cell can change both its radiative and non-radiative deexcitation rate, and thus its quantum yield and free space lifetime. Since different antibodies were used for all four labeleing schemes, we measured four free space lifetimes inside cells seeded on uncoated glass slides and calculated four quantum yield values. This will be explained in detail below ("Results II: methodological"). More details on the cell culture and staining procedures can be found in our publication [3].

The FLIM setup was exactly the same as described for the study of focal adhesions in the previous section. Due to the similar absorption and emission spectra of Atto 488 / Alexa Fluor 488 and Atto 647N / Alexa Fluor 633, even the same filters and the same excitation wavelengths could be used. The substrates for MIET measurements were improved slightly: We found that 2 nm titanium (instead of the 3 nm used when studying focal adhesions) sufficed for improving adhesion of gold to glass but lead to less absorption of fluorescence when traversing the substrate. Thus, the composition of the substrates was: glass cover slip (170 μm), 2 nm Ti, 15 nm Au, 2 nm Ti, 10 nm $SiO_2$. Since we were studying structures further away from the substrate, the silica spacer was not needed to avoid quenching of the dyes. We included it nevertheless to ensure that the same sample preparation procedures as for glass surfaces (as used for standard immunofluorescence measurements) could be used. For acquiring FLIM images, we used our three-step procedure established in [4] to exclude any occurrence of FRET: First, the fluorescence of the potential FRET acceptor A488 was measured. Second, a powerful red laser was used to completely bleach A488. Third, an image of the potential FRET donor A633 was acquired. Fluorescent lifetimes were extracted from the TCSPC data using a distributed tailfit as described in section 3.5, and converted to height

values using four distinct MIET curves (calculated for optical parameters of the four labeling schemes).

As a control, we also explicitly measured FRET between Lap2$\beta$-A488 and Nup358-A633 (sample type II) in HeLa cells on untreated glass cover slips. Both the experimental procedure and the data analysis were identical to the procedure described in section 4.1.3, and even the Förster radius of the two dye pairs was found to be the same, namely $R_0 = 5.0\,\text{ns}$.

### Results I: biological

Initially, we tested our staining procedure by acquiring fluorescence images with a standard confocal microscope. The axial resolution of a confocal microscope, and thus the size of the focal volume along the optical axis, can be described by [121]

$$\frac{\lambda}{n(1 - \cos\theta)}, \tag{4.15}$$

where $\lambda$ is the wavelength of the light, $n$ is the refractive index of the medium, and $\theta$ is the maximum angle of light collection of the objective. The fluorescence intensity images 4.19(c, d) and 4.20(c, d) were acquired using a water-immersion objective with numerical aperture 1.3 (i.e. $n = 1.33$, $\theta \approx 78,°$), all following FLIM images were taken with an oil-immersion objective with numerical aperture 1.49 (i.e. $n = 1.52$, $\theta \approx 79\,°$). This results in axial resolutions of $0.95\,\lambda$ and $0.82\,\lambda$, respectively, or roughly $500\,\text{nm}$ to $700\,\text{nm}$. The nucleus of a HeLa cell can be approximated as an oblate spheroid, where the diameter in the equatorial plane is between $10\,\mu\text{m}$ and $20\,\mu\text{m}$ and the height is around $2\,\mu\text{m}$ to $3\,\mu\text{m}$ [122]. Thus, when the focal plane of the microscope coincides with the equatorial plane of the nucleus, a staining of the INM or the ONM should result in ring-like bright structures. These can be clearly observed for Nup358-A488 and Lap2$\beta$-A633 in figure 4.19(c) and (d), respectively, and for Nup358-A633 and Lap2$\beta$-A488 in figure 4.20(c) and (d). Besides proving the quality of the labeling, these images nicely demonstrate the sectioning capabilites of the confocal microscope: Only light originating from the focal plane is detected with high efficiency. When focusing on the top of the substrate instead of the equatorial plane, we can therefore expect to excite only fluorophores on the basal side of the nuclear envelope.

With this, we turned to the study of exactly this part of the nuclear envelope using MIET. Figure 4.19(e) and (f) show the fluorescence intensity of Nup358-A488 and Lap2$\beta$-A633 (sample type I), respectively, while the panels (g) and (h) depict the corresponding height distributions. The same can be found for sample type II in figure 4.20(e)-(h). In order to estimate the height range in which the two proteins can be found, the height values of six cells per sample type were compiled in the histograms on the left side of figure 4.21. All localizations occurred above $50\,\text{nm}$, with a significant proportion of molecules found even above $150\,\text{nm}$. Thus, these structures are at the limit of the operating range of MIET, where the lifetime-versus-height curve is very flat (see e.g. figure 4.25). This results in a dramatic loss of axial localization accuracy compared to measurements closer to the metal layer, where accuracies of $2.5\,\text{nm}$ to $3\,\text{nm}$ have been demonstrated [38–40]. Using the same estimate for the lifetime uncertainty $\Delta\tau$ as in the study of focal adhesions, $\Delta\tau \approx 4.8\tau/\sqrt{N}$, where $N$ is the number of counts
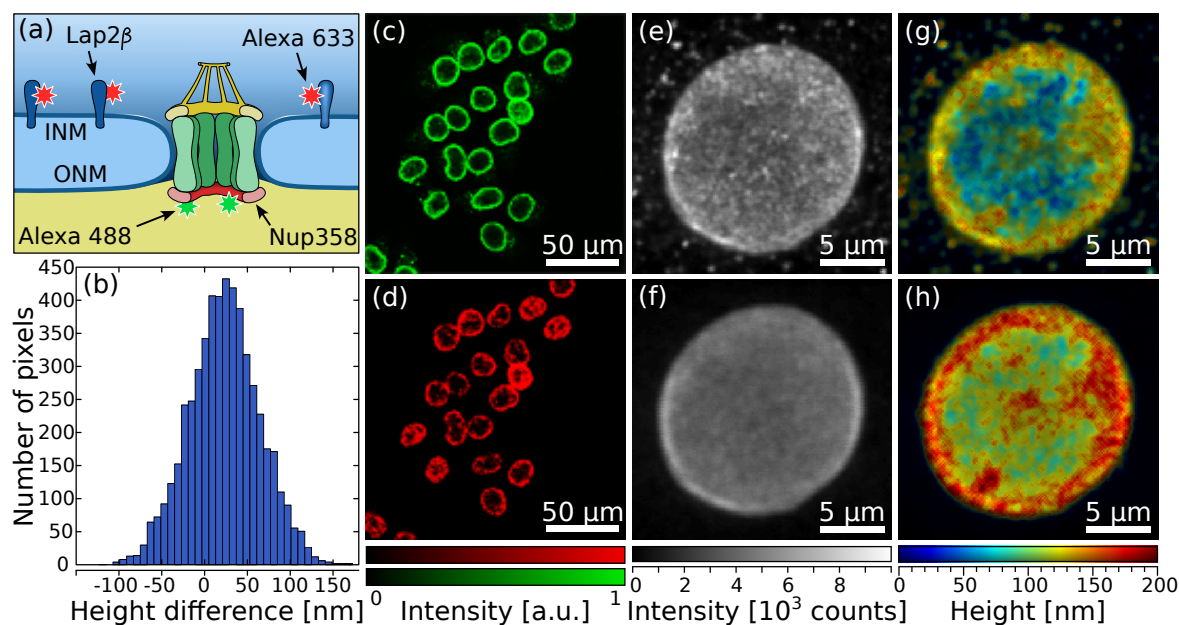
**Figure 4.19:** Sample type I: (a) Schematic depicting the labeling scheme, Lap2$\beta$-A633 and Nup358-A488. In the following, the top row shows results for Nup358-A488, the bottom row for Lap2$\beta$-A633. (c, d) Confocal fluorescence intensity images of the equatorial plane of the nuclear envelope showing ring-like structures. (e, f) Fluorescence intensity of FLIM images of the basal part of the nuclear envelope. The corresponding lifetimes were converted to the height maps shown in (g, h). The pixelwise difference between the height values of INM and ONM (i.e. "(h) minus (g)") are histogrammed in panel (b). Data recorded by Janine Pfaff (c-d) and Anna Chizhik (e-h), schematic (a) created by Alexey Chizhik. Image published in [3].
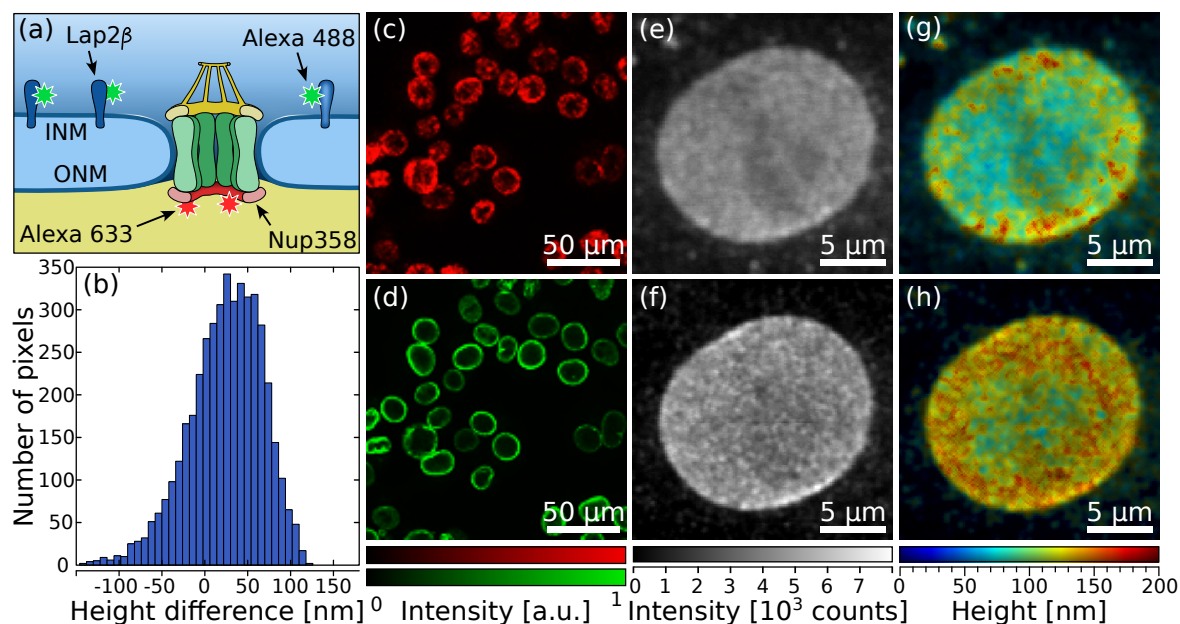


**Figure 4.20:** Sample type II: (a) Schematic depicting the labeling scheme, Lap2$\beta$-A633 and Nup358-A488. (c-h) Everything switched compared to sample type I. Again, the pixelwise distance between the height values of the INM and ONM are histogrammed in (b). Same contributions as above, published in [3].

162
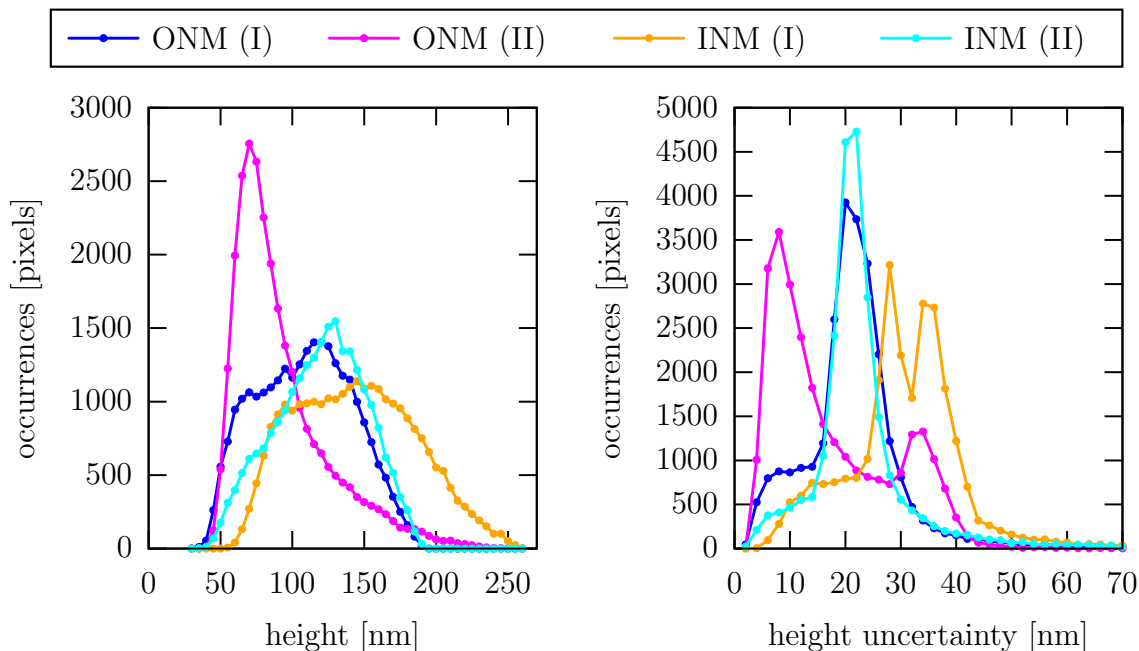
**Figure 4.21:** Overview of determined height values (left) and estimated height uncertainties (right) for the two membranes (ONM and INM) in both samples types (I and II). For each sample type, six cells were examined, the pixels belonging to the cells (rather than the background) identified, and the height values of Lap2$\beta$ and Nup358 from these pixels histogrammed. Height uncertainties $\Delta z$ were determined by multiplying the local slope of the MIET curve $\partial z(\tau)/\partial\tau$ with estimated lifetime uncertainties $\Delta\tau$. Note that the height uncertainty is larger for large $z$ due to the shape of the MIET curve. Data recorded by Anna Chizhik, image previously unpublished.

in the TCSPC histogram, we calculated estimates for the height uncertainty $\Delta z$ (right side of figure 4.21). While the details of the curves depend on the individual intensities, the general trend is as expected: The more large height values occur in a sample type, the more large uncertainties are observed, ranging from $\Delta z \approx 5$ nm to $\Delta z \approx 50$ nm. Despite these large uncertainties of the height values of individual pixels, the two height images in figure 4.19(g) and (h) – which, assuming that the two membranes closely follow each other, can be seen as two independent measurements of the same topology – show similar large-scale features (peaks and valleys). The same behaviour can be observed in figure 4.20(g) and (h). This finding indicates that while the *spread* of obtained height values is large, the *mean* is probably correct. In accordance with this theory, the histograms of the single-pixel height differences between Lap2$\beta$ and Nup358 in figures 4.19(b) and 4.20(b) are very broad, even containing approximately 20 % negative values, but have a maximum around the expected value of 30 nm to 50 nm. In order to investigate this point further, better statistics had to be obtained. To this end, we repeated the MIET measurements for six cells of each sample type. The corresponding INM-ONM distance distributions are shown as grey-scaled strips in figure 4.22. For each histogram, we determined the mean value (indicated as open red circles) and standard deviation (horizontal red lines). From these results, we calculated the
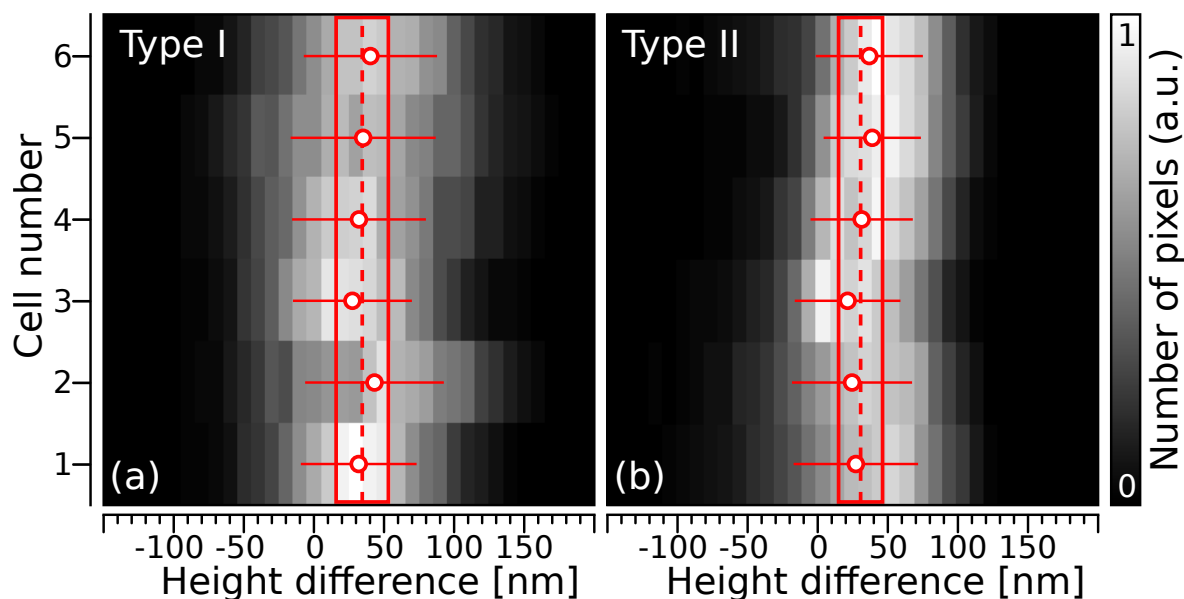
**Figure 4.22:** Comparison of INM-ONM distance for six different cells per sample type. The same kind of distance distribution as in figure 4.19(b) and 4.20(b) was obtained for each cell. Instead of twelve classical histograms, we chose two greyscale images to compactly represent these distributions. The mean of each distribution is indicated as an open red circle. Together with the standard deviations of the individual distributions (horizontal red lines = mean ± std. dev.), these mean values were used to determined the weighted arithmetic mean for each sample type (dashed red line). The standard deviation of the weighted arithmetic mean is represented by a red rectangle. The resulting average distance between INM and ONM is $(35 \pm 19)\,\text{nm}$ for sample type I and $(31 \pm 16)\,\text{nm}$ for sample type II. Data recorded by Anna Chizhik, image published in [3].

weighted arithmetic mean (dashed red line) and its standard deviation (red rectangle). For sample type I, we found a distance of $35 \pm 19\,\text{nm}$, and for sample type II $31 \pm 16\,\text{nm}$. These values are in excellent agreement with each other, but also with previous electron microscopy studies [118, 123] and the FRAP measurements in [119][10]. It should be pointed out that the electron microscopy studies [118, 123] showed local variations in the distance between the two membranes. Therefore, when histogramming the height difference between INM and ONM for the whole basal part of the nuclear envelope, some spread would be expected even for a technique with perfect axial localization.

Some further information can be gained from FRET measurements. We were able to observe a change in the donor lifetime in the presence of the acceptor as compared to the absence of the acceptor, see figure 6.4 in the appendix. The calculated distance between donor and acceptor was approximately 8 nm, however, as explained in the

---

[10]We remark that the latter, which is also an optical microscopy technique, reports a much smaller standard deviation than we do (2 nm - 4 nm). That is because Mudumbi *et al.* measured in the equatorial plane of the nucleus, where the distance between both membranes corresponds to a lateral distance that can be determined precisely using single-molecule localization microscopy. Furthermore, they only measure distances in a single spot (diameter 500 nm) and thus do not capture the heterogeneity of the distances across the whole basal part of the nuclear envelope as we do, see below.

previous section, this value is just a lower bound due to the likely interaction of one donor with several acceptors. We interpret the occurrence of FRET as a sign that a small proportion of Lap2$\beta$ is located at the ONM. This is feasable, since Lap2$\beta$ is synthesized in the cytoplasm and has to be translocated to the INM via the nuclear pore complex.

**Results II: methodological**

In the first proof-of-principle studies of MIET, labeled structures were either attached to the surface but immersed in a buffer [38] or single molecules were deposited on a surface and then surrounded by air [40]. In both cases, there was no ambiguity in the choice of refractive indices used to determine the MIET calibration curve. When labeling the basal cell membrane during the epithelial-to-mesenchymal transition (section 4.1.2), since the label was on the extracellular side of the membrane[11], using the refractive index of the buffer seems reasonable. However, one might ask if the presence of the cell close to the dye also has an impact on its lifetime. Since the nuclear envelope is located within the cell, and thus in a much more complex environment, the question arises which impact the modeling of this environment has on the resulting MIET curves.

We studied this subject by calculating MIET curves for four different model scenarios of increasing complexity, see figure 4.23. As a first step, we conducted an extensive analysis of the literature concerning refractive indices of cellular components. The refractive index of the plasma membrane has been reported as $n = 1.46$ [124] or $n = 1.48$ [125], whereas the refractive index of the nuclei of four different cells lines, including HeLa cells, has been determined as $n = 1.36$ [126]. Values found for the cytosol range between 1.36 and 1.39 [124, 127, 128]. Finally, the average refractive index of a HeLa cell was reported to be $n = 1.37$ [126, 129]. For MIET measurements, cells were mounted using Mowiol 4-88 (Calbiochem, Merck, Darmstadt, Germany). For this mounting medium, no official refractive index value exists, it seems to vary from batch to batch. We found reports of refractive indices between 1.41 and 1.49, and used the "worst case" $n = 1.49$ (i.e. largest deviation from the average cell value of $n = 1.37$). These findings were then used to construct four model systems:

a) The MIET substrate, followed by a general cell halfspace with refractive index $n = 1.37$.

b) The MIET substrate, followed by a 10 nm thick layer of Mowiol and then a general cell halfspace with refractive index $n = 1.37$.

c) The MIET substrate, followed by a general cell halfspace with refractive index $n = 1.37$ almost everywhere, except for a 6 nm thick lipid bilayer ($n = 1.46$) situated 2 nm below (INM) or above (ONM) the fluorophore.

d) The MIET substrate, followed by a 10 nm thick layer of Mowiol, a 6 nm thick lipid bilayer ($n = 1.46$) representing the plasma membrane, and then the same cell / nuclear membrane system as in c).

---

[11]In those experiments, we incubated cells with dye-containing culture medium for five minutes and then replaced the staining solution by normal cell culture medium. Since we did not permeabilze the membrane, no (or only very little) dye should have been able to penetrate the cell membrane.
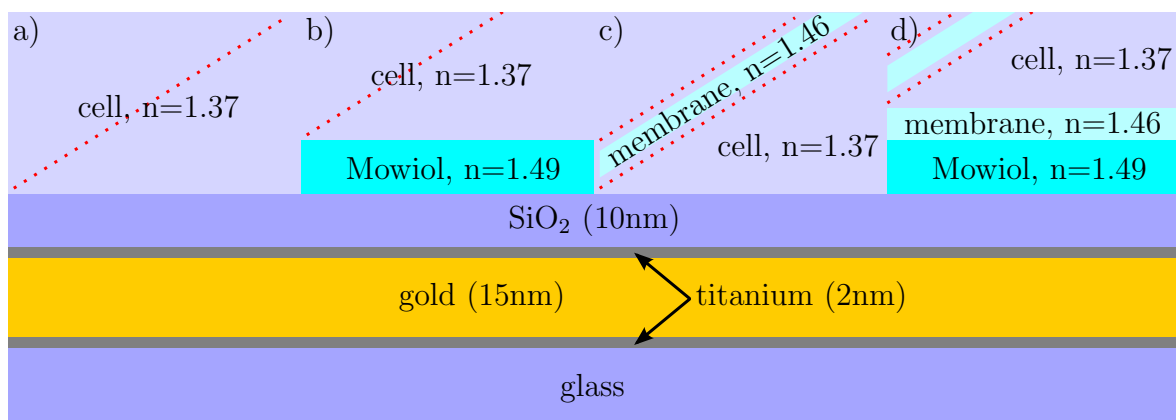
**Figure 4.23:** Four different models of the sample structure used to study the impact of the refractive index on the calculated height difference between INM and ONM. Red dots represent the dye molecules situated at different heights whose relative fluorescence lifetime makes up the MIET curve. The cell is modeled as a homogeneous space with $n = 1.37$ (a, b) or including the 6 nm thick nuclear membrane (c, d). A layer of 10 nm mounting medium (Mowiol) is included in models b) and d), the latter also takes the cell membrane into account. In c) and d), the nuclear membrane is situated 2 nm above or below the fluorophore for the ONM and INM, respectively (visualized by two rows of red dots). Image published in [3].

Since the refractive indices of nucleus and cytosol are rather similar, and partly overlap with the average refractive index reported for a HeLa cell, these are set to $n = 1.37$ in all four models. Both a) and c) assume that either the extracellular matrix possibly secreted by the cell prevents Mowiol from reaching the space between the cell and the substrate, or that the Mowiol layer does not have a large influence on the MIET curve because it is a just a thin, transparent layer that is not in the immediate vicinity of the fluorophore. The presence of the plasma membrane that surrounds the whole cell is ignored in all models except d), however, its refractive index is likely very similar to that of Mowiol, as we saw above. Furthermore, the presence of the nuclear membrane close to the fluorescent labels is only taken into account in c) and d).

After having obtained the MIET calibration curves for all four situations, we repeated the calculation of the average distance between inner and outer nuclear membrane for twelve different cells as described above and as shown in figure 4.22. The resulting average distances are summarized in the following table:

| sample | model a) | model b) | model c) | model d) |
|---|---|---|---|---|
| sample type I | $(35 \pm 18)$ nm | $(35 \pm 18)$ nm | $(34 \pm 19)$ nm | $(35 \pm 18)$ nm |
| sample type II | $(31 \pm 16)$ nm | $(29 \pm 15)$ nm | $(31 \pm 16)$ nm | $(29 \pm 15)$ nm |

For each sample type, the calculated average distance between INM and ONM changes by 1 nm to 2 nm when different refractive index structures are taken into account. Compared to the standard deviation of 15 nm to 19 nm, this is not a significant change. We therefore decided to use the simplest model, a homogeneous halfspace with refractive index $n = 1.37$, for all calculations.

Besides having a model of the sample, the second prerequisite for calculating a MIET

curve is knowing the optical parameters of the fluorophore. We measured the spectra of Alexa Fluor 488 and Alexa Fluar 633 inside the HeLa cells using a spectrograph equipped with a CCD camera, thus we were able to correctly average the MIET curves of the relevant wavelengths. The determination of the free space lifetime and quantum yield of the fluorophores is more challenging. In [107], Chizhik *et al.* used a nanocavity (see section 4.3) to study Alexa Fluor 488 in water, resulting in $\tau_{0,\text{ A488, water}} = 4.4\,\text{ns}$ and $\Phi_{\text{A488, water}} = 0.94$. For this work, we performed the same type of measurement for Alexa Fluor 633 in water and found $\tau_{0,\text{ A633, water}} = 3.2\,\text{ns}$ and $\Phi_{\text{A633, water}} = 0.59$. When conjugating the dyes to antibodies and introducing them into the HeLa cells, we expected both the radiative and non-radiative deexcitation rates to change, leading to a different quantum yield and free space lifetime. However, we could only directly measure the changed free space lifetime $\tau_{0,\text{ cell}} = (k_{\text{r, cell}} + k_{\text{nr, cell}})^{-1}$ by acquiring FLIM images of labeled HeLa cells on untreated glass cover slips. In order to find the quantum yield of the conjugated dye inside the cell, either $k_{\text{r, cell}}$ or $k_{\text{nr, cell}}$ had to be determined independently. In the previous work on focal adhesions (section 4.1.3), we assumed a functional relation between the refractive index and the radiative rate. Because $k_{\text{r}}$ in water can be determined as $\Phi_{\text{water}}/\tau_{0,\text{ water}}$, this allows to calculate the radiative rate in the cell. Thus, we started with the ansatz

$$k_{\text{r}}(n) = f(n)k_{\text{r}}^{\text{vac}}, \tag{4.16}$$

where $k_{\text{r}}^{\text{vac}}$ is the radiative rate in vacuum ($n = 1$). Several theoretical models provide possible forms for $f(n)$. Firstly, all our MIET calculations are based on the assumption that $k_{\text{r}}$ is proportional to the total energy flux per time, $S_{\text{tot}}$. Since for a dipole in free space, $S_{\text{tot}} = cnk_v^4p^2/3$ (see equation 2.162), this is equivalent to

$$f_{\text{lin}}(n) = n. \tag{4.17}$$

This relation was suggested already in 1976 by Nienhuis and Alkemade [130]. The fact that Drexhage's original experiments as well as several MIET studies from our group (including the work on focal adhesions in section 4.1.3) can be explained by the assumption $k_{\text{r}} \propto S_{\text{tot}}$ strongly supports this theory. However, these studies always examined the change in $S_{\text{tot}}$ due to the presence of an interface, and thus to a change in the angular distribution of radiation, never due to a change of the medium directly surrounding the dipole itself. Furthermore, the theory leading to equation 2.162 is based on a macroscopic version of Maxwell's equations. In particular, the dielectric material surrounding the emitter is modeled as a homogeneous medium (neglecting the fact that it actually consists of individual molecules) and the fluorophore itself is modeled as a point dipole (neglecting the size of the molecule). As a consequence, the space where the fluorophore resides is assumed to have the refractive index of the surrounding dielectric – *de facto*, the theory pretends the dipole emitter is not there. In reality, the dipole emitter displaces the neighbouring molecules of the dielectric, leading to a changed electric field. This effect can be taken into account by introducing the concept of a *local field* at the position of the dipole emitter, which differs from the field predicted by the macroscopic Maxwell's equations. A common approach to derive these local field corrections is to assume a spherical cavity inside the dielectric

surrounding the emitter. Depending on the nature of the cavity, two limiting cases have been described [131, 132]: In the *empty-cavity model*, the dipole emitter is inside a real, empty cavity, and one finds

$$f_{\text{e.c.}} = n \left( \frac{3n^2}{2n^2 + 1} \right)^2 . \tag{4.18}$$

Contrarily, if the cavity is only *virtual*, i.e. filled with a medium having the same average polarizability density as the surrounding dielectric, but the dipoles inside the cavity do not contribute to the local field, then

$$f_{\text{v.c.}} = n \left( \frac{n^2 + 2}{3} \right)^2 . \tag{4.19}$$

Alternatively, an expression for the spontaneous radiative emission rate can be derived starting from the Einstein coefficients for absorption and emission, and taking into account the refractive index of the surroundings as well as the broad absorption and emission spectra of molecules. The result [133] is known as the *Strickler-Berg equation*:

$$k_{\text{r, SB}}(n) = 2.88 \cdot 10^{-9} \, n^2 \, \frac{\int I(\tilde{\nu}) \mathrm{d}\tilde{\nu}}{\int I(\tilde{\nu}) \tilde{\nu}^{-3} \mathrm{d}\tilde{\nu}} \int \frac{\varepsilon(\tilde{\nu})}{\tilde{\nu}} \mathrm{d}\tilde{\nu}, \tag{4.20}$$

where $\tilde{\nu} = \lambda^{-1}$ is the wavenumber, $I(\tilde{\nu})$ is the fluorescence spectrum and $\varepsilon(\tilde{\nu})$ is the (wavelength-dependent) molar extinction coefficient. Thus, if absorption and emission spectra are independent of the refractive index, the Strickler-Berg equation suggests

$$f_{\text{SB}}(n) = n^2. \tag{4.21}$$

Some efforts have been made to determine experimentally which of these four equations truly describes the impact of the refractive index on the radiative rate of real fluorophores. Rikken *et al.* dissolved fluorescent europium $Eu^{3+}$ complexes in different liquids and found the lifetime values following the empty cavity model. However, exchanging the solvent might change not only the refractive index (and thus $k_{\text{r}}$), but also the non-radiative rate. Therefore, Schuurmans *et al.* [132] studied an $Eu^{3+}$ complex in dense supercritical $CO_2$, where they varied $n$ by increasing the pressure, which is not expected to influence $k_{\text{nr}}$. They also found excellent agreement with the empty-cavity model. Contrarily, Suhling *et al.* were able to explain their measurements of green fluorescent protein in various solvents by the Strickler-Berg equation [134]. Furthermore, they cite a number of studies that also found a quadratic dependence of the radiative rate of tryptophan residues or anthracene derivatives on $n$. This apparent contradiction can be resolved by comparing the four different models in the relevant refractive index range, see figure 4.24. While the linear model corresponds to the smallest change and the virtual cavity model predicts a drastic increse of $k_{\text{r}}$ with $n$, the empty-cavity model and the Strickler-Berg equation give almost identical results.
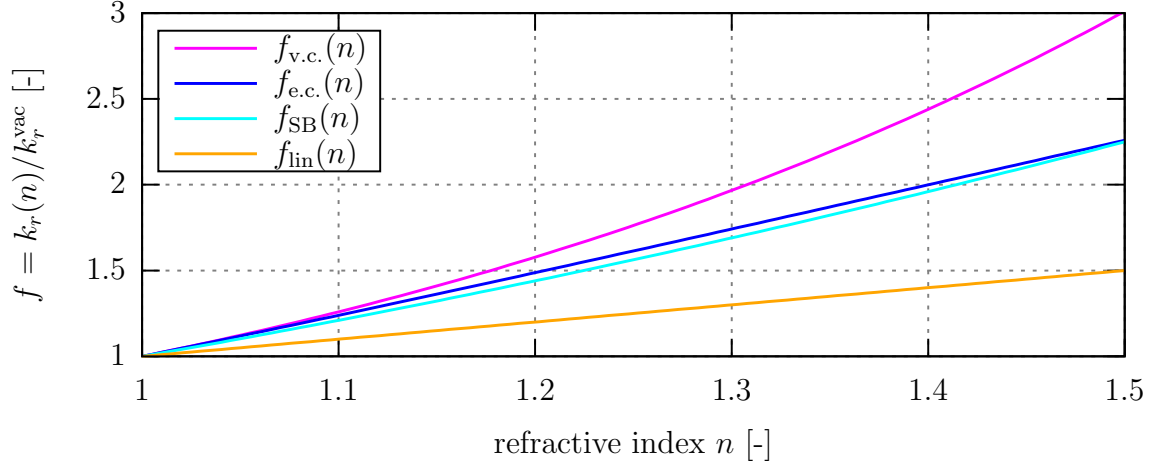
**Figure 4.24:** Comparison of the ratio $f(n)$ between the radiative rate in a medium with refractive index $n$ and in vacuum for the four different models described in equations (4.17) to (4.21), namely the linear, empty cavity, virtual cavity and Strickler-Berg model. When quantum yield values have been measured in one medium but need to be known in another, usually both media are liquids or polymers, limiting the relevant refractive index range to approximately $1.3 \leq n \leq 1.5$. Figure previously unpublished.
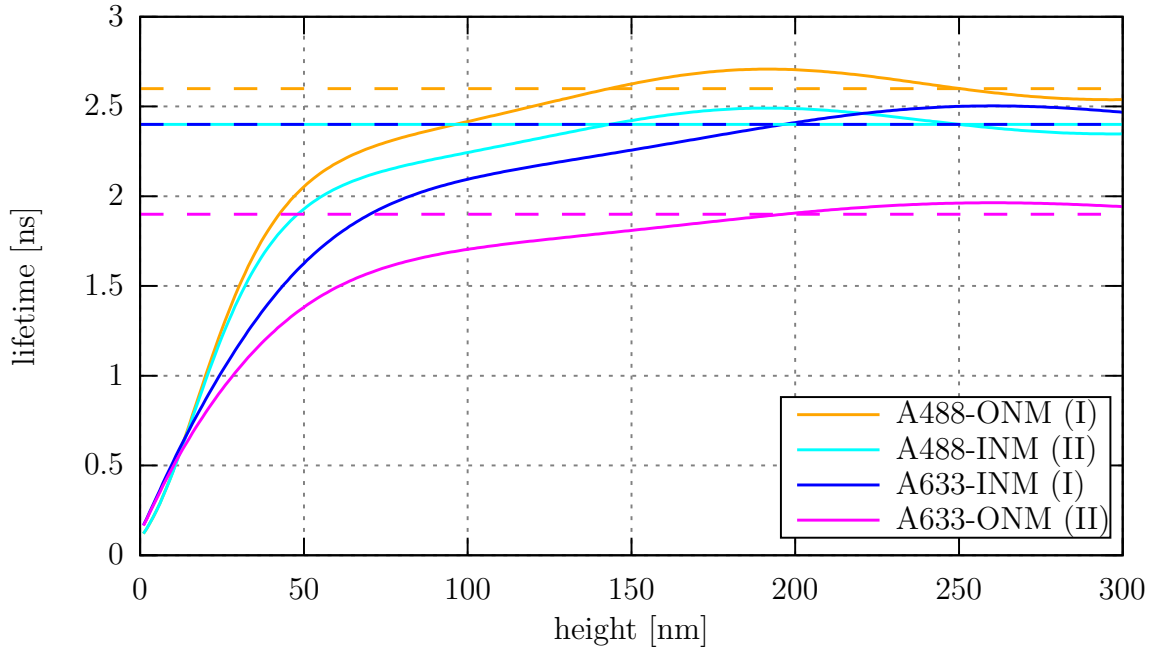


**Figure 4.25:** MIET lifetime-versus-height calibration curves for the four different label types: Alexa Fluor 488 labeling Nup358 (A488-ONM, sample type I) or Lap2$\beta$ (A488-INM, sample type II), and Alexa Fluor 633 labeling the same proteins in the inverse order (A633-INM, sample type I, and A633-ONM, sample type II). Dashed lines indicate the free space lifetimes $\tau_{0,\,\text{cell}}$ of the constructs inside the cells. Note that Lap2$\beta$ was almost exclusively found above 100 nm with significant populations above 150 nm, while Nup358 was mostly located above 50 nm. Thus, the height of especially the inner but also of parts of the outer nuclear membrane is in a range where the MIET curve is very flat, resulting in large height uncertainties for given lifetime uncertainties. Figure adapted from [3].

From these findings, we concluded that either the empty-cavity model or the Strickler-Berg equation can be used to estimate the radiative deexcitation rate inside the cell, and thus the quantum yield inside the cell. We chose the former and arrived at:

$$\Phi_{\text{cell}} = \Phi_{\text{water}} \cdot \frac{\tau_{0,\,\text{cell}}}{\tau_{0,\,\text{water}}} \cdot \frac{n_{\text{cell}}^5}{n_{\text{water}}^5} \cdot \frac{(2n_{\text{water}}^2 + 1)^2}{(2n_{\text{cell}}^2 + 1)^2}, \tag{4.22}$$

with $n_{\text{water}} = 1.33$ and $n_{\text{cell}} = 1.37$ as described above. The results for the free space lifetimes measured inside the cell and the derived quantum yields are summarized for the four different dye-antibody combinations in the following table:

| sample | | $\tau_{0,\,\text{cell}}$ [ns] | $\Phi_{\text{cell}}$ |
|---|---|---|---|
| A488-Nup358 | (sample type I) | 2.6 | 0.59 |
| A488-Lap2$\beta$ | (sample type II) | 2.4 | 0.54 |
| A633-Nup358 | (sample type II) | 1.9 | 0.37 |
| A633-Lap2$\beta$ | (sample type I) | 2.4 | 0.47 |

These values were then employed in the calculation of the four different MIET calibration curves shown in figure 4.25, which were used to transform lifetimes into height information.

## Summary

We have performed dual-colour MIET measurements to map the three-dimensional distance between the inner or outer nuclear membrane along the whole basal part of the nucleus. This was the first time that the nuclear envelope structure has been studied along the vertical axis using optical microscopy. With a height of approximately 100 nm to 150 nm above the substrate, the nuclear envelope is situated at the limit of the operating range of MIET, resulting in a strongly decreased localization accuracy. Even though the height uncertainties of some individual pixels were as large as the distance between INM and ONM reported previously (30 nm to 50 nm, [118]), the value that we found for the average distance between INM and ONM (30 nm to 35 nm) agrees well with previous works, both electron microscopy images of isolated nuclear membranes and horizontal measurements of single-point FRAP of the nuclear membrane in the equatorial plane.

We have studied the impact of four different sample models on the obtained distance between INM and ONM, and found no significant difference despite the fact that the nuclear envelope is situated in a highly complex environment. However, motivated by this work, this important question for estimating the accuracy of MIET will be studied more generally in the following section. Finally, we have compared different theories for relating the radiative deexcitation rate to the refractive index of the surrounding medium, which is crucial when estimating the quantum yield of fluorophores inside a cell. We found that the two models which are most endorsed in the literature – the empty-cavity model and the Strickler-Berg equation – yield very similar results in the relevant range of refractive indices, and were thus able to resolve this apparent contradiction.

## 4.2 Parametes influencing MIET z-localization accuracy

The MIET studies presented in the previous sections were concerned with structures situated in increasingly more complex environments: From microtubules on a coated glass cover slip surrounded by a buffer, via the outer membrane of a cell, to focal adhesions (which are close to the outer membrane of the cell) and finally the challenging regime of the nuclear envelope. As we stated in the introduction of chapter two, it is not feasible to model every detail of a real system. The main simplification which is used in all MIET calculations is the assumption of a stratified system. This means that the sample (excluding the dipole emitter itself) is modeled as a system of parallel planar layers with different thicknesses and refractive indices. As the studied structures get more complex, they deviate more from this model. Therefore, an important question when using MIET to study biological samples is which influence a deviation of the real structure from this ideal model has on the calculated height values. The second important parameter set that affects the MIET calibration curve consists of the optical properties of the fluorophore, namely the free space lifetime $\tau_0$ and the quantum yield $\Phi$. The former can be readily obtained by measuring the same sample on an untreated cover slip[12], therefore we assume that it is always correct. The latter is often only known in a different environment and with some measurement uncertainty, and has to be converted using the measured value of $\tau_0$ and the assumed value of the refractive index. Thus, it may also suffer from inaccuracies and introduce a bias in the calculated height values.

In this section, we study the effect of both the refractive index (RI) structure and the quantum yield values by calculating MIET curves for different scenarios and comparing the thus obtained height values $z_{\text{calc}}$ with the true heights $z_{\text{true}}$. Depending on the experiment, we are sometimes interested in the absolute height of a fluorophore above the substrate, and sometimes only in the $z$-position of two different types of fluorophores relative to each other. Therefore, both absolute height values $z$ and height differences $\Delta z = z_{\text{upper}} - z_{\text{lower}}$ will be examined.

The top row of figure 4.26 illustrates the situation studied in the first part of this section. A MIET substrate consisting of a glass cover slip ($n_{\text{glass}} = 1.52$) coated with thin metallic films – namely $2\,\text{nm}$ titanium, $10\,\text{nm}$ gold and $2\,\text{nm}$ titanium – and a $10\,\text{nm}$ thick silica spacer ($n_{\text{silica}} = 1.52$) is covered completely by a homogeneous halfspace with refractive index $n_{\text{cell}} = 1.37$. The composition of the MIET substrate is the same as that used for the examination of the nuclear envelope in the previous section, while 1.37 is the average RI of a HeLa cell. We now simulate a MIET curve for a hypothetical

---

[12]If the fluorophores are very close to the glass cover slip, e.g. when working with cells whose plasma membrane has been labeled, one has to take into account that a water/glass interface also induces a lifetime change, albeit a much smaller one of typically less than $10\,\%$. See figure 2.18a for a plot of $S_{\text{tot}}$ versus $z$ at a water/glass interface. In this case, one has to roughly estimate the height of the fluorophores, and use a curve such as that in figure 2.18a to determine a guess of the free space lifetime. This value may then be used to calculate the actual MIET calibration curve, and the refined height estimate can then be used to iterate the procedure. Due to the much weaker $z$-dependence of the lifetime at a water/glass rather than a water/metal interface, this calculation usually converges quickly.
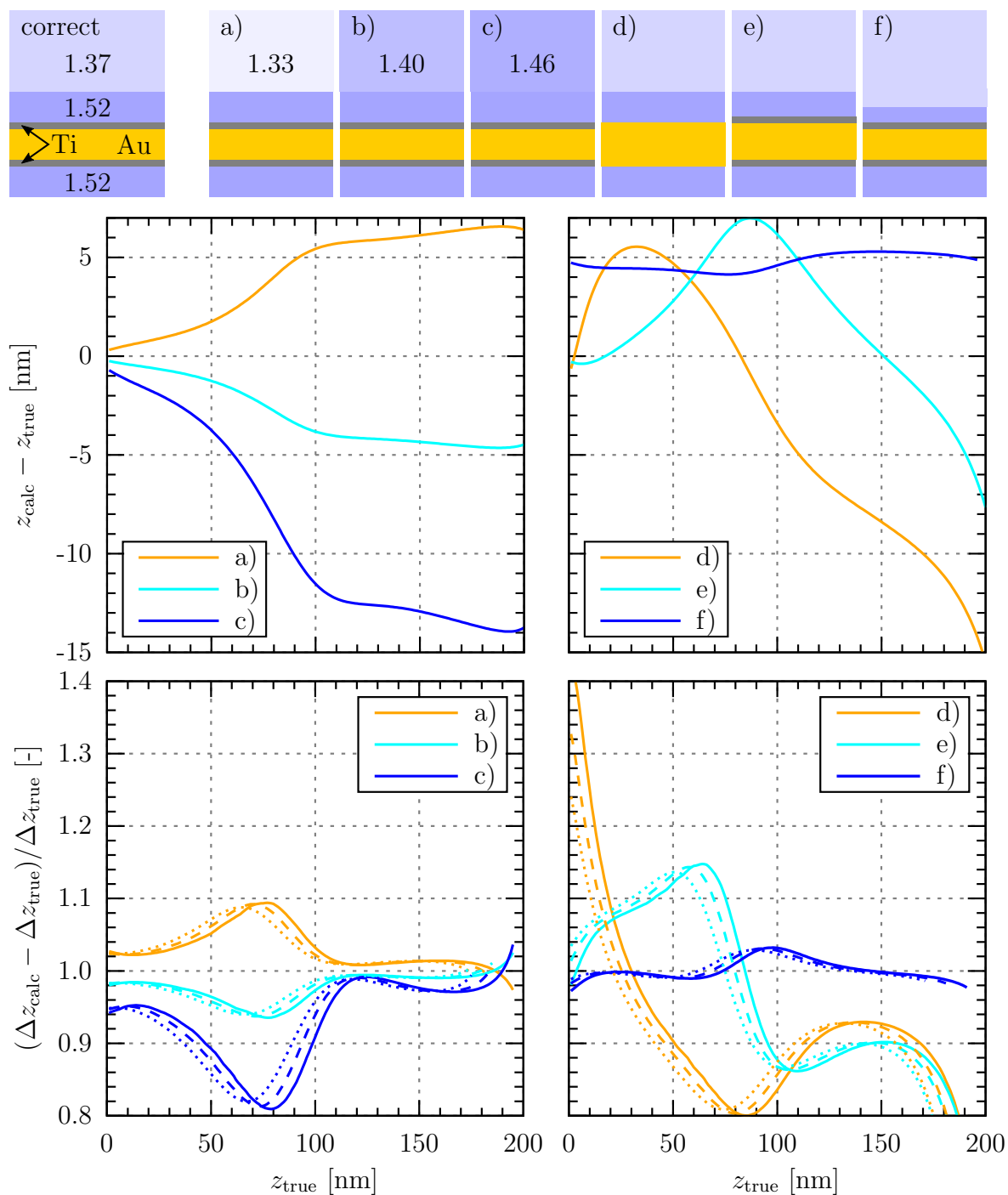
**Figure 4.26:** Impact of incorrectly modeled sample parameters on $z$-localization accuracy. **Top row:** Schematic of the true sample (left) and six "wrong" models (a-f) used to calculate MIET curves. Models (a-c) correspond to incorrect assumptions about the sample itself, while (d-f) concern wrong modeling of the MIET substrate, see text for more details. **Central row:** Deviation between the height values $z_{calc}$ calculated using the MIET curves (a-f) and the true height values $z_{true}$ for various emitter positions. **Bottom row:** Relative error of height differences $\Delta z$ between an upper and a lower fluorophore, plotted against the $z$-coordinate of the lower fluorophore. Calculations were done for true distances $\Delta z_{true} = 5\,\text{nm}$ (solid lines), $15\,\text{nm}$ (dashed lines) and $25\,\text{nm}$ (dotted lines). In all four panels, the emission wavelength is $\lambda = 600\,\text{nm}$ and the quantum yield $\Phi = 1$.

fluorophore emitting at $\lambda = 600\,\text{nm}$ and having a quantum yield of $\Phi = 1$ that is situated at many different heights $z_{\text{true}}$ above the silica layer[13]. The thus obtained lifetimes are then transformed back into height values $z_{\text{calc}}$ using MIET curves calculated for the following "wrong" sample models (see top row of figure 4.26):

a) Homogeneous halfspace filled with water ($n_{\text{water}} = 1.33$).

b) Homogeneous halfspace filled with mounting medium Fluoroshield ($n_{\text{FS}} = 1.40$).

c) Homogeneous halfspace filled with mounting medium Mowiol, using the average refractive index found in the literature ($n_{\text{Mow}} = 1.46$).

d) Replacing the three-layered metal system Ti-Au-Ti by a single gold layer whose thickness (14 nm) equals the sum of the three metal layers (2 nm-10 nm-2 nm).

e) Assuming the gold layer has a thickness of 12 nm instead of 10 nm.

f) Assuming the silica layer is only 5 nm thick.

Thus, a)-c) model the situation that the refractive index of the sample itself was guessed incorrectly, while d)-f) treat errors in the composition of the MIET substrate, namely an oversimplification of the sample by ignoring the presence of the titanium, or wrong thickness assumptions as they might occur due to errors during the sample preparation. The central row of figure 4.26 depicts the error $z_{\text{calc}} - z_{\text{true}}$ for all six situations. An incorrect refractive index in the sample halfspace leads to a bias of the height, where a too small RI leads to larger $z_{\text{calc}}$ (a) and a too large RI results in smaller $z_{\text{calc}}$ (b, c). We only study the height range where the MIET curves are monotonically increasing, i.e. up to $z_{\text{true}} = 200\,\text{nm}$. Within this range, the absolute error rises with $z$, however, it levels off at approximately $z = 100\,\text{nm}$, leading to a decreasing relative error. In contrast, the error due to incorrectly modeled metal layers can lead to both over- and underestimation of the height. Finally, assuming that the silica layer is only 5 nm instead of 10 nm thick results in a MIET curve whose shape is almost unchanged but that is shifted by 5 nm. Thus, when adding the silica spacer thickness and $z_{\text{calc}}$ to obtain the distance of the emitter from the top of the metal layers, the result is almost correct. This indicates that the refractive indices of dielectric materials which are not in direct contact with the emitter do not have a large impact on the MIET curve. The same calculations were repeated for a quantum yield of 30 % but are not shown here because they only differ in the fourth significant digit from the results for $\Phi = 1$.

In co-localization experiments such as the study of focal adhesions in section 4.1.3 or the distance determination of INM and ONM in the nuclear envelope in section 4.1.4, the main quantity of interest is the height difference $\Delta z$ between two fluorophores. Neglecting that these are usually dual-colour measurements, we present the relative error of $\Delta z_{\text{calc}}$ for $\Delta z_{\text{true}} = 5\,\text{nm}$, 15 nm and 25 nm in the bottom row of figure 4.26. For all three fluorophore distances, the relative error of a model with incorrect sample RI (a-c) shows a similar behaviour that is consistent with the plots in the central row: At small ($z_{\text{true}} < 30\,\text{nm}$) and large ($z_{\text{true}} > 120\,\text{nm}$) distances from the substrate, the height

---

[13]Since the MIET curve is proportional to the free space lifetime and the latter is assumed to be known, its value is arbitrary for these calculations.

bias is almost constant, leading to a small relative error of $\Delta z$ in the order of $5\,\%$ or less. At intermediate distances, where the height bias $z_{\mathrm{calc}} - z_{\mathrm{true}}$ changes rapidly, the relative error of $\Delta z$ has a maximum around $6\,\%$ to $20\,\%$, depending on the difference between the correct and the actually used refractive index. If the sample is modeled correctly but the metal layers are represented wrongly (d-e), very large relative deviations up to $40\,\%$ can occur. Contrarily, assuming a smaller thickness of the silica spacer (f) results in relative errors that are always smaller than $4\,\%$, consistent with the almost parallel curve in the central row.

These results explain why we only saw a very small impact of the four different sample geometries on the calculated average distance between the two nuclear membranes in section 4.1.4. Including a $10\,\mathrm{nm}$ thick layer of Mowiol with a higher RI than the cell in the model is similar to changing the thickness of the silica layer, thus only small deviations of $\Delta z_{\mathrm{calc}}$ between models are expected. Similarly, the inclusion of the plasma membrane (i.e. a thin dielectric layer) in the model does not result in large changes. In the study of focal adhesions in section 4.1.3, vinculin was found at heights around $10\,\mathrm{nm}$ to $20\,\mathrm{nm}$, while actin was located at approximately $20\,\mathrm{nm}$ to $50\,\mathrm{nm}$. At these small distances from the substrate, even when comparing the smallest and largest refractive indices used in the simulation, namely $1.33$ and $1.46$, the calculated heights differ by less than $6\,\mathrm{nm}$ (difference between a) and c) in the top row of figure 4.26), or the relative height difference between actin and vinculin by less than $6\,\%$ (bottom row). Since the true refractive index inside the cells is probably in between the RI of cytosol ($n = 1.37$) and the RI of the mounting medium ($n = 1.4$), using the latter as sample RI is a valid approximation that should not introduce significant errors.

With that, we turn to the influence of incorrect quantum yield values on the determination of height values. The sample we study is the same as before, but now the fluorophore either has a quantum yield of $\Phi_{\mathrm{true}} = 1.0$ or $\Phi_{\mathrm{true}} = 0.3$. Figure 4.27 shows the same kind of height deviation plots as figure 4.26, now for varying assumed quantum yields. The curves in the top row demonstrate that there is one point at $z_{\mathrm{true}} = 161\,\mathrm{nm}$ where all MIET curves predict the same lifetime, independent of the quantum yield. At this height, the total energy $S_{\mathrm{tot}}(z)$ emitted by an ideal dipole emitter with $\lambda = 600\,\mathrm{nm}$ equals the free-space value $S_0$, thus, $\tau(z) = \tau_0/(1 - \Phi + \Phi \cdot S_{\mathrm{tot}}(z)/S_0) = \tau_0$ irrespective of $\Phi$. For $z_{\mathrm{true}} < 161\,\mathrm{nm}$, an underestimation of the quantum yield leads to an underestimation of the height, and vice versa. For $z_{\mathrm{true}} > 161\,\mathrm{nm}$, the opposite is true. The effect is stronger for small quantum yield values than for large $\Phi_{\mathrm{true}}$. For example, the height error is always less than $1\,\mathrm{nm}$ if one assumes a quantum yield of $98\,\%$ instead of $100\,\%$. The plots in the lower row of figure 4.27 illustrate the relative error of the determined relative distance $\Delta z_{\mathrm{calc}}$ for $\Delta t_{\mathrm{true}} = 5\,\mathrm{nm}$, $15\,\mathrm{nm}$ or $25\,\mathrm{nm}$. Here, a complex behaviour can be observed, but the general trend is the same for all seven cases: For fluorophores that are close to the substrate ($z_{\mathrm{true,\ lower}} \leq 30\,\mathrm{nm}$), the relative error is quite small, while it increases drastically at heights above $150\,\mathrm{nm}$. As before, the effect is smaller for fluorophores with a higher quantum yield.

Overall, the height deviation observed when employing a wrong quantum yield is larger than the height bias introduced by using a wrong refractive index. This highlights the importance of knowing the correct quantum yield of a fluorophore. However, this quantity is often not published by the manufacturers of fluorescent substances, because it is harder to determine than, say, a fluorescent lifetime or an emission spectrum.
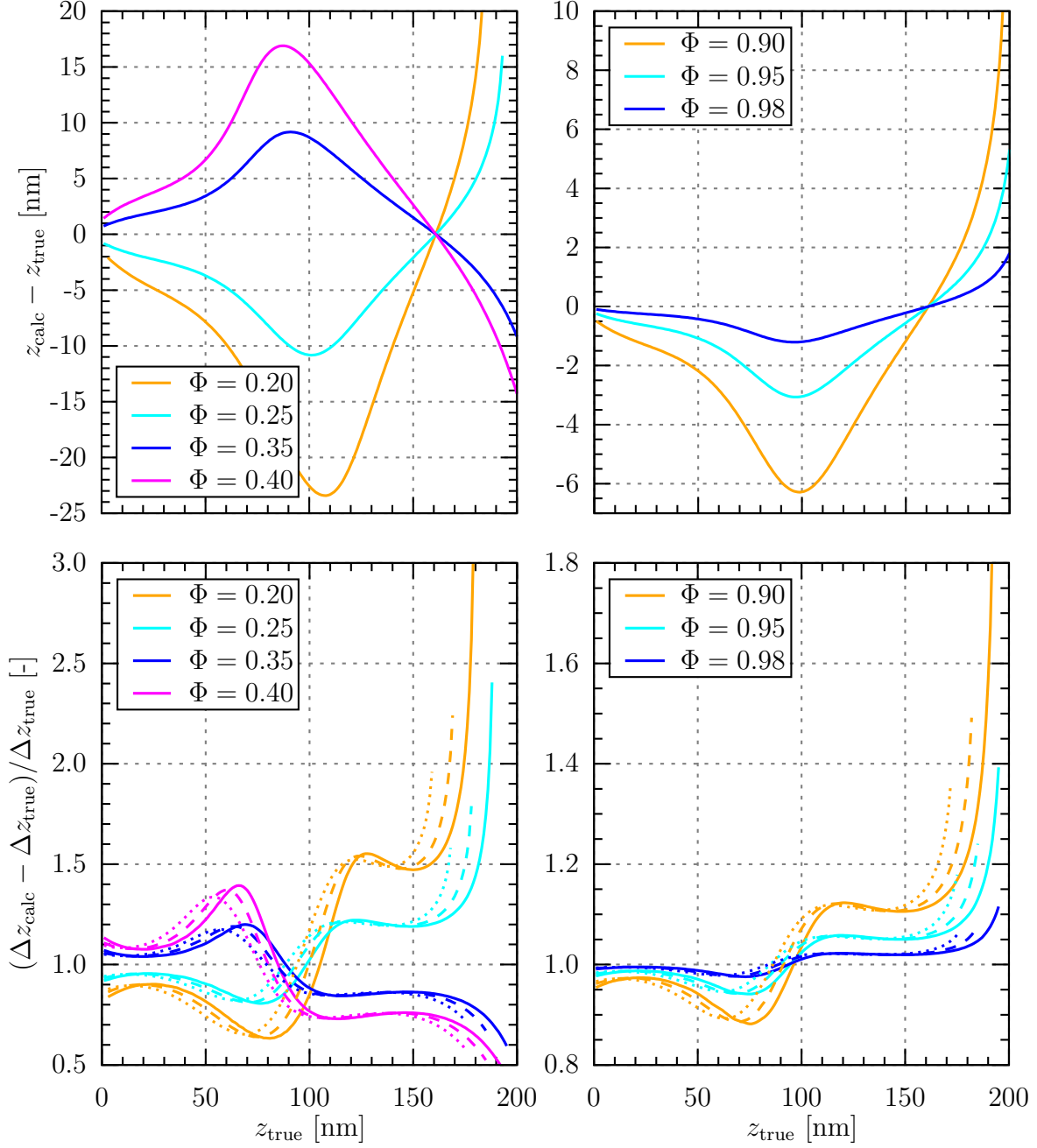
**Figure 4.27:** Impact of incorrect quantum yield values on $z$-localization accuracy. In the left column, the "true" quantum yield is $\Phi_{\text{true}} = 0.3$, in the right column $\Phi_{\text{true}} = 1.0$. **Top row:** Deviation between the height values $z_{\text{calc}}$ calculated using MIET curves with "wrong" quantum yield values $\Phi$ as indicated and the true height values $z_{\text{true}}$ for various emitter positions. **Bottom row:** Relative error of height differences $\Delta z$ between an upper and a lower fluorophore, plotted against the $z$-coordinate of the lower fluorophore. Calculations were done for true distances $\Delta z_{\text{true}} = 5\,\text{nm}$ (solid lines), $15\,\text{nm}$ (dashed lines) and $25\,\text{nm}$ (dotted lines). In all four panels, the emission wavelength is $\lambda = 600\,\text{nm}$.

## Applications

Therefore, the next section introduces a method for determining quantum yield values using a metal nanocavity.

# 4.3 Nanocavity-based quantum yield measurements

As mentioned previously, the fluorescence quantum yield (QY) $\Phi$ is defined as the average number of fluorescence photons emitted per absorbed excitation photons. In fluorescence microscopy, the QY used to be an interesting quantity mainly because, together with the absorption cross-section (i.e. the probability of the fluorophore to absorb excitation light), it determines the average number of fluorescence photons that can be detected at a given excitation power. Thus, fluorophores with a larger QY generally yield a higher signal, and consequently better images. More recently, the QY has also attracted attention because it influences effects such as metal-induced energy transfer, as we saw in detail in the previous sections. Therefore, methods to accurately determine the fluorescence quantum yield are of high importance. Before describing a relatively new technique that utilizes the fluorescence lifetime changes close to metal surfaces, which was refined and applied to various systems over the course of this work, we briefly describe alternative methods that have been used in the past. This summary is mostly based on a review article by Demas and Crosby [135].

Historically, the first reliable technique for obtaining absolute QY values was described by Wawilow in 1924 [136]. It was used with slight modifications until the 1960s [135]. The basic idea behind Wawilow's method was based directly on the definition of the quantum yield,

$$\Phi = \frac{\text{number of fluorescence photons}}{\text{number of absorbed excitation photons}}. \tag{4.23}$$

In a cuvette, the fluorophore to be studied was dissolved in a solvent at a sufficiently high concentration that an excitation light beam was almost completely absorbed. Normal to the front surface of the cuvette, a detector then measured the fluorescene signal. In a second step, the fluorescent sample was exchanged with a scattering sample that allowed to determine the intensity of the excitation light at the position of the cuvette. Taking into account some correction factors, the ratio of these two measurements yielded the fluorescence quantum yield. However, the accuracy of this method depended on several parameters that were difficult to control, such as the quality of the scattering sample, the wavelength-depenence of some of the correction factors, or distortion of the measured fluorescence due to reabsorption of the light in the highly concentrated solution. Therefore, even under optimal conditions, the relative uncertainty of the obtained quantum yield values was around 5 %, and for most molecules, a relative error of 10 % was "more realistic" [135]. The development of more reliable scattering samples, as well as better instrumentation, later allowed more accurate measurements.

As soon as the quantum yield values of a number of fluorophores had been confirmed independently by several groups, these could be used as reference samples, making the (sometimes unpredictable [135]) scattering samples unneeded: If both the solution of the compound to be studied and the reference could be excited with light of the same wavelength, and if both solutions had the same optical density, the ratio of the measured intensities equaled the ratio of the quantum yields (with some corrections, see e.g. [135]). However, for these measurements, too, there were some experimental difficulties. To name just one, if the emission spectra of the fluorophore of interest and the reference sample were not identical, wavelenght-dependent detection efficiencies

required an exact calibration of the detector. Furthermore, this type of measurement could only be conducted if a reference sample with a similar absorption spectrum existed. And finally, any error in the quantum yield value of the reference resulted in an error in the quantum yield of the studied fluorophore. Nevertheless, relative quantum yield measurements are probably the most common technique used today.

A complementary approach to measure fluorescence quantum yields is calorimetry, a technique that measures temperature changes during irradiation. When a nonluminescent material is irradiated, the absorbed light is transformed entirely to heat. Contrarily, when a luminescent sample is irradiated, only a fraction of the absorbed energy – namely $(1 - \Phi)$ of the absorbed photons – is transformed to heat. Thus, when optically exciting solutions of a nonluminscent reference and the sample to be studied, both with the same optical density, and comparing the temperature change, one can deduct the fraction of non-radiative deexciation events, and thus $\Phi$. This type of measuremet requires very precise temperature sensors, a good thermal isolation of the sample chamber, and a careful design to prevent systematic errors e.g. from back-reflected fluorescence that heats the sample. This method is still used today, but it is not so widespread due to the typically custom-designed and more expensive equipment ([137] and citations therein). Some of the experimental difficulties of the methods discussed so far can be overcome by using an integrating sphere, a large globe with a reflective coating on the inside. The excitation light enters through a window and reaches the sample in the center of the sphere. Subsequently, the fluorescence light is reflected multiple times at the inside of the sphere before exiting through a second window, where the detector is placed. Due to the multiple reflections, polarization effects are eliminated, and a non-uniform angular distribution of radiation of the sample (e.g. in crystals) is equalized. Nowadays, integrating spheres are mostly used for absolute quantum yield measurements [137–139], while relative methods mostly employ absorption and fluorescence spectromenters [137]. Although this introduction presented a number of different techniques for determining QY values, a new class of fluorophores has exposed a common weakness of all these methods: Photoactivatable proteins have two distinct states, one where they are not luminescent (dark state), and one where they fluoresce (bright state) [140, 141]. By illumination with light of a certain wavelength, the proteins can be stochastically switched from the bright to the dark state. If the proteins absorb light in their dark state, then all techniques introduced so far would yield an effective quantum yield value that depends on the ratio of bright and dark molecules in the solution. However, the true quantum yield of the proteins in the bright state could only be obtained on a sample where all molecules have switched from the bright to the dark state, which one generally cannot guarantee. Therefore, a technique is needed that is only sensitive to fluorescent molecules.

Such a method was developed by Enderlein, Chizhik *et al.* in [107]. It is based on fluorescence lifetime measurements in metal nanocavites, and the main physical mechanism exploited to determine the QY is MIET. The theoretical refinement of this method, the programming of a software package to improve the accessibility for new users, and the application to special samples was a large part of this work. Therefore, the theoretical basics are explained in detail in the next subsection, followed by a short presentation of the software implementation and finally the application of the technique to dye molecules in lipid membranes.

## 4.3.1 Theory

As we saw previously, metal-induced energy transfer imaging is based on the observation that the fluorescence lifetime $\tau$ of a dipole emitter decreases close to a metal halfspace or even just a thin metal layer. The central equations describing this effect are

$$\tau(z,\theta) = \frac{1}{k_{\mathrm{r}}(z,\theta) + k_{\mathrm{nr}}},$$

$$\text{and} \quad k_{\mathrm{r}}(z,\theta) = \frac{S_{\mathrm{tot}}(z,\theta)}{S_0} \, k_{\mathrm{r},0}. \tag{4.24}$$

Here, $z$ is the distance between the fluorophore and the metal interface (i.e. the "height" of the emitter) and $\theta$ is the angle between the dipole moment and the optical ($z$-) axis, i.e. $\theta = 0$ for a vertical and $\theta = \pi/2$ for a parallel dipole. Furthermore, $k_{\mathrm{r}}(z,\theta)$ is the radiative deexcitation rate of a dipole emitter at $z$ with orientation $\theta$, $k_{\mathrm{nr}}$ denotes the position- and orientation-independent non-radiative deexcitation rate, $S_{\mathrm{tot}}(z,\theta)$ is the electromagnetic energy emission per time for a fluorophore at $z$ with orientation $\theta$ (derived in sections 2.4.5 and 3.4), and $S_0$ is the energy emission per time for the same fluorophore in the same medium but in the absence of the metal. The quantum yield $\Phi$ of the fluorophore in the absence of the metal relates the directly measurable free space lifetime $\tau_0$ (i.e. the lifetime of the fluorophore in the absence of the metal) to the not directly accessible rates $k_{\mathrm{r},0}$ and $k_{\mathrm{nr}}$:

$$\Phi := \frac{k_{\mathrm{r},0}}{k_{\mathrm{r},0} + k_{\mathrm{nr}}} = k_{\mathrm{r},0} \cdot \tau_0 = 1 - k_{\mathrm{nr}} \cdot \tau_0. \tag{4.25}$$

This leads to the equation we use to calculate MIET calibration curves:

$$\tau(z,\theta) = \frac{\tau_0}{\Phi \cdot [S_{\mathrm{tot}}(z,\theta)/S_0] + 1 - \Phi}. \tag{4.26}$$

Theoretically, by preparing a sample with known metal-fluorophore distance $z$ and known orientation[14] $\theta$, and by subsequently measuring the lifetime $\tau(z,\theta)$, one can deduce $\Phi$. In practice, the accuracy of such a measurement can be improved (and possible errors in the assumptions of $\tau_0$, $S_0$ and the optical properties of the sample entering $S_{\mathrm{tot}}(z,\theta)$ can be discovered) by measuring lifetimes for a number of $z$-values and fitting equation (4.26) with $\Phi$ as the free parameter.

While this is a valid approach, it has some drawbacks. Firstly, the most straightforward way to achieve different $z$-values is by fabricating samples with spacer layers of different thickness, which is tedious. Secondly, in order to minimize the uncertainty of the $z$-position, fluorophores either have to be deposited directly on the spacer or be embedded in a very thin layer of a transparent material[15]. However, this might change the non-radiative transition rate. Moreover, as we argued in section 4.1.4, it is not trivial to

---

[14]Which usually means averaging over a random distribution of $\theta$-values.

[15]The former can be achieved, for example, by spin-coating a droplet of fluorophore-containing solution on the sample which evaporates quickly, leaving behind a thin layer of fluorophores. The second can be produced in a similar manner, but using e.g. a polymer or optical glue instead of a quickly evaporating solvent. However, the thickness of a layer prepared in this manner is difficult to control.
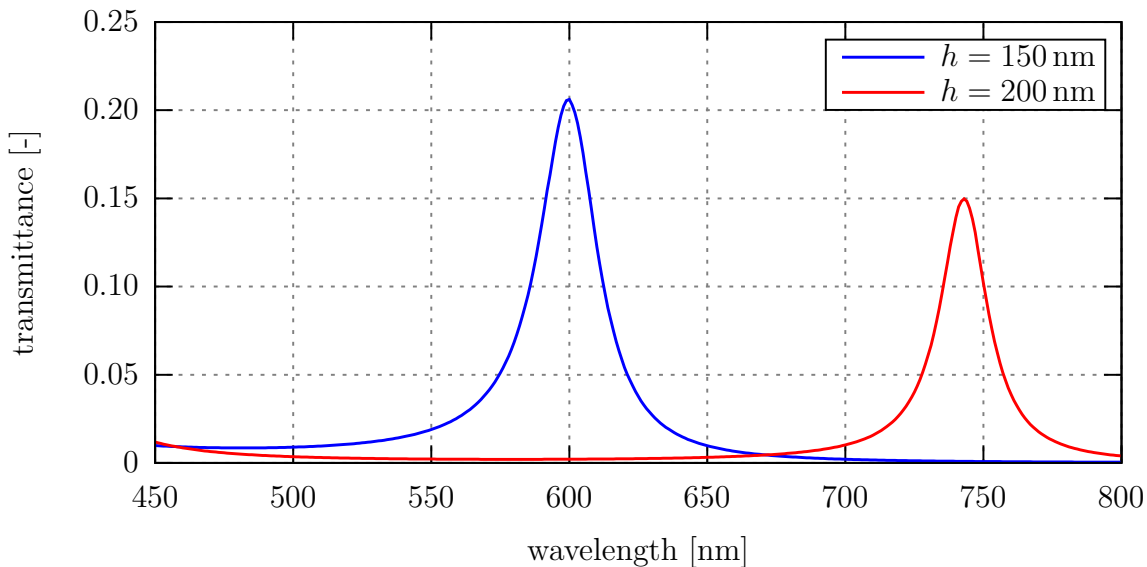
**Figure 4.28:** White-light transmittance spectra $T(\lambda) = |t_{p,s}(\lambda)|^2$ for two nanocavities, one with height $h = 150\,\text{nm}$ (blue line) and one with height $h = 200\,\text{nm}$ (red line). In both cases, the bottom silver mirror has a thickness of $30\,\text{nm}$ and the upper a thickness of $60\,\text{nm}$, the cavity itself is filled with DMSO (refractive index 1.48). In our setup, the white light lamp is placed directly above the sample, which corresponds to normal incidence of the light. Therefore, there is no distinction between $p$- and $s$-waves. Note that the transmission maximum shifts towards longer wavelengths for an increased cavity height. The magnitude of the peak decreases because the reflectance of the silver mirrors increases with $\lambda$.

convert quantum yield values for materials with differing refractive indices. Thus, it is desirable to measure the quantum yield in the same medium in which one intends to use the fluorophore later – which generally means a liquid of some kind. In that case, fluorophores are present at many different $z$-positions simultaneously, and the measured lifetime is the average lifetime of all fluorophores within the focal volume. Since the latter has a $z$-extension of approximately $500\,\text{nm}$ to $700\,\text{nm}$ (see equation 4.15 in section 4.1.4) and since MIET causes lifetime changes only in the first $250\,\text{nm}$ to $300\,\text{nm}$ (see e.g. figure 2.29), the change of the measured average lifetime with the quantum yield is small, which leads to large uncertainties of the obtained $\Phi$ values. Here, Chizhik *et al.* found a simple, yet powerful solution with several practical advantages [107]: Instead of using just one metal-coated surface, they placed two parallel metal-coated glass surfaces – denoted (semi-transparent) "mirrors" in the following – approximately $h = 100\,\text{nm}$ to $h = 200\,\text{nm}$ apart, with the fluorescent solution in between. This geometry is called a nanocavity, a microresonator, or a Fabry-Pérot interferometer. It is mainly known for the fact that its transmission spectrum (as a function of wavelength $\lambda$) displays distinct peaks corresponding to constructive interference of the multiply reflected waves. Figure 4.28 shows exemplary transmittance spectra for two dimethyl sulfoxide (DMSO)-filled cavities with silver mirrors of thickness $30\,\text{nm}$ and $60\,\text{nm}$ and with cavity heights $h = 150\,\text{nm}$ and $200\,\text{nm}$. The transmittance $T$ is the square of the absolute value of the effective Fresnel transmission coefficient $t(\theta, \lambda)$ of
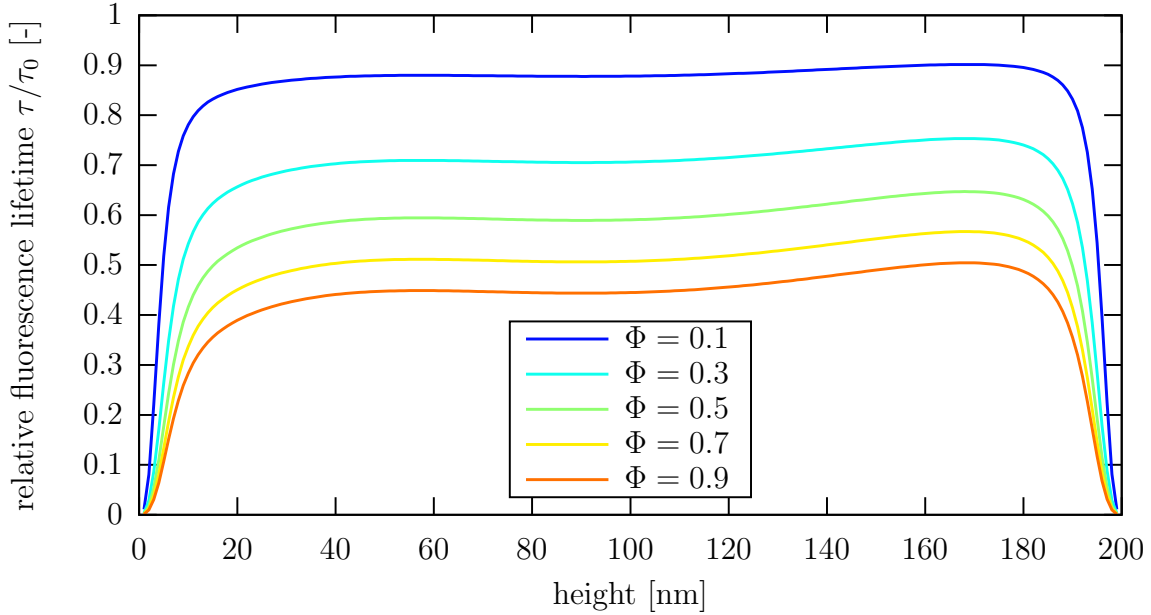
**Figure 4.29:** Height-dependent fluorescence lifetime of quickly rotating fluorophores (emission wavelength 650 nm) with varying quantum yield $\Phi$ inside a DMSO-filled nanocavity. Thicknesses of the silver mirrors are 30 nm (bottom mirror) and 60 nm (top mirror), which is why the curves are not symmetric.

the whole sample, i.e. the stratified system glass-silver-DMSO-silver-glass, calculated as described in section 3.2. It is assumed that the light source is positioned above the sample, i.e. $\theta = 0$, thus there is no distinction between $p$- and $s$-waves. Since the mirrors are relatively thin, their reflectance is low (monotonically increasing with $\lambda$ in the visible range, 0.6-0.9 for the 30 nm thick mirror and 0.90-0.97 for the 60 nm thick layer), which means that the resonance peaks are slightly smeared out. Nevertheless, by measuring these transmission spectra and fitting $|t(\lambda)|^2$ to the data with $h$ as the fit parameter, the cavity height can be determined with a high precision in the order of 1 nm. Furthermore, deviations of the metal layer thicknesses from the design values can be detected as a deviation between model and data which cannot be eliminated by changing $h$. This renders any further checks of the sample fabrication unnecessary.

The second advantage of a nanocavity is that all fluorophores are close to a metal layer, which increases the sensitivity of the lifetime measurement with respect to quantum yield changes. Figure 4.29 compares the lifetimes of dipole emitters at different heights inside a cavity with $h = 200$ nm for varying quantum yields of the fluorophores. It becomes apparent that a higher quantum yield leads to a significantly lower average fluorescence lifetime. The exact amount of this lifetime change varies for different $h$ (not shown in the figure)[16].

Thirdly, the need for spacer layers of varying thickness to generate independent data points is eliminated by a nanocavity: One can simply vary the distance $h$ between the

---

[16]If this were not the case, changing the MIET curve by using a different free space lifetime $\tau_0$ might result in the same expected average lifetime. Then, one would have to determine the exact value of $\tau_0$ separately in order to find the correct value of $\Phi$. However, because the lifetime change with $h$ is non-linear, we can define $\tau_0$ as a second fit parameter, as will be shown later in this section.

two mirrors and measure the average lifetime $\tau_{\mathrm{meas}}$ each time. Experimentally, this can be achieved elegantly by using a silver-coated convex lens as the second mirror. The curvature of the lens surface is so small that it can be approximated as plane within the focal volume. By moving the sample laterally, one moves between regions with different distances between the planar bottom mirror and the lens, and thus between cavities with different height $h$. For each lifetime measurement, the current height of the cavity can be obtained by acquiring a white-light transmission spectrum and fitting it to $|t(\lambda)|^2$ with $h$ as fit parameter, as explained above.

Thus, the main challenge of using a nanocavity to determine quantum yield values is not of practical but of theoretical nature: For each cavity height, a single TCSPC curve is acquired, and has to be related to the quantum yield $\Phi$ via an accurate theoretical model. We do this in a two-step approach. Firstly, the average photon arrival time $\tau_{\mathrm{meas}}$ is extracted from the TCSPC histogram. Secondly, this value is modeled taking into account a number of different effects:

1. The lifetime of a single excited dipole emitter depends both on the $z$-position and the orientation $\theta$ of the emitter, as described by equation (4.26).

2. The probability to excite a dipole emitter with a focused laser beam depends on $z$ and $\theta$.

3. If a radiative decay occurs, the probability to actually detect the resulting photon also depends on $z$ and $\theta$.

4. Generally, the fluorophores are dissolved in a buffer, and thus free to rotate. Depending on the relative magnitudes of the rotational diffusion time and the fluorescence lifetime, the fluorophores can be treated as fixed, or the rotational diffusion has to be included in the model.

Based on [107], the last three aspects will now be discussed separetely, and finally combined to yield an expression of the measured average photon arrival time $\tau_{\mathrm{meas}}$ that depends on the cavity height $h$ and the quantum yield $\Phi$. By measuring $\tau_{\mathrm{meas}}$ for several values of $h$ and fitting the model to the data, an estimate of $\Phi$ is obtained.

**Excitation probability**

Assuming that the excitation intensity is small enough that non-linear saturation effects can be ignored, the excitation rate of a molecule with position $\boldsymbol{r}_m = (\rho_m, \varphi_m, z_m)$ and absorption dipole moment $\boldsymbol{p}$ is proportional to $|\boldsymbol{p} \cdot \boldsymbol{E}(\boldsymbol{r}_m)|^2$, where $\boldsymbol{E}$ is the exciting electric field that will now be determined. The nanocavity lifetime measurements are performed on a confocal FLIM setup as already described in this work. The calculation of the electric field of the excitation laser inside the nanocavity is performed on the basis of the concept developed by Wolf and Richards [142, 143]. The details are given in [107], we will only briefly outline the basic idea here and show the resulting formulas. The focusing of a linearly polarized laser beam by an objective with high numerical aperture can be described by decomposing the field in plane $p$- and $s$-waves, which are then treated individually. This is depicted schematically in figure 4.30: When the focusing is close to diffraction limited, one can assume that the field in the back
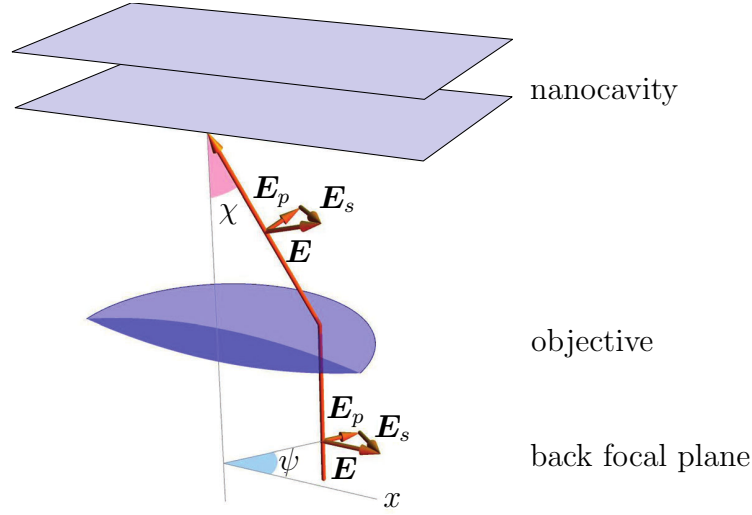
**Figure 4.30:** Focusing of a linearly polarized laser beam into a nanocavity. In the back focal plane of the objective, the electric field is approximated as a plane wave. At each point, the electric field vector $\boldsymbol{E}$ is decomposed into a radial and a tangential component, which will turn into $p$- and $s$-waves in the cavity, respectively. The further such a point is from the optical axis, the larger the angle $\chi$ that the wavevector makes with the optical axis after passing the lens. The maximum angle $\chi_{\max}$ is determined by the numerical aperture of the objective, $NA = n \cdot \sin \chi_{\max}$. Figure adapted from [107].

focal plane of the objective is a single plane wave, and thus has the same amplitude at each point. At each such point, the electric field vector is decomposed into a radial and a tangential component, which will become $p$- and $s$-waves, respectively, when the light from this point is focused into the nanocavity. The further the point is from the optical axis, the larger the angle $\chi$ between the beam and the optical axis after the beam passes through the lens system. Since the numerical aperture ($NA$) of the objective is related to the largest angle $\chi_{\max}$ under which light can leave the objective, $NA = n \sin \chi_{\max}$ where $n$ is the refractive index of the objective's mounting medium, it defines which plane waves have to be taken into account when calculation the field inside the nanocavity. Upon reaching the nanocavity, the plane waves are reflected and transmitted multiple times in the stratified system. This situation was already treated in section 3.4 (*Energy flux through infinite plane*), with the core idea to sum all upwards- or downwards-traveling waves independently. When denoting the angle between the optical axis and the propagation direction within the fluorophore's medium as $\chi_m$ and the refractive index of said medium as $n_m$, then Snell's law gives $n \sin \chi = n_m \sin \chi_m$, connecting the propagation directions outside and inside the cavity. Combining these steps leads to the final result for the electric field [107]:

$$
\begin{aligned}
E_x(\rho, \varphi, z) &= f_0(\rho, z) + f_2(\rho, z) \cdot \cos(2\varphi) \\
E_y(\rho, \varphi, z) &= f_2(\rho, z) \cdot \sin(2\varphi) \\
E_z(\rho, \varphi, z) &= f_1(\rho, z) \cdot \cos(\varphi),
\end{aligned}
\tag{4.27}
$$

where it was assumed that the electric field in the back focal plane is polarized along the $x$-axis, and the point $\boldsymbol{r} = (\rho, \varphi, z)$ in the nanocavity is given by cylindrical coordinates
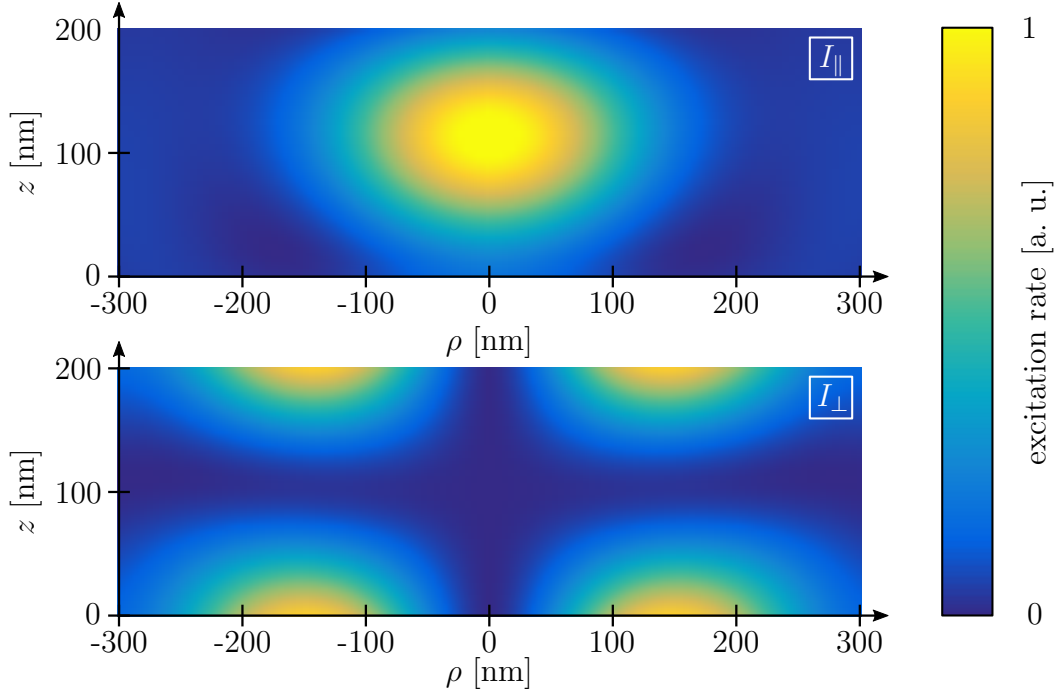
**Figure 4.31:** Excitation rate for a dipole oriented parallel ($I_{\parallel}^{\mathrm{exc}}$) or perpendicular ($I_{\perp}^{\mathrm{exc}}$) to the cavity mirrors, calculated for excitation light of wavelength $\lambda = 585\,\mathrm{nm}$ and for a cavity that is filled with DMSO ($n_m = 1.48$). The bottom mirror extends from $z = -30\,\mathrm{nm}$ to $z = 0\,\mathrm{nm}$, the top mirror from $z = 200\,\mathrm{nm}$ to $z = 260\,\mathrm{nm}$, both are made of silver.

to account for the symmetry of the situation. The functions $f_{0,1,2}(\rho, z)$ are independent of the angle $\varphi$, they are obtained by integrating over all plane wave components, taking into account all optical properties of the sample, including the cavity height. The exact form of these functions is given in section 6.1.3 in the appendix.

As mentioned above, the excitation rate of a molecule situated at $\boldsymbol{r}_m$ is proportional to $|\boldsymbol{p} \cdot \boldsymbol{E}(\boldsymbol{r}_m)|^2$. Since only the orientation of the absorption dipole moment $\boldsymbol{p}$ relative to the optical axis (given by the angle $\theta$), and not in the plane perpendicular to it, influences the lifetime, the dipole orientation can be averaged over the azimuthal angle around the optical axis. Then, the excitation rate $I^{\mathrm{exc}}$ is given by:

$$
\begin{aligned}
I^{\mathrm{exc}}(\theta, \boldsymbol{r}_m) &= I_{\perp}^{\mathrm{exc}}(\boldsymbol{r}_m) \cdot \cos^2\theta + I_{\parallel}^{\mathrm{exc}}(\boldsymbol{r}_m) \cdot \sin^2\theta \\
\text{with} \quad I_{\perp}^{\mathrm{exc}}(\boldsymbol{r}_m) &= \frac{1}{2}\left[\, |E_x(\boldsymbol{r}_m)|^2 + |E_y(\boldsymbol{r}_m)|^2 \,\right] \\
I_{\parallel}^{\mathrm{exc}}(\boldsymbol{r}_m) &= |E_z(\boldsymbol{r}_m)|^2.
\end{aligned}
\tag{4.28}
$$

As an example, figure 4.31 shows the excitation rate for dipoles that are oriented parallel ($\theta = \pi/2$) or perpendicular ($\theta = 0$) to the cavity mirrors and that are located in the $x$-$z$-plane in the cavity. Note that the dipoles parallel to the mirrors are excited much more efficiently than the ones perpendicular to the interfaces. Furthermore, the excitation maximum is close to the center of the cavity for the former, while it is near
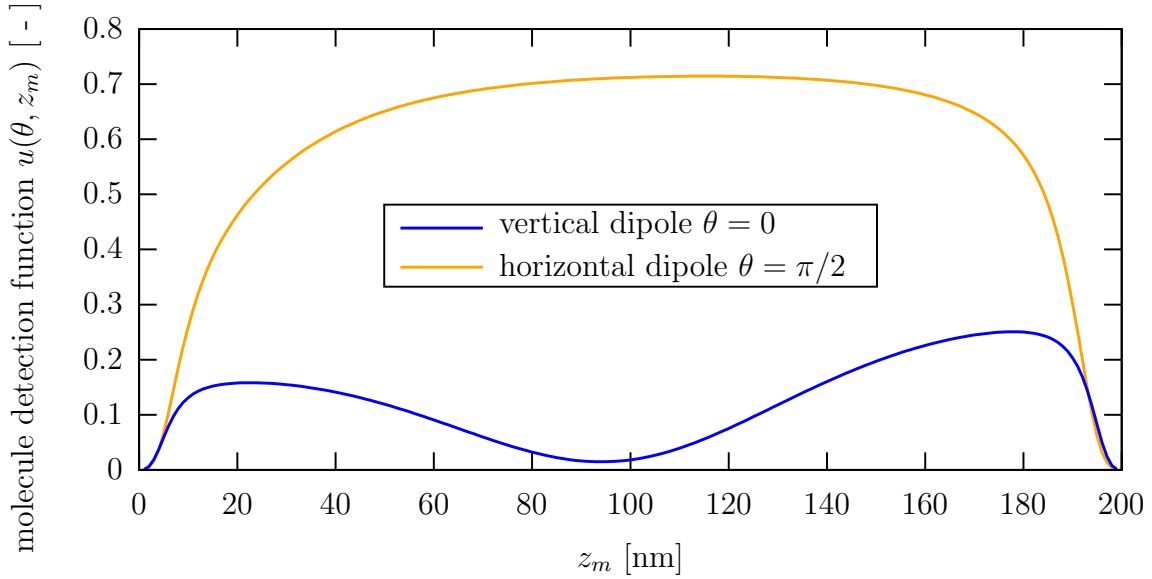
**Figure 4.32:** Molecule detection function $u(\theta, z_m)$ inside a cavity for dipoles oriented perpendicular ($\theta = 0$) or parallel ($\theta = \pi/2$) to the mirrors. The top silver mirror has a thickness of 60 nm, the bottom mirror is 30 nm thick, the cavity is filled with DMSO ($n_m = 1.48$). The height of the cavity is $h = 200$ nm, the numerical aperture of the objective is 1.49 (oil immersion with $n = 1.52$, i.e. $\chi_{\max} = 79°$) and the emission wavelength is $\lambda = 650$ nm.

the mirrors for the latter[17]. Since the fluorescence lifetime of a fluorophore only depends on $\theta$ and $z_m$, not $\rho_m$ or $\varphi_m$, $I^{\text{exc}}(\theta, \boldsymbol{r}_m)$ is integrated over $\rho_m$ and $\varphi_m$ to yield $I^{\text{exc}}(\theta, z_m)$ for the later evaluation of the lifetime data. As the final result will be normalized, anyway, this does not change the calculated lifetimes.

**Molecule detection function**

In order to model the average detected photon arrival time of all molecules inside the focal volume, not only the excitation probabilities of fluorophores at different positions have to be taken into account, but also the probabilities to actually detect emitted photons. The molecule detection function $u(\theta, z_m)$ is proportional to the probability that a photon is detected if a dipole emitter situated at height $z_m$ and with its emission dipole moment at an angle $\theta$ relative to the optical axis undergoes a radiative transition from the excited to the ground state. The proportionality factor is determined by parameters such as the quantum yield of the detector or losses in the optical path between objective and detector, and thus independent of the sample. It will cancel later and is therefore ignored. Physically, a photon can be detected if it is emitted in a direction that is within the objective's cone of light collection, which is defined as all angles $\chi$ relative to the optical axis that are smaller than the angle $\chi_{\max} = \arcsin(NA/n)$ defined by the numerical aperture of the objective and the refractive index $n$ of the objective's immersion oil. Thus, $u(\theta, z_m)$ can be obtained by integrating the angular distribution of radiation of the emitter (ADR, see section 2.4.5) over the solid angle of

---

[17]The probabilities are not symmetric in $z$ because of the different thicknesses of the silver mirrors.

the cone of light collection, and dividing by $S_{\mathrm{tot}}$, i.e. the total amount of energy emitted per time by the dipole:

$$u(\theta, z_m) = \frac{S_{\mathrm{det},\perp}(z_m)\cos^2\theta + S_{\mathrm{det},\parallel}(z_m)\sin^2\theta}{S_\perp(z_m)\cos^2\theta + S_\parallel(z_m)\sin^2\theta}, \tag{4.29}$$

where $S_{\mathrm{det},\perp,\parallel}$ is the result of the integration of the ADR over the cone of light collection for a dipole that is perpendicular or parallel to the planar system. Similarly, $S_{\perp,\parallel}(z_m)$ is $S_{\mathrm{tot}}(z_m)$ for a vertical or horizontal dipole. The large differences between the ADR of dipole emitters with different orientations was already discussed in section 2.4.5, we now only show examples of $u(\theta, z_m)$ for $\theta = 0, \pi/2$ and varying $z_m$ in figure 4.32. The general trend is the same as for the excitation rate: If the dipole moment is oriented parallel to the mirrors, the detection is most efficient in the center of the cavity, while orthogonally oriented dipoles are detected best when they are situated close to the mirrors.

### Rotational diffusion

Up to now, we have always considered a dipole emitter whose dipole moment has a fixed orientation $\theta$ relative to the optical axis. For this case, we derived expressions for the fluorescence lifetime $\tau(\theta, z_m)$, the excitation probability $I(\theta, z_m)$ and the molecule detection function $u(\theta, z_m)$. This assumption is valid if all fluorophores are embedded in a solid, e.g. in optical glue, or if the fluorescing molecules are so large that their rotation is much slower than the decay of their excited state. For small fluorophores in solution, however, this model breaks down. If an excited fluorophore rotates quickly, its coupling to the electric field also changes quickly. Employing the different models presented in section 2.5.3, one could say that the one-dimensional oscillator changes the direction of its oscillation and is thus driven by different components of the reflected electric field, which of course have different magnitudes. Alternatively, one can state that the total amount of emitted energy per time $S_{\mathrm{tot}}$ varies rapidly as the molecule rotates. Both explanations result in the prediction that the overall decay rate is a superposition of the decay rates associated with the orientations $\theta$ that are sampled by the molecule as it rotates. The explicit mathematical description of this process is challenging because the rotational diffusion is intrinsically a stochastic process, and the exact "path" $\theta(t)$ that the dipole moment of a single dipole emitter describes while the emitter is in its excited state is unknown.

In the following, we will use two different approaches to determine the lifetime of a rotating fluorophore in a nanocavity. The first is the analytical calculation used in [107] which employs probability distributions. The second approach is a numerical simulation of many rotating molecules that proves the applicability of the first approach to this situation.

Assume that a fluorophore with dipole moment $\boldsymbol{p}$ is located at the height $z_m$ inside the cavity and shows no lateral diffusion on the timescale of the fluorescence lifetime. At time $t = 0$, the sample is illuminated with a short light pulse. Provided that the

dipole emitter was excited by this light pulse, the probability that its dipole moment was oriented at an angle $\theta$ relative to the optical axis at time $t = 0$ is given by:

$$
\begin{aligned}
p(\theta, z_m, t = 0) &= \frac{I^{\text{exc}}(\theta, z_m)}{\int_0^\pi I^{\text{exc}}(\theta, z_m) \sin\theta \mathrm{d}\theta} \\
&= \frac{3[I_\perp^{\text{exc}}(z_m) \cdot \cos^2\theta + I_\parallel^{\text{exc}}(z_m) \cdot \sin^2\theta]}{2[I_\perp^{\text{exc}}(z_m) + 2I_\parallel^{\text{exc}}(z_m)]}.
\end{aligned}
\tag{4.30}
$$

The function $p(\theta, z_m, t)$ describes the probability that there is an excited dipole emitter at height $z_m$ with orientation $\theta$ at time $t$[18]. As time progresses, there are two processes that can change $p$: The dipole can undergo a (radiative or non-radiative) transition from the excited to the ground state, or it can change its orientation $\theta$. These two possibilities are contained in the following rotational diffusion equation:

$$
\frac{\partial p(\theta, z_m, t)}{\partial t} = \frac{D}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial p(\theta, z_m, t)}{\partial\theta} - K(\theta, z_m) p(\theta, z_m, t).
\tag{4.31}
$$

Here, $D$ is the rotational diffusion constant (linked to the rotational diffusion time $\tau_{\text{rot}}$ via $D = (6\tau_{\text{rot}})^{-1}$) and $K$ is the total deexcitation rate. The latter can always be expressed in the form

$$
\begin{aligned}
K &= k_{\text{nr}} + k_r(z_m, \theta) \\
&= [k_{\text{nr}} + k_{\text{r},\perp}(z_m)] \cos^2\theta + [k_{\text{nr}} + k_{\text{r},\parallel}(z_m)] \sin^2\theta \\
&=: K_\perp(z_m) \cos^2\theta + K_\parallel(z_m) \sin^2\theta,
\end{aligned}
\tag{4.32}
$$

where $k_{\text{nr}}$ is the position- and angle-independent non-radiative deexcitation rate, and $k_{\text{r},\perp}$ and $k_{\text{r},\parallel}$ are the radiative deexcitation rates of a vertical and horizontal dipole at $z_m$, respectively. In equation (4.31), explicitly performing the differentiation with respect to $\theta$ can be avoided by expanding $p(\theta, z_m, t)$ into a series of Legendre polynomials $P_\ell(\cos\theta)$ with time-dependent expansion coefficients $a_\ell(z_m, t)$,

$$
p(\theta, z_m, t) = \sum_{\ell=0}^\infty a_\ell(z_m, t) \, P_\ell(\cos\theta),
\tag{4.33}
$$

because Legendre polynomials fulfill

$$
\begin{aligned}
0 &= \frac{\partial}{\partial x} \left[ (1 - x^2) \frac{\partial P_\ell(x)}{\partial x} \right] + \ell(\ell+1) P_\ell(x) \\
&= \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left[ \sin\theta \frac{\partial P_\ell(\cos\theta)}{\partial\theta} \right] + \ell(\ell+1) P_\ell(\cos\theta).
\end{aligned}
\tag{4.34}
$$

---

[18]Note that the integral $\int_0^\pi \mathrm{d}\theta \, \sin\theta \, p(\theta, z_m, t = 0)$ equals one, in accordance with our assumption: At time $t = 0$, we know that there is an excited dipole at height $z_m$.

Inserted into equation (4.31), this gives:

$$\sum_{\ell=0}^{\infty} \frac{\partial a_\ell(z_m,t)}{\partial t} P_\ell(\cos\theta) = D\sum_{\ell=0}^{\infty} \ell(\ell+1)a_\ell(z_m,t)P_\ell(\cos\theta) - K\sum_{\ell=0}^{\infty} a_\ell(z_m,t)P_\ell(\cos\theta).$$

(4.35)

As a next step, the deexcitation rate $K$ is rewritten as

$$K = \cos^2\theta K_\perp + \sin^2\theta K_\parallel = \cos^2\theta(K_\perp - K_\parallel) + K_\parallel =: \cos^2\theta\Delta K + K_\parallel.$$

(4.36)

By multiplying equation (4.35) by $P_{\ell'}(\cos\theta)$ and integrating over $\sin\theta \mathrm{d}\theta$ from 0 to $\pi$, one can exploit the orthogonality of the Legendre polynomials[19],

$$\int_{-1}^{1} \mathrm{d}x\ P_\ell(x)P_{\ell'}(x) = \frac{2}{2\ell+1}\delta_{\ell\ell'},$$

(4.37)

to transform (4.35) into an infinite set of ordinary differential equations for $a_\ell(z_m,t)$:

$$\frac{\partial a_\ell(z_m,t)}{\partial t} = -D\ell(\ell+1)\ a_\ell(z_m,t) - \sum_{\ell'=0}^{\infty} M_{\ell\ell'}(z_m)a_{\ell'}(z_m,t),$$

(4.38)

with the transition matrix

$$M_{\ell\ell'}(z_m) = \begin{cases} \frac{(\ell-1)\ell}{(2\ell-3)(2\ell-1)}\Delta K & \text{for } \ell' = \ell+2 \\ \frac{2\ell(\ell+1)-1}{(2\ell-1)(2\ell+3)}\Delta K + K_\parallel & \text{for } \ell' = \ell \\ \frac{(\ell+1)(\ell+2)}{(2\ell+3)(2\ell+5)}\Delta K & \text{for } \ell' = \ell-2 \\ 0 & \text{else.} \end{cases}$$

(4.39)

In matrix notation, the set of differential equations for fixed $z_m$ can be compactly written as

$$\frac{\partial}{\partial t}\boldsymbol{a}(t) = -\underline{\underline{\tilde{M}}}\boldsymbol{a} \quad \text{with} \quad \tilde{M}_{\ell\ell'} = M_{\ell\ell'} + \ell(\ell+1)D\delta_{\ell\ell'}.$$

(4.40)

If $\underline{\underline{\tilde{M}}}$ were a diagonal matrix, the solutions would simply be given by $a_\ell(z_m,t) = a_\ell(z_m,0) \cdot \exp[-\tilde{M}_{\ell\ell}(z_m) \cdot t]$. Since this is not the case, we numerically diagonalize $\underline{\underline{\tilde{M}}}$, i.e. we find a matrix $\underline{\underline{V}}$ and a diagonal matrix $\underline{\underline{L}}$ such that $\underline{\underline{\tilde{M}}} = \underline{\underline{V}}\,\underline{\underline{L}}\,\underline{\underline{V}}^{-1}$. Then, the differential equation is given by

$$\frac{\partial}{\partial t}\left(\underline{\underline{V}}^{-1}\boldsymbol{a}\right) = -\underline{\underline{L}}\left(\underline{\underline{V}}^{-1}\boldsymbol{a}\right).$$

(4.41)

Denoting $\boldsymbol{b} := \underline{\underline{V}}^{-1}\boldsymbol{a}$, the solutions are

$$b_\ell(z_m,t) = b_\ell(z_m,0) \cdot \mathrm{e}^{-L_{\ell\ell}t}.$$

(4.42)

---

[19]And use the recursion relation $x \cdot P_\ell = \frac{\ell+1}{2\ell+1}P_{\ell+1} + \frac{\ell}{2\ell+1}P_{\ell-1}$ twice to eliminate the $\cos^2\theta$-term.

The initial values $\boldsymbol{b}(z_m, t = 0)$ can be found by projecting the initial values $\boldsymbol{a}(z_m, t = 0)$ on the eigenvectors of $\underset{\displaystyle\tilde{=}}{\tilde{M}}$,

$$\boldsymbol{b}(z_m, 0) = \underline{\underline{V}}^{-1}\boldsymbol{a}(z_m, 0), \tag{4.43}$$

where $\boldsymbol{a}(z_m, 0)$ can be obtained from the probability distribution $p(\theta, z_m, t = 0)$ given in equation (4.30):

$$a_0(z_m, t = 0) = \frac{1}{2}$$

$$a_2(z_m, t = 0) = \frac{I_\perp^{\mathrm{exc}}(z_m) - I_\parallel^{\mathrm{exc}}(z_m)}{I_\perp^{\mathrm{exc}}(z_m) + 2I_\parallel^{\mathrm{exc}}(z_m)}$$

$$a_\ell(z_m, t = 0) = 0 \qquad\qquad \forall \ell \neq 0, 2. \tag{4.44}$$

The coefficients $\boldsymbol{a}(z_m, t)$ are then obtained from $\boldsymbol{b}(z_m, t)$ via

$$\boldsymbol{a}(z_m, t) = \underline{\underline{V}}\boldsymbol{b}(z_M, t) = \sum_\ell \hat{\boldsymbol{e}}_\ell \sum_{\ell'} V_{\ell\ell'} b_{\ell'}(z_m, 0) \mathrm{e}^{-L_{\ell'\ell'}t}. \tag{4.45}$$

The final result, i.e. the probability distribution $p(\theta, z_m, t)$, is then given by:

$$p(\theta, z_m, t) = \sum_{\ell=0}^\infty P_\ell(\cos\theta) \sum_{\ell'=0}^\infty V_{\ell\ell'}(z_m) b_{\ell'}(z_m, 0) \mathrm{e}^{-L_{\ell'\ell'}t}. \tag{4.46}$$

This function was used in [107] for the evaluation of the measured data. However, since it only makes statements about the probability to find an excited molecule at a certain orientation at time $t$, while a single molecule's dipole moment follows one specific trajectory $\theta(t)$, we wanted to check that the predicted fluorescence lifetime of the ensemble average truly corresponds to the average of many fluorophores with their specific paths $\theta(t)$. In order to do this, we first had to derive the predicted lifetime $\tau(z_m)$ using equation (4.46). The fluorescence lifetime is defined as the average time between the excitation of the fluorophore and the emission of a photon. The probability that a photon is emitted at time $t$ is given by the probability that the molecule undergoes a transition from the excited to the ground state at time $t$, multiplied by the probability that such a transition is radiative (as opposed to non-radiative). The latter is simply the quantum yield $\Phi(z_m, \theta)$, while the former equals the probability that the fluorophore is in its excited state at time $t$, multiplied by the total deexcitation rate. Thus, the number of photons $I_{\mathrm{em}}(z_m, t)$ emitted by a fluorophore at height $z_m$ at time $t$ is proportional to:

$$I_{\mathrm{em}}(z_m, t) \propto \int_0^\pi \mathrm{d}\theta \; \sin\theta \; p(\theta, z_m, t) \cdot K(z_m, \theta) \cdot \Phi(z_m, \theta) \tag{4.47}$$

Note that the quantum yield $\Phi(z_m, \theta)$ is the local quantum yield in the presence of the metal mirrors, not the free space quantum yield which we want to determine with the help

of the cavity. The formula can be simplified by noting that $K(z_m, \theta) \cdot \Phi(z_m, \theta) = k_{\text{r}}(z_m, \theta)$. With this, the average emission time of a photon – and thus the lifetime $\tau$ – is given by

$$
\begin{aligned}
\tau &= \frac{\int_0^\infty \mathrm{d}t \; I(t) \cdot t}{\int_0^\infty \mathrm{d}t \; I(t)} \\
&= \frac{\int_0^\pi \mathrm{d}\theta \sin\theta \; k_{\text{r}}(z_m, \theta) \cdot \sum_\ell P_\ell(\cos\theta) \sum_{\ell'} V_{\ell\ell'} b_{\ell'}(z_m, 0) \int_0^\infty \mathrm{d}t \; t \cdot \exp(-L_{\ell'\ell'} t)}{\int_0^\pi \mathrm{d}\theta \sin\theta \; k_{\text{r}}(z_m, \theta) \cdot \sum_\ell P_\ell(\cos\theta) \sum_{\ell'} V_{\ell\ell'} b_{\ell'}(z_m, 0) \int_0^\infty \mathrm{d}t \; \exp(-L_{\ell'\ell'} t)}.
\end{aligned} \tag{4.48}
$$

The integration over $t$ can be carried out analytically, it yields $L_{\ell'\ell'}^{-2}$ in the numerator and $L_{\ell'\ell'}^{-1}$ in the denominator. Thus, only the integration over $\theta$ has to be performed numerically. Two points deserve to be mentioned here. Firstly, the free space quantum yield $\Phi$ is implicitly contained in this expression via $\underline{V}$ and $\underline{L}$, which were derived from $\underline{\underline{\tilde{M}}}$ with $\tilde{M}_{\ell\ell'} = M_{\ell\ell'} + \ell(\ell+1) D \delta_{\ell\ell'}$, where $M_{\ell\ell'}$ contains the rates $K_\perp$ and $K_\parallel$. The latter are related to the quantum yield $\Phi$ and the free space lifetime $\tau_0$ via

$$
\begin{aligned}
K_{\perp,\parallel} &= k_{\text{nr}} + k_{\text{r},\perp,\parallel}(z_m) \\
&= k_{\text{nr}} + k_{\text{r},\perp,\parallel,0} \cdot \frac{S_{\text{tot},\perp,\parallel}(z_m)}{S_0} \\
&= \left[ (1 - \Phi) + \Phi \cdot \frac{S_{\text{tot},\perp,\parallel}(z_m)}{S_0} \right] \cdot \tau_0^{-1}.
\end{aligned} \tag{4.49}
$$

In the same way, $\Phi$ is also contained in the term $k_{\text{r}}(z_m, \theta) = k_{\text{r},\perp}(z_m) \cos^2\theta + k_{\text{r},\parallel}(z_m) \sin^2\theta$. Secondly, the relative magnitudes of $D$ and $K$ have an impact on $\underline{\underline{\tilde{M}}}$, and as a consequence also on $\tau$. Figure 4.33 displays the average relative lifetime $\tau/\tau_0$ for three different free space quantum yields and for varying rotational diffusion times. For the calculation, we assume a DMSO-filled nanocavity with height $h = 200\,\text{nm}$, an emission wavelength of $\lambda = 650\,\text{nm}$, and a fluorophore at height $z_m = 100\,\text{nm}$. The dashed lines represent the case of purely horizontal excitation ($I_\perp^{\text{exc}} = 0$, $p(\theta, t = 0) = 3/4\,\sin^2\theta$), while the solid lines represent purely vertical excitation ($I_\parallel^{\text{exc}} = 0$, $p(\theta, z_m, t = 0) = 3/2\,\cos^2\theta$). For short rotational diffusion times, both curves converge to a common limit, namely

$$
\tau_{\text{fast rot.}}(z_m) = \frac{\tau_0}{1 - \Phi + \Phi \cdot [2 S_{\text{tot},\parallel}(z_m) + S_{\text{tot},\perp}(z_m)]/3 S_0}. \tag{4.50}
$$

Thus, the decay rate of the excited state of a quickly rotating fluorophore is obtained by averaging over the decay rates of fluorophores with all possible orientations, where the orientations are uniformly distributed. Contrarily, at the limit of extremely slow rotation, the calculated average lifetime is identical to the average lifetime of an ensemble of fixed dipoles, where the orientations follow $p(\theta, z_m, t = 0)$:

$$
\tau_{\text{slow rot.}}(z_m) = \int_0^\pi \mathrm{d}\theta \; \sin\theta \; \frac{\tau_0}{1 - \Phi + \Phi \cdot [S_{\text{tot}}(\theta, z_m)/S_0]} \cdot p(\theta, z_m, t = 0). \tag{4.51}
$$

These predictions, which were derived employing the theory published in [107], were tested by means of a numerical simulation as part of the present work. The core idea is to simulate the rotational diffusion of the dipole moment $\boldsymbol{p}$ concurrently with the
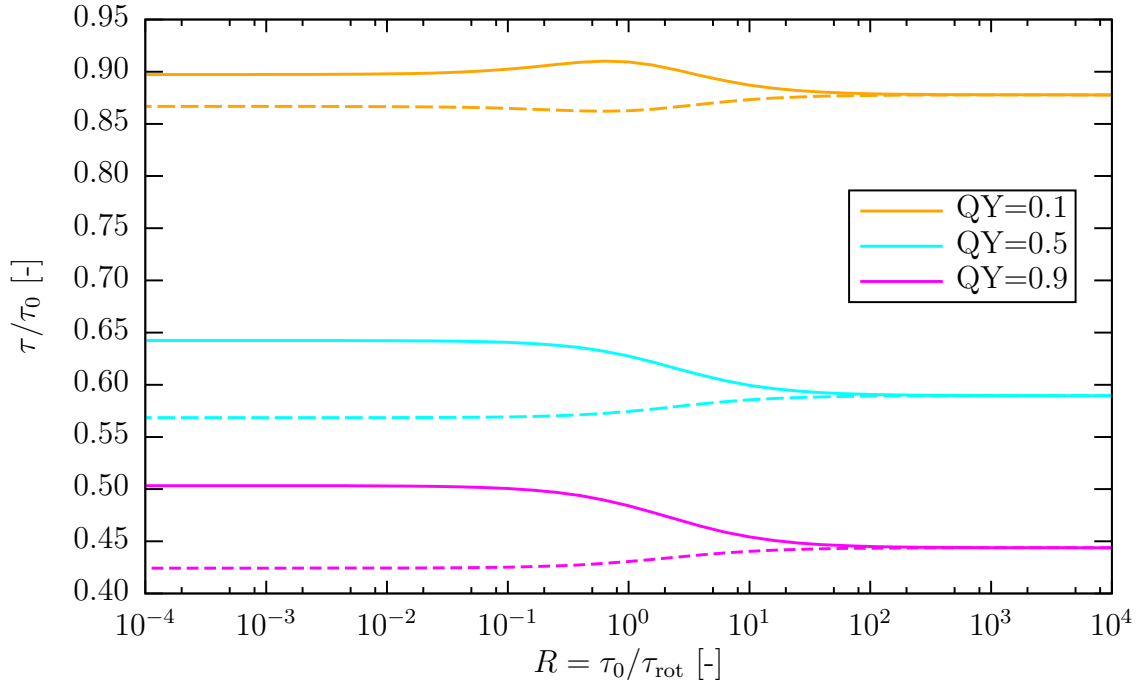
**Figure 4.33:** Simulation of the average excited state lifetime of a fluorophore with quantum yield $\Phi = 0.1$ (orange), 0.5 (cyan) or 0.9 (pink) that is situated at $z_m = 100\,\mathrm{nm}$ inside a DMSO-filled nanocavity with height $h = 200\,\mathrm{nm}$ and mirror thicknesses $30\,\mathrm{nm}$ and $60\,\mathrm{nm}$, assuming either purely vertical (solid lines) or purely horizontal excitation (dashed lines). The emission wavelength is $\lambda = 650\,\mathrm{nm}$.

(radiative or non-radiative) decay of the excited state, taking into account that the decay rate depends on the angle $\theta$ between $\boldsymbol{p}$ and the optical axis. The resulting decay times $t_{\mathrm{decay}}$ of many such simulation runs are then histogrammed, and an average lifetime $\tau_{\mathrm{sim}}$ is determined as

$$\tau_{\mathrm{sim}} = \frac{1}{M} \sum_{m=1}^{M} t_{\mathrm{decay}}(m). \tag{4.52}$$

First, an algorithm to simulate the rotational diffusion of $\boldsymbol{p}$ had to be designed. The orientation of $\boldsymbol{p}$ is completely defined by the angle $\theta$ with the optical axis and the angle $\varphi$ with the $x$-axis. Thus, the rotational diffusion of $\boldsymbol{p}$ can be modeled as a random walk on the unit sphere, with the trajectory $\{(\theta_i, \varphi_i)\}$, where the index $i$ denotes the time steps. Since the diffusion is isotropic, at each time point $i$, the next step can be taken in any direction with equal probability. Mathematically, this can be defined straightforwardly if the position at time $i$ is the "North pole" (i.e. $\theta_i = 0$) by randomly choosing a longitude between 0 and $2\pi$, and then taking a step of fixed step length $l$. Based on this concept, our procedure for the rotational diffusion is (see figure 4.34):

1. At time $i$, the position of the random walker on the unit sphere is defined by the angles $(\theta_i, \varphi_i)$.

1. Position of walker at time $i$      2./3. Step in random direction at North pole: $N \to P$



4. Rotate coordinate system      5. Append $\mathcal{R}(P)$ to trajectory
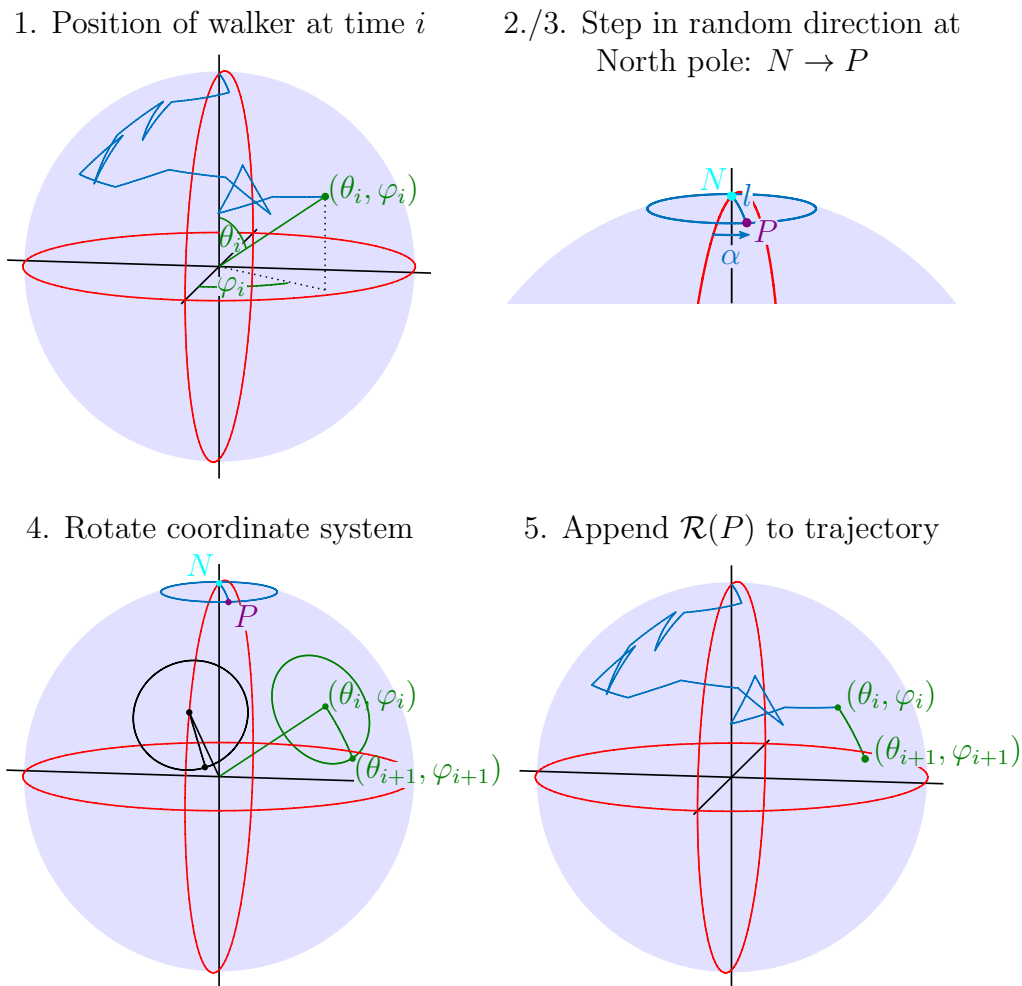


**Figure 4.34:** Algorithm for simulating the rotational diffusion of the dipole moment $\boldsymbol{p}$. **1.** At time $i$, the position is described by the angles $(\theta_i, \varphi_i)$. **2.** A random meridian at longitude $\alpha$ is chosen. **3.** Starting at the North pole $N$, a step with step lenght $l$ is taken along the meridian, ending at the point $P$. **4.** The coordinate system is rotated such that $\mathcal{R}(N) = (\theta_i, \varphi_i)$. **5.** The rotation of $P$ yields the next point of the random walker's trajectory, $(\theta_{i+1}, \varphi_{i+1}) = \mathcal{R}(P)$.

2. For the transition $i \to i+1$, draw a random number $\alpha$ from a uniform distribution between 0 and $2\pi$.

3. Take a step of step length $l$ from the North pole $N = (0,0)$ along the meridian with longitude $\alpha$, ending in the point $P = (l, \alpha)$[20].

4. Rotate the coordinate system such that the North pole is moved to $(\theta_i, \varphi_i)$, $\mathcal{R}(N) = (\theta_i, \varphi_i)$.

5. By applying the same rotation to $P$, the next position of the random walker is obtained, $\mathcal{R}(P) = (\theta_{i+1}, \varphi_{i+1})$.

---

[20]Since the walker moves on the unit sphere, the arc length $l$ of a step along a meridian equals the angle $\theta$ that is covered by this step.

This procedure was tested and characterized by simulating a number of trajectories $\{(\theta_i, \varphi_i)\}$, creating $\theta_i$-distributions at several time points $i$, and comparing with the analytical probability distribution $p(\theta_i)$. The distribution of start values $(\theta_0, \varphi_0)$ was given as

$$p(\theta_0, \varphi_0) = p(\varphi_0) \cdot p(\theta_0) = \frac{1}{2\pi} \cdot \frac{3}{2} \cos^2 \theta_0 = \frac{1}{2\pi} \cdot \left[ \frac{1}{2} P_0(\cos\theta_0) + P_2(\cos\theta_0) \right], \quad (4.53)$$

corresponding to excitation with vertically polarized light. The pure rotational diffusion equation (without the decay constant $K$) results in the differential equations

$$\frac{\partial a_\ell(t)}{\partial t} = -D\ell(\ell+1)a_\ell(t), \quad (4.54)$$

thus, $a_0 = \mathrm{const.} = 1/2$ and

$$a_2(t) = a_2(0) \cdot \mathrm{e}^{-6Dt} = \mathrm{e}^{-6Dt}, \quad (4.55)$$

while all other $a_\ell(t)$ equal zero. By fitting the discrete $\theta_i$-distributions $H(\theta_i)$ with

$$H(\theta_i) = \frac{1}{2}P_0(\cos\theta_i) + a_2(t_i)P_2(\cos\theta_i) \quad (4.56)$$

with $a_2(t_i)$ as the only fit parameter, and then using

$$D = -\frac{\log[a_2(t_i)]}{6t_i} \quad (4.57)$$

we could therefore obtain a relationship between $D$ and the step length $l$, the only adjustable parameter of our algorithm. Figure 4.35(a) exemplarily shows $\theta_i$-distributions $H(\theta_i)$ for $i = 0, 10, 20$ and $30$ and a step length $l = 0.2$, together with the best fits of equation (4.56) to the data. The resulting fit values of $D$ for $i \geq 1$ were averaged to find $D(l = 0.2)$. At very long times, the distribution $p(\theta_i)$ approaches $P_0(\cos\theta_i)$ as $a_2(\theta_i) \to 0$, then the fit of $a_2(\theta_i)$ becomes unreliable. At which time point the fit breaks down depends on the magnitude of $D$ and thus on the step length $l$. Therefore, the number of time points that were included in the calculation of $D$ was different for each step length $l$. Figure 4.35(b) shows an overview of diffusion constants for step lengths between 0.01 and 0.5. For small $l$, $\sqrt{D}$ rises linearly with $l$:

$$D(l) = \frac{1}{4} l^2 \left[ \frac{1}{\text{time step}} \right] \qquad \text{for } l \leq 0.2. \quad (4.58)$$

For step lengths between 0.3 and 0.5, $D(l)$ starts to become larger than $l^2/4$. Therefore, we only used $l \leq 0.3$ in the subsequent simulations.

For the simultaneous simulation of the rotational diffusion of $\boldsymbol{p}$ and the decay of the excited state, several parameters had to be chosen, starting with the ratio $R$ between the free space fluorescence lifetime $\tau_0$ and the rotational diffusion time $\tau_{\mathrm{rot}}$, $R = \tau_0/\tau_{\mathrm{rot}}$.

a)



b)
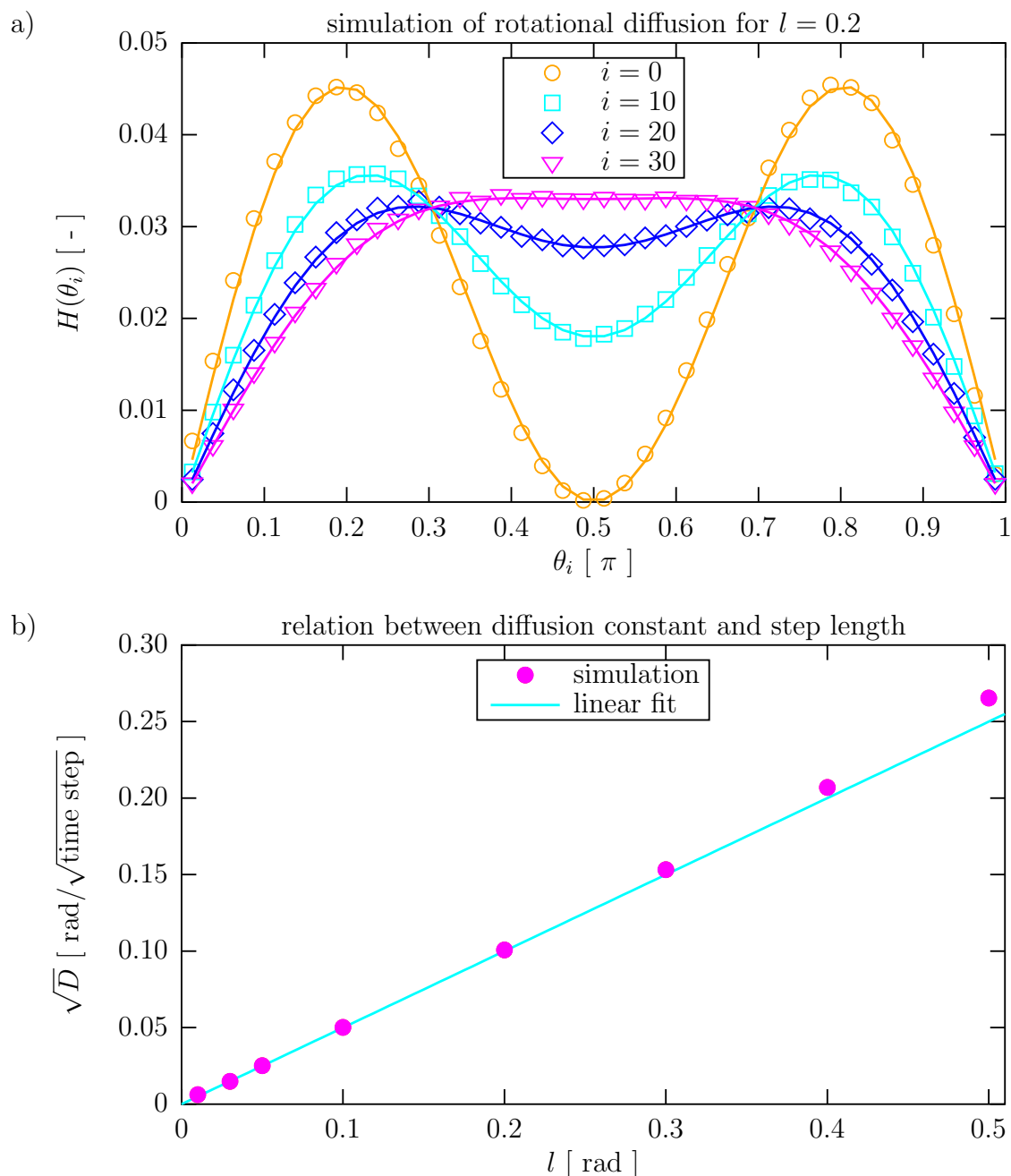


**Figure 4.35:** Simulations to determine the relation between the step length $l$ of the rotational diffusion algorithm and the diffusion constant $D$. For each step length, $10^6$ trajectories with 41 time steps $i = 0, \ldots, 40$ were simulated, starting with the distribution $p(i = 0) = \frac{1}{2} P_0(\cos \theta_0) + P_2(\cos \theta_0)$. At each time point $i$, the $10^6$ coordinates $(\theta_i, \varphi_i)$ were binned with respect to $\theta_i$, resulting in distributions $H(\theta_i)$ as shown in (a) for $l = 0.2$. Fits of equation (4.56) to $H(\theta_i)$ for $i \geq 1$ yielded 40 estimates of $D$ for each step length $l$. Fits at late time points were unreliable because $a_2(t_i) \to 0$, thus for each step length, a different subset of these 40 values was used to determine the average diffusion constant $D$. In (b), the square root of $D$ is plotted against $l$ (pink dots), the standard deviations are smaller than the marker size and therefore not shown as errorbars. The blue line corresponds to $l/2$, showing that for small step lengths, we have the relation $D = l^2/4$.
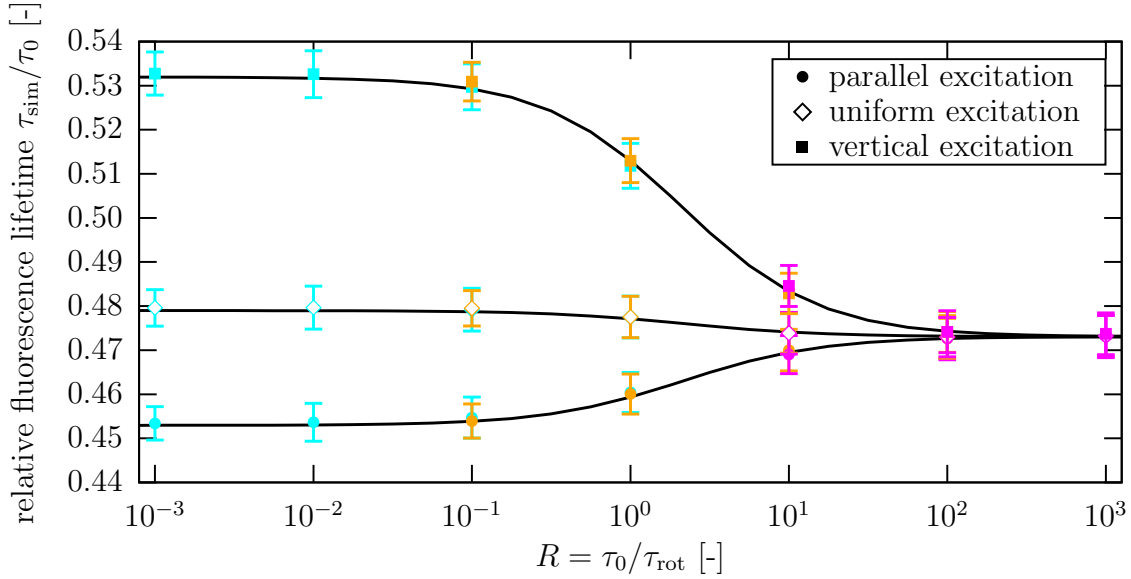
**Figure 4.36:** Comparison of simulated average photon arrival times $\tau_{\text{sim}}$ (points with errorbars) and the values predicted by the probability-distribution based approach in equation (4.48) (black lines). The initial distribution at time zero corresponded to vertically polarized excitation light (squares), horizontally polarized light (circles) or uniform polarization (open diamonds). For the simulation, step lengths of $l = 0.01$ (cyan), $l = 0.1$ (orange) and $l = 0.3$ (magenta) were employed to cover different $R$-ranges without having to choose too small excited state decay rates $k_{\text{tot}}$. Over the whole $R$-range, simulations with different step length for the same $R$-value are consistent, and all simulated values agree excellently with the analytical predictions. It becomes apparent that the lifetime depends strongly on the polarization of the excitation light for fixed or slowly rotating fluorophores (small $R$), while all curves converge to the same value $\tau_{\text{fast rot.}}$ (eq. 4.50) for quickly rotating fluorophores. Simulation for a DMSO-filled cavity with $h = 200\,\text{nm}$, silver mirror thicknesses $30\,\text{nm}$ and $60\,\text{nm}$, and a fluorophore at height $z = 100\,\text{nm}$ with $\Phi = 0.8$.

Based on $R$ and the step length $l$ of the rotational diffusion, the total decay rate of the fluorophore's excited state in units of [time step$^{-1}$] was determined:

$$k_{\text{tot}} = \frac{1}{\tau_0} = \frac{1}{R \cdot \tau_{\text{rot}}} = \frac{6D(l)}{R} = \frac{3l^2}{2R}. \tag{4.59}$$

The simulation of a single excited molecule lasted until the excited state decayed, or a pre-defined number of time steps was exceeded. In order to prevent a bias towards shorter decay times, this number of time steps had be sufficiently large. We chose $8/k_{\text{tot}}$, since $\exp(-8) \approx 3.4 \cdot 10^{-4}$ was deemed suffciently small. As a small decay rate entails many time steps before a decay event, it is desirable from a numerical point of view to avoid too small numbers of $k_{\text{tot}}$. Therefore, step lengths of $l = 0.01, 0.1$ or $0.3$ were employed to cover different $R$-ranges from $10^{-3}$ to $10^3$. The Matlab code used for the simulation is given in section 6.1.4 in the appendix. For each parameter set $(l, R)$, 1000 "experiments" were simulated, each consisting of the simulation of 5000

"fluorophores" whose decay times were averaged to obtain $\tau_{\mathrm{sim}}$. The averages of these 1000 values for $\tau_{\mathrm{sim}}$, together with their standard deviations, are depicted in figure 4.36. They are compared with the average lifetimes $\tau(z_m)$ predicted by the probability distribution-based approach in equation (4.48), shown as solid black lines. Both agree very well over the whole $R$-range. Moreover, we find that the simulated values of $\tau_{\mathrm{sim}}$ converge towards the limits $\tau_{\mathrm{fast\ rot.}}$ and $\tau_{\mathrm{slow\ rot.}}$ from equations (4.50) and (4.51) for fast and slow rotation, respectively. It becomes apparent that the polarization of the excitation light, and thus the initial distribution of orientation of $\boldsymbol{p}$, has a strong impact on the average fluorescence lifetime when fluorophores are fixed or only slowly rotating. In contrast, quickly rotating fluorophores sample the whole $\boldsymbol{p}$-orientation space several times before decaying, resulting in lifetimes that are independent of the initial $\theta$-distribution.

Due to this confirmation of the probabilistic approach, and because it is computationally much faster than the single-fluorophore simulations, it was employed for the nanocavity-based determination of fluorescence quantum yield values as described below.

**Measured average arrival time**

We will now combine the results from the previous paragraphs to predict the measured average photon arrival time $\tau_{\mathrm{meas}}$ for a given sample geometry and given optical properties of the fluorophore. We start by revisiting equation (4.47), which describes the time dependence of the number of photons *emitted* by a fluorophore situated at $z_m$:

$$I_{\mathrm{em}}(z_m, t) \propto \int_0^\pi \mathrm{d}\theta \ \sin\theta \ p(\theta, z_m, t) \cdot k_{\mathrm{r}}(z_m, \theta).$$

The number of photons that is actually detected is lower than $I_{\mathrm{em}}$: The molecule detection function $u(\theta, z_m)$ specifies which fraction of emitted photons enters the objective. Additionally, the detection efficiency of the detector and possible losses in the optical path between objective and detector could be taken into account, however, since they are independent of the sample itself, they cancel in the calculation of $\tau_{\mathrm{meas}}$. Thus, we find for the number of photons $I_{\mathrm{det}}$ detected from a photon at height $z_m$:

$$I_{\mathrm{det}}(z_m, t) \propto \int_0^\pi \mathrm{d}\theta \ \sin\theta \ p(\theta, z_m, t) \cdot k_{\mathrm{r}}(z_m, \theta) \cdot u(z_m, \theta). \tag{4.60}$$

Note that, when inserting the definitions of $k_r$ and $u$, the formulas simplify due to

$$k_r(\theta, z_m) \cdot u(\theta, z_m) = \frac{S_{\mathrm{tot},\perp} \cos^2\theta + S_{\mathrm{tot},\parallel} \sin^2\theta}{S_0} \cdot \frac{S_{\mathrm{det},\perp} \cos^2\theta + S_{\mathrm{det},\parallel} \sin^2\theta}{S_{\mathrm{tot},\perp} \cos^2\theta + S_{\mathrm{tot},\parallel} \sin^2\theta}$$

$$= \frac{S_{\mathrm{det},\perp}(z_m) \cos^2\theta + S_{\mathrm{det},\parallel}(z_m) \sin^2\theta}{S_0}. \tag{4.61}$$

All calculations performed so far assumed a single emission wavelength $\lambda$. However, real fluorophores can emit with different probabilities at various wavelengths, described by the emission spectrum $F(\lambda)$[21]. As described in section 2.5.1, this is due to the presence

---

[21]In the following, we assume that the emission spectrum has been normalized, i.e. $\sum_\lambda F(\lambda) = 1$.

of many vibrational and rotational sublevels in the electronic ground state, which all have slightly different energies. In an exact treatment, the radiative decay rate has to be viewed as the sum of the rates of all possible transitions from the excited state to any of these sublevels of the ground state. Since MIET is wavelength-dependent, each of these rates is changed by a slightly different amount by the presence of the nanocavity:

$$
\begin{aligned}
k_{\mathrm{r}}(\theta, z_m) &= \sum_\lambda k_{\mathrm{r},\lambda}(\theta, z_m) \\
&= \sum_\lambda k_{\mathrm{r},\lambda,0} \cdot \frac{S_{\mathrm{tot},\lambda}(\theta, z_m)}{S_{0,\lambda}} \\
&= k_{\mathrm{r},0} \cdot \sum_\lambda F(\lambda) \cdot \frac{S_{\mathrm{tot},\lambda}(\theta, z_m)}{S_{0,\lambda}}.
\end{aligned}
\tag{4.62}
$$

Equivalently, the molecule detection function is different for each $\lambda$. Since $p(\theta, z_m, t)$ contains the radiative decay rate, this probability distribution, too, depends on $\lambda$. Therefore, the correct expression for the number of detectable photons originating from a fluorophore at height $z_m$ is:

$$
\langle I_{\mathrm{det}} \rangle_\lambda(z_m, t) \propto \sum_\lambda F(\lambda) \int_0^\pi \mathrm{d}\theta \, \sin\theta \, p_\lambda(\theta, z_m, t) \cdot k_{\mathrm{r},\lambda}(z_m, \theta) \cdot u_\lambda(z_m, \theta).
\tag{4.63}
$$

In practice, the calculation of $p_\lambda(z_m, \theta, t)$ is one of the most time-consuming steps due to the matrix diagonalization. The repeated determination of this probability distribution can be avoided by assuming that the emission spectrum is not too broad, and the variation of the radiative rates therefore not too large. Then, the wavelength-averaged radiative rate

$$
\langle k_{\mathrm{r}} \rangle(\theta, z_m) = \left( \sum_\lambda k_{\mathrm{r},\perp,\lambda}(z_m) \right) \cos^2\theta + \left( \sum_\lambda k_{\mathrm{r},\|,\lambda}(z_m) \right) \sin^2\theta
\tag{4.64}
$$

can be used to calculate the wavelength-averaged distribution $\langle p \rangle_\lambda(z_m, \theta, t)$. Thus, we have:

$$
\langle I_{\mathrm{det}} \rangle_\lambda(z_m, t) \propto \int_0^\pi \mathrm{d}\theta \, \sin\theta \, \langle p \rangle_\lambda(\theta, z_m, t) \cdot \sum_\lambda F(\lambda) k_{\mathrm{r},\lambda}(z_m, \theta) \cdot u_\lambda(z_m, \theta).
\tag{4.65}
$$

It remains to take into account that the detector collects photons from the whole focal volume simultaneously by summing over $z_m$. Since the rotational diffusion calculations were performed under the condition $\int_0^\pi \mathrm{d}\theta \sin\theta p(z_m, \theta, t = 0) = 1$, i.e. with the assumption that there was an excited molecule at height $z_m$ at time zero, the relative weights for the different heights $z_m$ have yet to be determined. This is can be achieved by realising that the probability to excite a molecule at height $z_m$ is proportional to the

intensity of the electric field at $z_m$. Thus, the needed weights are simply proportional to $I_\perp^{\mathrm{exc}}(z_m) + 2I_\parallel^{\mathrm{exc}}(z_m)$:

$$\langle I \rangle_{\lambda, z_m}(t) = \int_0^h \mathrm{d}z_m \langle I_{\mathrm{det}} \rangle_\lambda(z_m, t) \cdot \left[ I_\perp^{\mathrm{exc}}(z_m) + 2I_\parallel^{\mathrm{exc}}(z_m) \right]. \tag{4.66}$$

With this, we arrive at the desired expression for the average measured photon arrival time $\tau_{\mathrm{meas}}$:

$$\tau_{\mathrm{meas}} = \frac{\int_0^\infty \mathrm{d}t \; t \cdot \langle I \rangle_{\lambda, z_m}(t)}{\int_0^\infty \mathrm{d}t \; \langle I \rangle_{\lambda, z_m}(t)}. \tag{4.67}$$

After the rigorous and thus rather complex derivation of this equation, it is worth to recapitulate which parameters influence $\tau_{\mathrm{meas}}$. The refractive indices and thicknesses of all sample layers – the glass cover slip and the convex lens, the silver layers serving as mirrors, and the medium between the mirrors containing the fluorophores – influence $k_{\mathrm{r}}$, $u$ and $I^{\mathrm{exc}}$. These parameters are assumed to be known. Another member of this class of parameters is the height $h$ of the cavity, which is determined by measuring a white light transmission spectrum such as the ones shown in figure 4.28 and fitting the data with $h$ as the free parameter. The other group of parameters concerns the fluorophore itself, it consists of the quantum yield $\Phi$, the free space lifetime $\tau_0$ and the rotational diffusion time $\tau_{\mathrm{rot}}$. It is the aim of the nanocavity measurements to determine $\Phi$. In principal, the other two parameters can be determined by additional experiments. Alternatively, they are kept as free parameters in the fit. Then, comparing the fit values of $\tau_0$ and $\tau_{\mathrm{rot}}$ with results obtained by other methods can serve as a control of the whole fitting process, including the result for $\Phi$.

Before introducing the analysis software for nanocavity measurements that was developed as a part of this work, we present a short example of real data. The squaraine-based membrane-staining dye dSQ12S developed by the group of Klymchenko *et al.* [144] was shown to fluoresce brightly in dioxane and DMSO, as well as when it is incorporated in a lipid bilayer, but to be almost dark when dissolved in water. This makes dSQ12S interesting for staining cell membranes with a low background signal. The quantum yield of this dye was only determined by comparison with the commercially available dye DiD (1,1'-dioctadecyl-3,3,3',3'-tetramethylindodicarbocyanine perchlorate), and the result was published without specifying the details of the measurement [144]. Therefore, we performed a nanocavity-based measurement of dSQ12S in DMSO ($n = 1.48$), with cavity mirror thicknesses of $30\,\mathrm{nm}$ (bottom silver mirror evaporated on glass cover slip) and $60\,\mathrm{nm}$ (top silver mirror evaporated on convex lens). The spectrum of dSQ12S in DMSO as measured by us is given in figure 6.5 in the appendix. For the main experiment, the cavity was moved laterally to allow the acquisition of TCSPC histograms and white light transmission spectra at various cavity heights $h$. By fitting the absolute square of the effective Fresnel transmission coefficient $|t_{p,s}(\lambda)|^2$ to the transmission spectrum with $h$ as the free parameter, the cavity heights were determined. In figure 4.37, the average photon arrival times $\tau_{\mathrm{meas}}$ are plotted against $h$. The errorbars indicate the estimated uncertainties of the lifetimes, determined as $\Delta\tau_{\mathrm{meas}} \approx 4.8\tau_{\mathrm{meas}}/\sqrt{N}$, where $N$ is the number of photons used to determine $\tau_{\mathrm{meas}}$ ([4] and section 4.1.3). Note that the signal is strongest when the emission maximum of the dye is in resonance with
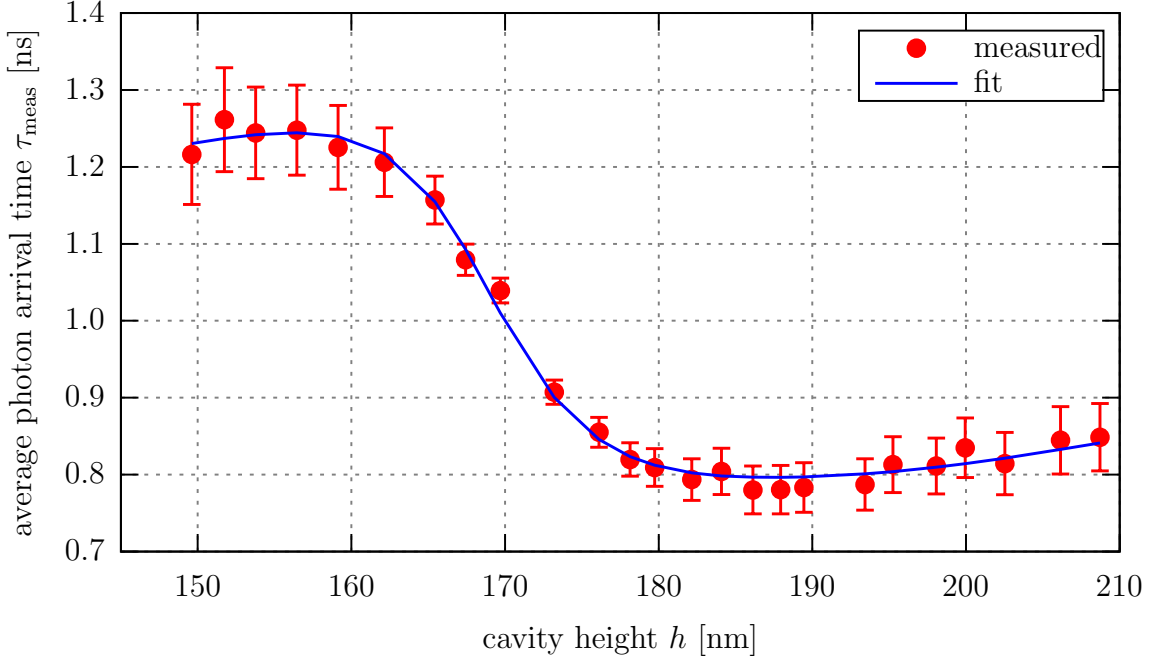
**Figure 4.37:** Quantum yield measurement of the squarine-based dye dSQ12S. Measured average photon arrival time $\tau_{\text{meas}}$ plotted against cavity height $h$ (red circles), lifetime uncertainty $\Delta\tau_{\text{meas}}$ indicated by errorbars. Solid blue line corresponds to the best fit of equation (4.67) to the data. Fit parameters: $\Phi = 0.49$, $\tau_0 = 1.4\,\text{ns}$ and $\tau_{\text{rot}} = 0.08\,\text{ns}$. The nanocavity consisted of two silver mirrors (thicknesses $30\,\text{nm}$ and $60\,\text{nm}$) on glass, separated by DMSO ($n = 1.48$) containing the dye. Data recorded by Alexey Chizhik.

the cavity, and grows weaker as the cavity height is changed. Therefore, the lifetime values obtained at the edges of the measurement interval are known with less accuracy than those in the center of the interval. The solid line corresponds to the best fit of equation (4.67) to the data, with $\Phi$, $\tau_0$ and $\tau_{\text{rot}}$ as fit parameters. We found $\Phi = 0.49$, $\tau_0 = 1.4\,\text{ns}$ and $\tau_{\text{rot}} = 0.08\,\text{ns}$ (i.e. $R = \tau_0/\tau_{\text{rot}} \sim 20$). In their original publication [144], Klymchenko *et al.* claimed a quantum yield of 0.73 in DMSO. The plausibility of these two conflicting results can be tested by comparing with an independent measurement: We stained the plasma membrane of blood platelets with dSQ12S and found a lifetime of $\tau_{0,\text{cell}} = 2.5\,\text{ns}$ in the absence of any metal. The refractive index of the medium was determined to be $n_{\text{cell}} = 1.33$, as opposed to $n_{\text{DMSO}} = 1.48$. The empty-cavity model (see section 4.1.4) therefore predicts a quantum yield of

$$\Phi_{\text{cell}} = \Phi_{\text{DMSO}} \cdot \frac{\tau_{0,\,\text{cell}}}{\tau_{0,\,\text{DMSO}}} \cdot \frac{n_{\text{cell}}^5}{n_{\text{DMSO}}^5} \cdot \frac{(2n_{\text{DMSO}}^2 + 1)^2}{(2n_{\text{cell}}^2 + 1)^2} \approx 0.72 \qquad (4.68)$$

inside the cell. By inserting Klymchenko's value $\Phi_{\text{DMSO}} = 0.73$ instead, we find a predicted quantum yield of $\Phi_{\text{cell}} \approx 1.03$, which is impossible. We therefore conclude that our measurement was more accurate than that reported in [144].

## 4.3.2 Numerical implementation: graphical user interface

As the previous section has shown, the extraction of a quantum yield value from lifetime measurements in a nanocavity requires a considerable knowledge of the underlying theory. In order to enable non-expert scientists to perform and evaluate this type of quantum yield measurements, a Matlab-based software package equipped with a graphical user interface (GUI) similar to the MIET GUI was created. Figure 4.38 presents a screenshot and explains all relevant elements. Notably, the user can choose between a *Standard mode* that assumes a simple cavity consisting of two metal mirrors and one type of material in between, the latter containing the fluorophore, and an *Expert mode* that enables the definition of more complex sample geometries. Furthermore, prior knowledge about the free space lifetime and the rotational diffusion time can be incorporated in the *Expert mode* in the form $\tau_0 \pm \Delta\tau_0$ or $\tau_{\text{rot}} \pm \Delta\tau_{\text{rot}}$, respectively.

The data evaluation starts with the determination of the cavity heights $h$ from the transmission spectra. Since the white light lamp generally does not provide a uniform intensity of all wavelengths, and since the detection efficiency of the detector is usually wavelength dependent, before fitting, all transmission spectra are divided by a transmission spectrum recorded without the sample. The resulting heights $h$, together with the predefined sample parameters, are then used to calculate the excitation rates $I_\perp^{\text{exc}}(z_m, h)$ and $I_\parallel^{\text{exc}}(z_m, h)$, the relative radiative rates $S_{\text{tot},\perp}(z_m, h)/S_0$ and $S_{\text{tot},\parallel}(z_m, h)/S_0$, and the relative amount of detected radiation $S_{\text{det},\perp}(z_m)/S_0$ and $S_{\text{det},\parallel}(z_m)/S_0$. This greatly speeds up the subsequent fitting of the lifetime data to equation (4.67). The fit itself employs a Nelder-Mead downhill simplex method [145] to minimize the sum of the squared residuals of the lifetime values. Finally, a figure is produced showing the fitted lifetime values together with the measured values to allow a visual inspection of the fit quality, see figure 4.37 for an example.

If desired, the fitting is repeated for one hundred simulated data sets which are based on the measured data set to obtain estimates of the uncertainties of the fit parameters (*bootstrapping*). The individual simulated data sets only differ in the lifetimes $\tau_{\text{meas}}(h)$. If the measurement uncertainty is known, these lifetime values are drawn from Gaussian distributions with mean values $\tau_{\text{meas}}(h)$ and standard deviations $\Delta\tau_{\text{meas}}(h)$. Otherwise, data sets are produced by randomly discarding 20 % of the data points and only fitting the remaining 80 %. The latter technique is less reliable than the former because the estimated uncertainties depend on the fraction of discarded data points, however, it is able to reveal the order of magnitude of the uncertainties. In the last part of this section, a general analysis of the accuracy of the determined quantum yield values will be presented.

This software was used to evaluate a large number of experiments during the course of this work. One published example is given in the next section.
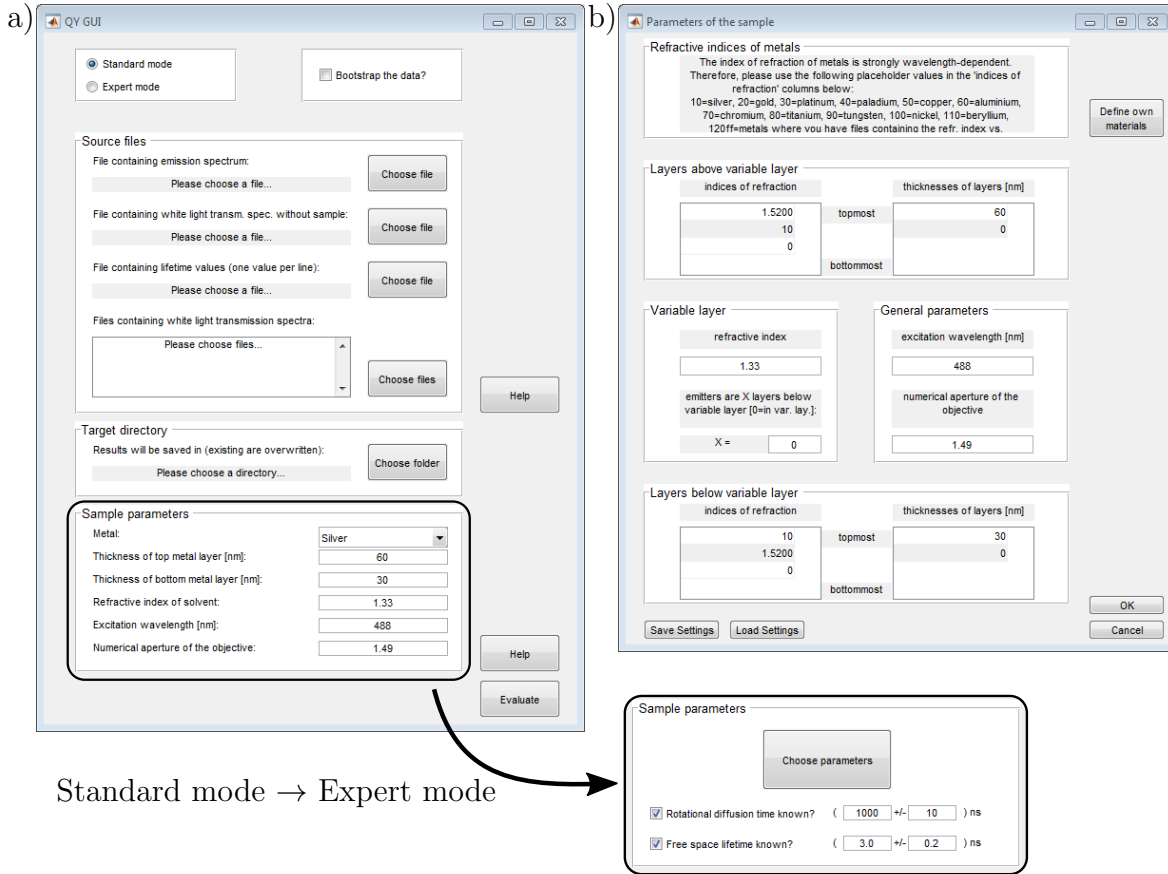
**Figure 4.38:** Screenshot of the graphical user interface for the evaluation of quantum yield measurements. a) Main window. The user can switch between the *Standard mode* and an *Expert mode* which allows unusual sample compositions. In both cases, the files containing the emission spectrum of the fluorophore, a white light transmission spectrum acquired without the sample, a list of all measured lifetimes, and all white light transmission spectra have to be defined. In the standard mode, the cavity is assumed to consist of a metal-coated cover slip and a metal-coated lens with a fluorescent liquid or solid in between. Then, the type and thickness of the metal, the refractive index of the medium containing the fluorophores, and the numerical aperture have to be chosen. Alternatively, in the expert mode, one can freely define a stratified system: By clicking the button *Choose parameters*, the second window (b) opens. Here, two stacks of layers can be defined, the bottom one containing a cover slip and the top one situated on a lens. In between, there is a *variable layer*, whose height varies due to the curvature of the lens. In this mode, it is possible that the fluorophores are not inside the variable layer, e.g. when they are confined to a lipid bilayer that is located at the bottom of a water-filled cavity. Furthermore, the expert mode enables the user to incorporate prior knowledge about the free space lifetime or the rotational diffusion time. In both operation modes, it is possible to obtain estimates of the uncertainties of the fit parameters by ticking the box *Bootstrap the data?*. If the lifetime file also contains lifetime uncertainties, several data sets are simulated whose lifetimes at height $h$ are randomly drawn from Gaussian distributions with means $\tau_{\mathrm{meas}}(h)$ and standard deviations $\Delta\tau_{\mathrm{meas}}(h)$. The spread of the resulting fit parameters then yields $\Delta\Phi$, $\Delta\tau_0$ and $\Delta\tau_{\mathrm{rot}}$. Contrarily, when no information regarding $\Delta\tau_{\mathrm{meas}}$ exists, bootstrapping is done by fitting data sets containing only 80 % of the data points (chosen randomly).

### 4.3.3 Quantum yield measurements of fluorophores in lipid bilayers using a plasmonic nanocavity

One of the main advantages of the nanocavity-based quantum yield measurement is its flexibility regarding the sample type. For example, the required amount of sample in a nanocavity is very small, since typical cavity heights range from 100 nm to 200 nm, enabling to study compounds that are expensive or difficult to synthesize. In contrast, common techniques using integrating spheres or spectrometers usually employ cuvettes of 10 mm × 10 mm [137]. The sandwich-structure of the sample also allows more complex sample compositions, for example fluorophores inside a lipid membrane. This was used to full capacity in a study comparing the quantum yields of two dyes in water and in a lipid membrane, which is summarized in this section:

> [2] F. Schneider, **D. Ruhlandt**, I. Gregor, J. Enderlein, and A. I. Chizhik, "Quantum yield measurements of fluorophores in lipid bilayers using a plasmonic nanocavity," *Journal of Physical Chemistry Letters*, vol. 8, no. 7, pp. 1472–1475, 2017. F. S. and **D. R.** contributed equally to this work.

The experiments were performed by Falk Schneider and Alexey Chizhik, I provided the data analysis software, and we all analyzed the data.

**Methods**

The setup used for the nanocavity measurements is shown in figure 4.39. It differs from the standard FLIM setup previously shown in this work in two respects: Firstly, a white light source is placed above the sample, and a flipping mirror allows to guide the transmitted light to a spectrometer and a CCD camera for acquiring white light transmission spectra. Secondly, the sample contains not only a plane metal-coated cover slip as for MIET imaging, but also a metal-coated glass lens. The thicknesses of these silver layers were 30 nm for the bottom mirror and 60 nm for the top mirror, both were prepared by vapor deposition as described in the MIET section. For details of the used instruments and optical components, see [2].

We studied two different commercially available fluorophores, the hydrophobic dye Atto 647N and the hydrophilic dye Atto 655 from Atto-Tec (Siegen, Germany). In the first part of the study, lipid conjugates were examined, namely Atto 647N-1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (Atto 647N-DOPE) and Atto 655-1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine (Atto 655-DPPE). The fluorescent lipid analogues were incorporated into a supported lipid bilayer (SLB), a commonly used model of cell membranes [146–148]. The SLBs were formed by spin-coating a solution of 1 mg/ml of the lipid 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) in a 2:1 mixture of chloroform and methanol onto the substrate at 3000 rpm for 45 s. The dye-lipid conjugates were incorporated into the SLBs by adding them to the DOPC solution at ratios between 1:20 000 and 1:50 000 (w/w) prior to spin-coating. Subsequently, the lipid bilayers were hydrated and washed with a buffer containing 150 mM NaCl and 10 mM HEPES. After removal of excess buffer, the silver-coated lens was placed on top of the buffer and the SLB. Since the SLB was assumed to be very close to the substrate, a 30 nm thick silicon dioxide spacer was deposited above the bottom silver mirror to decrease direct
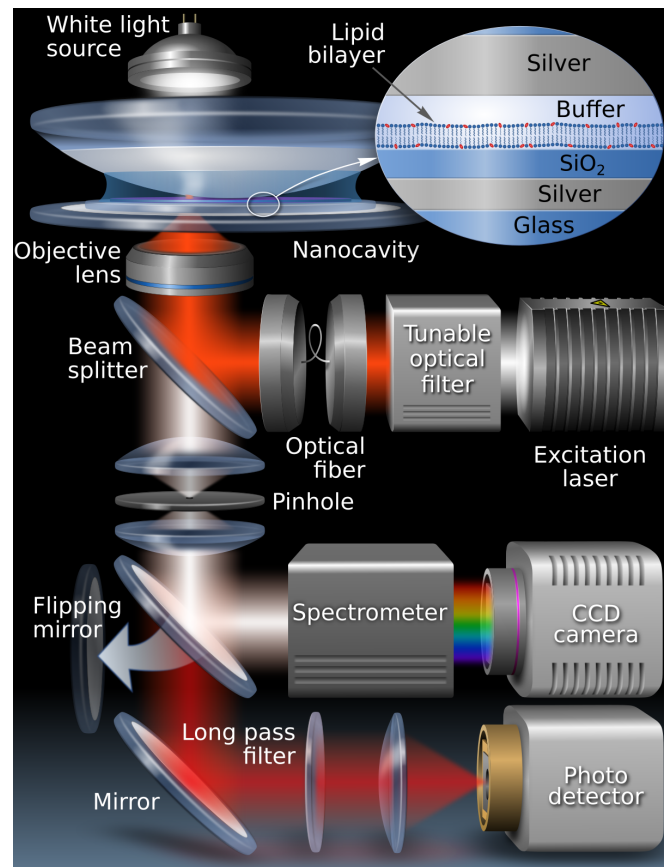
**Figure 4.39:** Schematic of the setup used to measure fluorescence lifetimes and white light transmission spectra inside a nanocavity. The inset shows the structure of the sample inisde the cavity, where small red dots represent the fluorophores. Image created by Alexey Chizhik and published in [2].
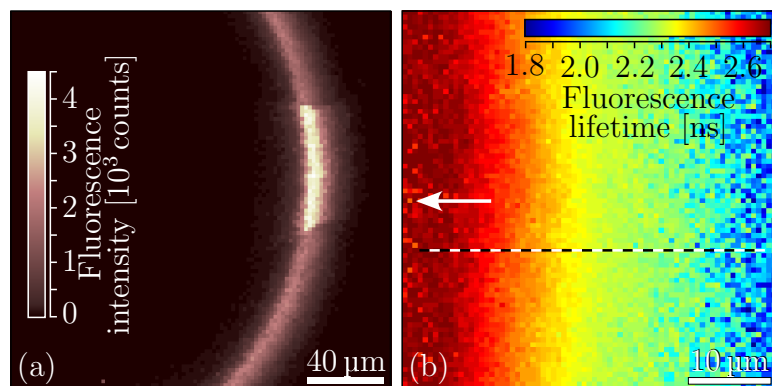


**Figure 4.40:** Confocal scan images of Atto 655-DPPE incorporated in a supported lipid bilayer and placed inside a nanocavity. (a) Fluorescence intensity in the $\lambda/2$ region, visible as a bright ring. The segment with approximately doubled fluorescence intensity corresponds to a region with an accidentally formed double bilayer. (b) Fluorescence lifetime distribution within a part of the $\lambda/2$ region where there is only a single bilayer. The dashed line represents a typical scan used for the quantum yield measurement, while the arrow points towards the center of the cavity where the lens touches the cover slip. Data recorded by Falk Schneider and Alexey Chizhik, figure published in [2].

quenching of the dye by the metal, see also the inset in figure 4.39. In a confocal fluorescence intensity image of such a sample, distinct bright rings are visible. These correspond to regions where the emission maximum $\lambda_{\max}$ of the dye is in resonance with the cavity. Our measurements were always performed in the innermost ring, called the $\lambda/2$ region, since a smaller cavity height ensures a stronger modification of the fluorescence lifetime. Besides guiding the choice of measurement positions, the lifetime images also acted as controls of the integrity of the lipid bilayers, since a disruption of the SLB would lead to loss of fluorescence. An interesting example (which was not used for the quantum yield measurement) of Atto 655-DPPE inside the SLB is shown in figure 4.40(a). There, coincidentally a small region with two SLBs on top of each other formed, resulting in a higher fluorescence intensity. This observation confirmed our trust in the intensity images as monitors of the SLB integrity. Figure 4.40(b) depicts a FLIM image of part of the $\lambda/2$ region, illustrating the lifetime change with cavity height. For determining the quantum yield, both the fluorescence lifetime and the white light transmission spectra were measured along the dashed line. When fitting the lifetime values to the theoretical model, we took into account the unusual sample geometry: Underneath the SLB, there is a thin water film with a height of one to two nanometers [149]. The lipid bilayer itself was assumed to be approximately 6 nm thick [108], with dye molecules placed within $\pm 2$ nm of either surface of the bilayer (due to the carbon chain linking the fluorophore to the lipid anchor).

In the second part of the study, the pure dyes (or, more accurately, the maleimide derivative of Atto 647N and the N-hydroxysuccinimide ester of Atto 655) were separately dissolved in water at a submicromolar concentration. A droplet of the solution was placed between the cavity mirrors, and both average photon arrival times and white light transmission spectra were obtained at several positions as described above. Since the fluorophores were able to explore the full cavity in this scenario, no spacer was used. Additionally, free space fluorescence lifetimes $\tau_{0,\mathrm{meas}}$ of all four samples were obtained on pure glass cover slips in order to compare them with the fit results for $\tau_0$.

### Results

The measured fluorescence lifetimes are plotted against the cavity height in figure 4.41. Filled circles correspond to measurements of the lipid conjugates incorporated into the SLB, while open circles denote values obtained in water. The best fit of our model to the data is shown as solid and dashed lines, respectively. The fit results are ($k_{\mathrm{r}} = \Phi/\tau_0$, $k_{\mathrm{nr}} = (1 - \Phi)/\tau_0$):

| sample | $\Phi$ [-] | $\tau_0$ [ns] | $k_{\mathrm{r}}$ [ns$^{-1}$] | $k_{\mathrm{nr}}$ [ns$^{-1}$] | $\tau_{0,\mathrm{meas}}$ [ns] |
|---|---|---|---|---|---|
| Atto 647N, water | 0.62 | 3.6 | 0.17 | 0.11 | 3.6 |
| Atto 647N, SLB | 0.77 | 3.9 | 0.20 | 0.06 | 4.0 |
| Atto 655, water | 0.31 | 1.9 | 0.16 | 0.36 | 1.9 |
| Atto 655, SLB | 0.46 | 3.0 | 0.16 | 0.19 | 2.9 |

In all four cases, the fit results of $\tau_0$ differ by 0.1 ns or less from the measured values $\tau_{0,\mathrm{meas}}$, which indicates a high fit quality. For both dyes, the quantum yield inside the SLB is larger than in water. At the same time, the free space lifetime of Atto 655
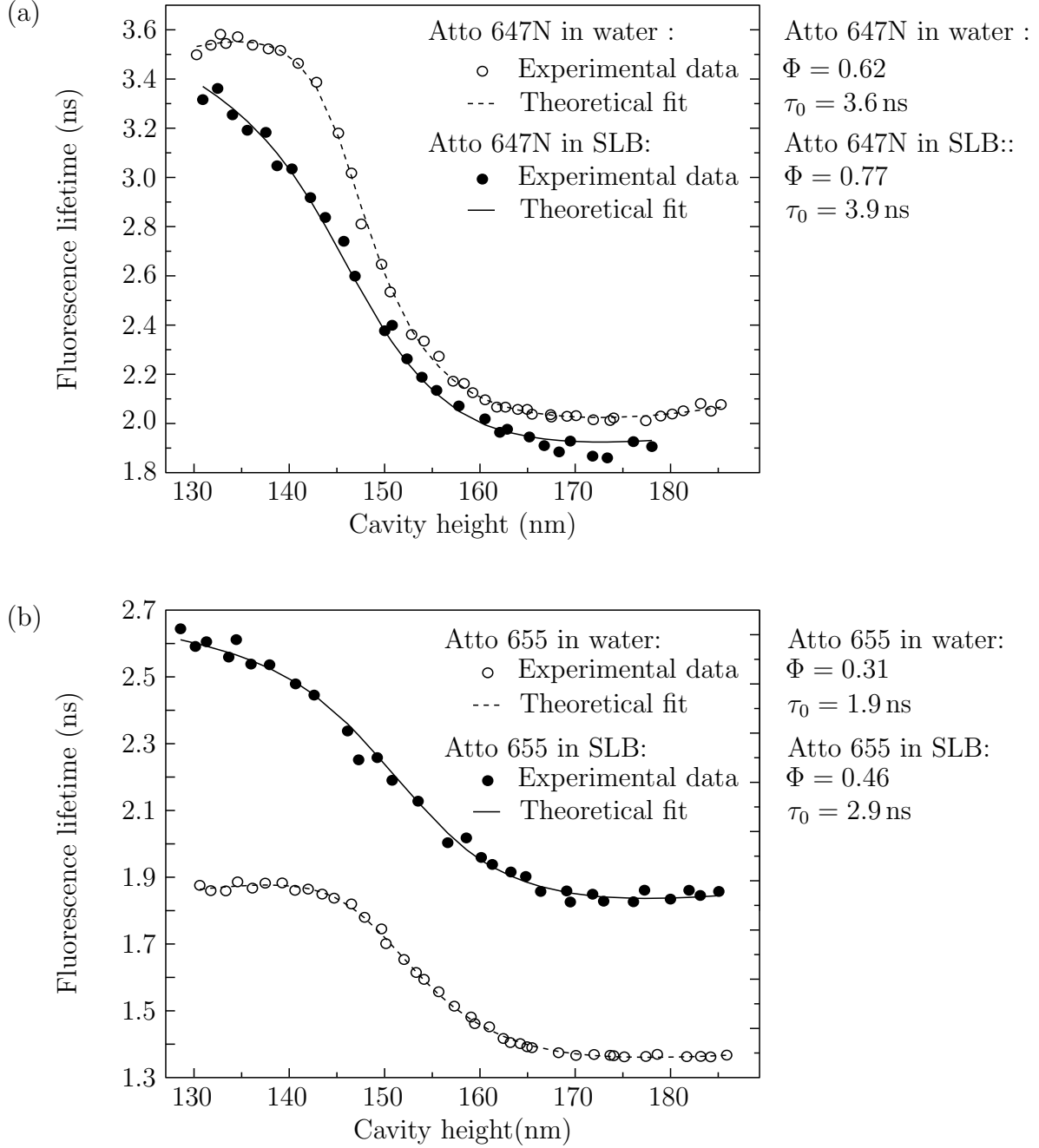
**Figure 4.41:** Average photon arrival time of Atto 647N (a) and Atto 655 (b) as a function of the cavity height. Solid circles denote measurements of the lipid analogues in an SLB, while open circles correspond to the pure dye in water. Solid and dashed curves are the best fit of the theoretical model to the data, where the free parameters are the quantum yield $\Phi$, the free space lifetime $\tau_0$ and the rotational diffusion time $\tau_{\text{rot}}$ (not shown). Data recorded by Falk Schneider and Alexey Chizhik, image published in [2].

increases by approximately 50 %, while the lifetime of Atto 647N is only weakly changed ($\sim 8\,\%$).

In order to explain these results, the environment of the fluorophores has to be considered: They might enter the lipid bilayer, or remain in the water phase (which is possible due to the carbon chain linking the dye to its lipid anchor). Assuming that the free space radiative rate only depends on the refractive index of the medium in which the fluorophore is embedded, $k_r$ should be constant if the dye remains in the water phase. Contrarily, since the refractive index of the lipid membrane is much higher than that of water, $n_{\mathrm{lipid}} \approx 1.46$ [124] as opposed to $n_{\mathrm{water}} = 1.33$, models such as the empty cavity model or the Strickler-Berg-equation (see section 4.1.4) predict that the radiative rate of the dye should increase if the dye enters the lipid bilayer. For Atto 647N, based on $k_r = 0.17\,\mathrm{ns}^{-1}$ in water, both models predict a radiative rate of $0.20\,\mathrm{ns}^{-1}$ in an environment with refractive index 1.46, in excellent agreeemnt with our measurements. In contrast, the radiative rate of Atto 655 remains constant, indicating that the dye does not enter the lipid membrane. This is in accordance with results from Hughes *et al.* [150], who found that pure Atto 647N interacts strongly with lipid bilayers, while Atto 655 only interacts very weakly with them.

The non-radiative rates of both dyes decrease by a factor of two when incorporated into the lipid bilayer. Environmental effects on non-radiative decay rates are a very complex topic: While detailed descriptions of decay mechanism have been given for some fluorophores, there is no general theory covering all emitters because a number of very different processes can lead to a non-radiative decay of the excited state (see e.g. [57], chapter 6 for an overview). Therefore, we do not attempt to explain the exact nature of the change of $k_{\mathrm{nr}}$ here. Instead, we only note that it is plausible that the very different environment inside or close to a lipid membrane leads to an altered non-radiative decay rate.

### 4.3.4 Accuracy of the quantum yield measurements

Several factors determine the accuracy of the quantum yield estimation: the quality of the raw data, the chosen fitting procedure, and the shape of the error landscape. The first aspect is quite intuitive, the less noisy the data, the more reliable the fit results. Contrarily, the last two points are non-trivial. We fit the measured average photon arrival times to the theoretical model by minimizing the sum of squared residuals (SSR) as a function of $\Phi$, $\tau_0$ and $\tau_{\mathrm{rot}}$. Finding the global minimum is a difficult non-linear three-dimensional optimization problem, see e.g. [151] for an overview of global optimization algorithms. The Nelder-Mead downhill simplex method [145] used in the present work to minimize the SSR does not guarantee that the obtained minimum is a global minimum, however, it is generally faster and easier to implement than true global minimization algorithms [151]. Yet, even if a true global minimization procedure would be used, the fit results could still be wrong if the position of the global minimum of the SSR does not coincide with the true values of $\Phi$, $\tau_0$ and $\tau_{\mathrm{rot}}$. Figuratively speaking, this situation can occur, for example, if the error landscape is relatively flat in a region around the true parameters, such that noise in the data can shift the global minimum away from the true values. In this case, restricting the allowed parameter space (e.g. by defining bounds for $\tau_0$ or $\tau_{\mathrm{rot}}$ based on independent measurements) can help to find a
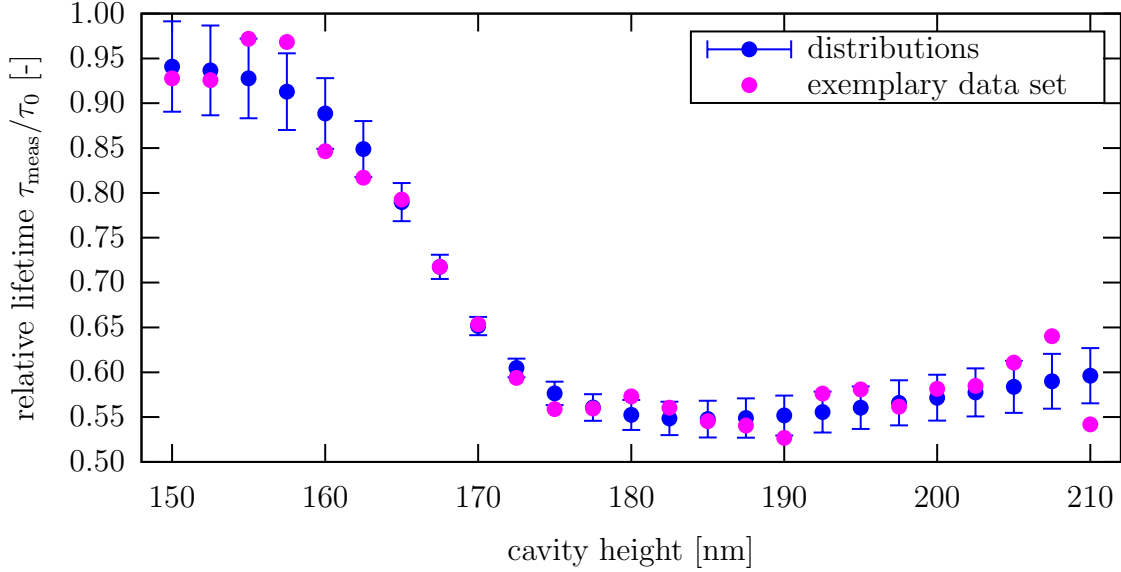
**Figure 4.42:** Data sets used for the simulation of lifetime data acquired in a nanocavity with known ground truth $\Phi = 0.5$, $\tau_0 = 1$ a.u. and $\tau_{\text{rot}} = \tau_0$ (other values tested in the simulations: $\tau_0/\tau_{\text{rot}} = 10^{\pm 4}$). The blue points with errorbars denote the mean and standard deviation of Gaussian distributions, from which single lifetime values are drawn. One exemplary data set is shown as pink dots; typically, 100 such data sets were evaluated per parameter set.

local minimum which is close to the true parameters. Conversely, it is useful to know which conclusions regarding the accuracy of the fitted quantum yield can be drawn from a comparison of the fit results of $\tau_0$ or $\tau_{\text{rot}}$ with measured values.

In order to test the robustness of the quantum yield fits, and to see how the uncertainty $\Delta\Phi$ changes when the other parameters are restricted, we performed fits of simulated data for which the ground truth was known. The results are presented in this section.

The simulated system was similar to the experiment done on the membrane staining dye dSQ12S: a nanocavity with silver mirrors (thicknesses 30 nm and 60 nm) and filled with DMSO ($n = 1.48$), an objective with numerical aperture 1.49, an excitation wavelength of 585 nm, a quantum yield of the dye of 0.5 and the same emission spectrum as we measured for dSQ12S (figure 6.5 in the appendix). We assumed different $R$-values ($R = \tau_0/\tau_{\text{rot}}$) and calculated the expectation values of the measured average photon arrival times $\langle\tau_{\text{meas}}\rangle(h)$ for cavity heights $h$ within the $\lambda/2$ region of the cavity. Noisy data was simulated by drawing random values from Gaussian distributions with means $\langle\tau_{\text{meas}}\rangle(h)$ and standard deviations $\langle\tau_{\text{meas}}\rangle(h) \cdot \sigma_{\text{rel}}$ using the relative uncertainties $\sigma_{\text{rel}} = \Delta\tau_{\text{meas}}(h)/\tau_{\text{meas}}(h)$ from the dSQ12S data set. In figure 4.42, the Gaussian distributions for $R = 1$ are visualized as blue points with errorbars, together with an exemplary data set (pink dots).

In a first step, 100 such data sets were evaluated using the fitting procedure described previously, with completely free fit parameters $\Phi$, $\tau_0$ and $\tau_{\text{rot}}$. Figure 4.43(a-c) show scatter plots of the obtained parameter pairs $(\tau_0, \Phi)$, $(\tau_{\text{rot}}, \Phi)$ and $(\tau_0, \tau_{\text{rot}})$, respectively. The ground truth ($\Phi = 0.5$, $\tau_0 = 1$ a.u. and $\tau_{\text{rot}}/\tau_0 = 10^{-4}$, 1 or $10^4$) is indicated as vertical or horizontal dashed lines. It becomes apparent that the fitted quantum yield
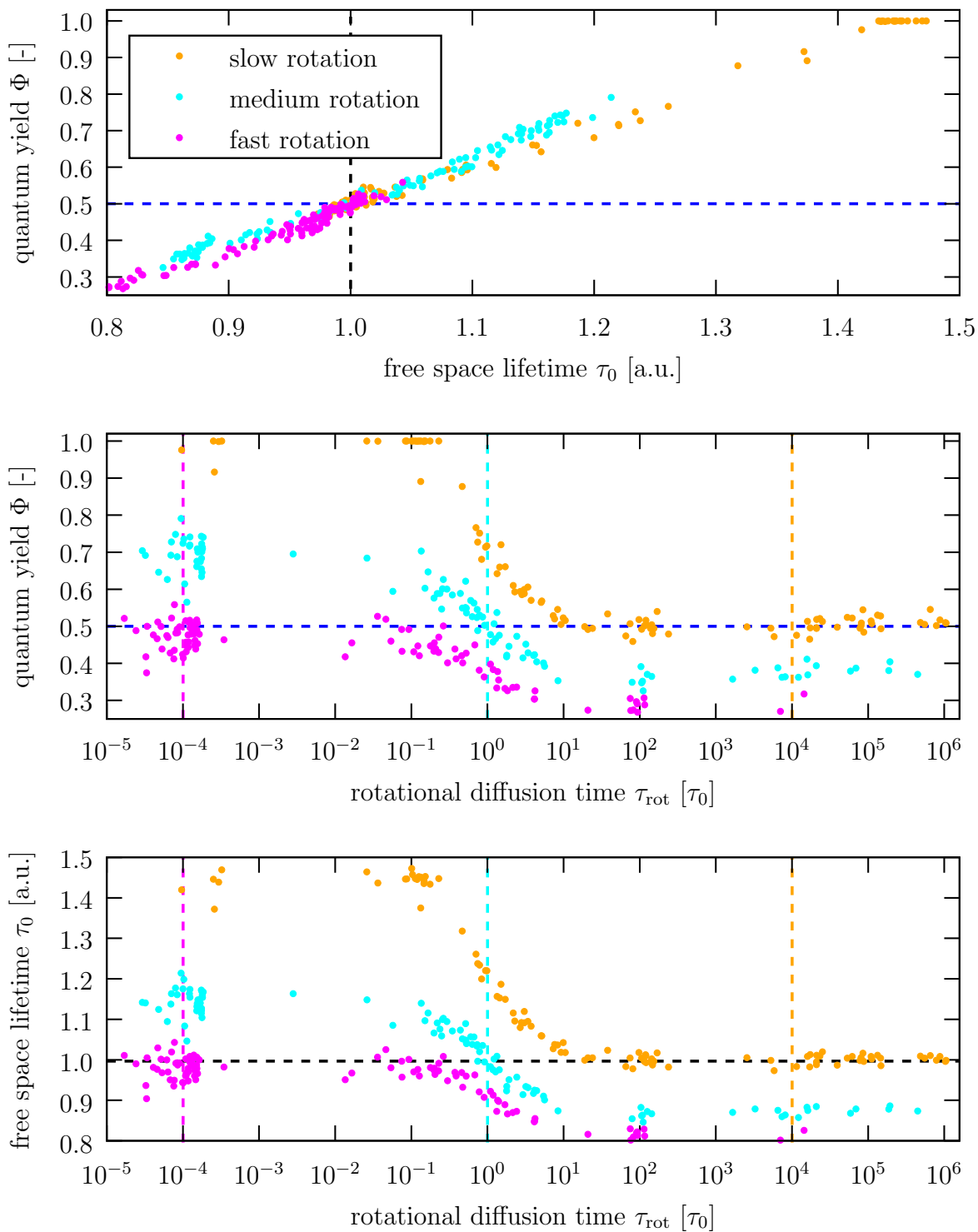
**Figure 4.43:** Fit parameters obtained by evaluating 100 data sets such as the one shown in figure 4.42. The only restrictions placed on the fit parameters were positivity for $\tau_0$ and $\tau_{rot}$ and $0 \leq \Phi \leq 1$. The simulations were carried out for three ratios $R = \tau_0/\tau_{rot}$, namely $R = 10^{-4}$ ('slow rotation', orange), $R = 1$ ('medium rotation', cyan) and $R = 10^4$ ('fast rotation', pink). Each point corresponds to one data set. Vertical and horizontal lines indicate the underlying true parameter values. Note that correct values of $\tau_0$ or $\tau_{rot}$ coincide with correct values of $\Phi$. The legend is valid for all three plots.
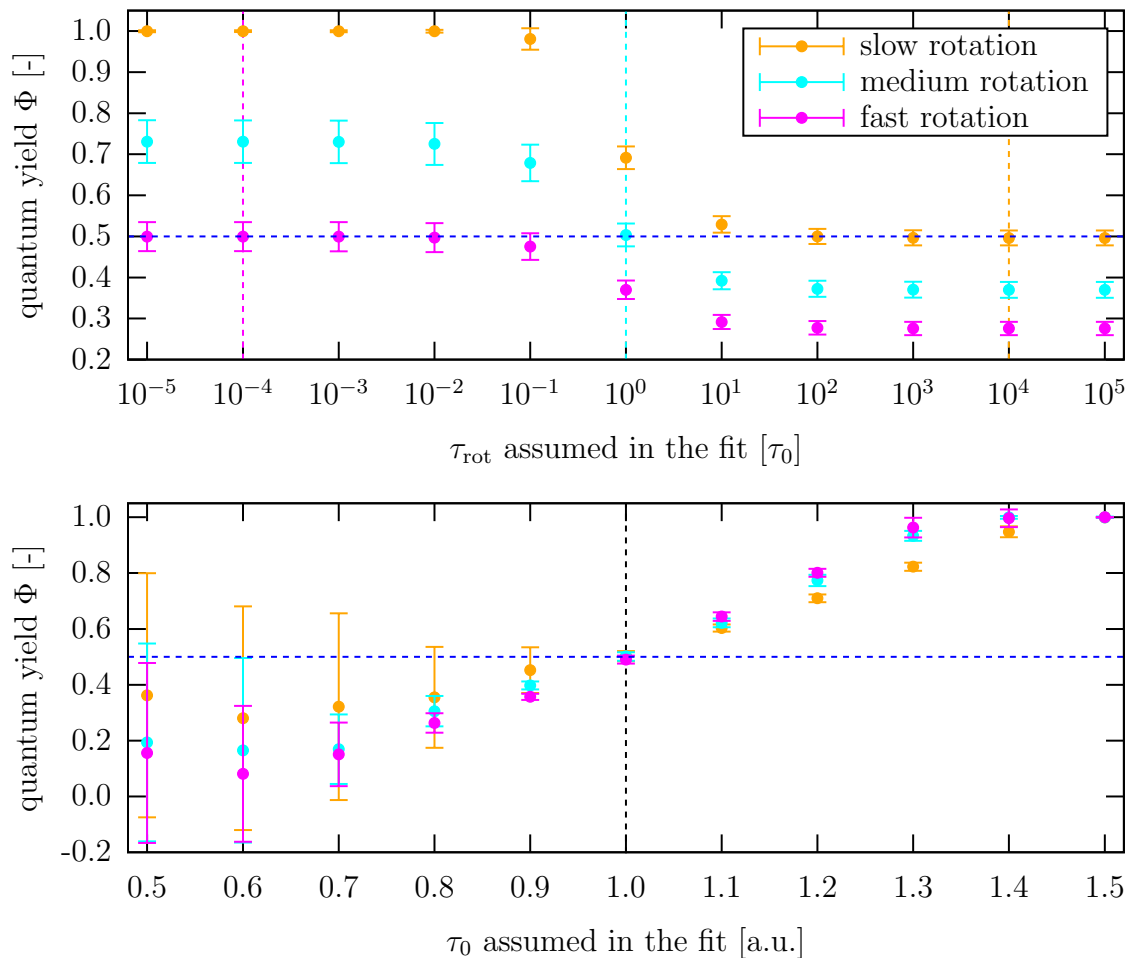
**Figure 4.44:** Fit parameters obtained by evaluating 100 data sets such as the one shown in figure 4.42. Either $\tau_{\mathrm{rot}}$ (top) or $\tau_0$ (bottom) were fixed to the values indicated on the axis. The simulations were carried out for three ratios $R = \tau_0/\tau_{\mathrm{rot}}$, namely $R = 10^{-4}$ ('slow rotation', orange), $R = 1$ ('medium rotation', cyan) and $R = 10^4$ ('fast rotation', pink). Vertical and horizontal lines indicate the underlying true parameter values. The legend is valid for both plots.

value is only correct if the fitted free space lifetime and rotational diffusion time are also close to their true values, with one exception: For slowly rotating molecules (true $\tau_{\mathrm{rot}} = 10^4\,\tau_0$), all fitted rotational diffusion times larger than $10^2\,\tau_0$ coincide with correct quantum yield fits. As a consequence, it seems advantageous to conduct an independent measurement of $\tau_0$ or $\tau_{\mathrm{rot}}$. Then, the measured values can either be used to retroactively check the fit quality, or to directly restrict the parameter space used for fitting.

Therefore, the second step of the study examined the performance of this approach. We hypothesized that setting $\tau_0$ or $\tau_{\mathrm{rot}}$ to their correct value would improve both the accuracy and the precision of the estimate of the QY, while fixing them at a wrong value would also lead to an incorrect $\Phi$. The top panel of figure 4.44 shows the mean and standard deviation of $\Phi$ obtained by fitting 100 data sets with fixed $\tau_{\mathrm{rot}}$. The correct value of $\Phi = 0.5$ is recovered for $\tau_{\mathrm{rot,\,guess}}/\tau_0 \leq 10^{-2}$ if the true ratio was $10^{-4}$,

and for $\tau_{\text{rot, guess}}/\tau_0 \geq 10^2$ if the true ratio was $10^4$. Thus, if the rotation is by at least two orders of magnitude smaller or larger than the fluorescence lifetime, the exact value is unimportant, and it suffices to know that the rotation is 'slow' or 'fast'. Contrarily, if the rotational diffusion time is similar to $\tau_0$, then the fit result depends sensitively on the assumed value of $\tau_{\text{rot, guess}}$. Depending on the true rotation behaviour of the fluorophore, the quantum yield values obtained for different assumed rotational diffusion times differ by as little as 0.25 (fast rotation) or as much as 0.5 (slow rotation). Therefore, the rotational diffusion time should only be fixed to a specific value – or a small range of values – if it is known with reasonable certainty.

A similar conclusion can be drawn from the fit results for fixed $\tau_0$ in the bottom panel of figure 4.44. Close to the correct value of $\tau_0$, the quantum yield changes approximately linearly with the assumed value of $\tau_0$, with the slope depending on the rotational behaviour of the emitter. In these cases, if the lifetime has been determined with an error of $\pm 10\%$, the resulting error of the quantum yield is between 0.08 (slow rotation) and 0.14 (fast rotation). We performed the same simulations for a true quantum yield of $\Phi = 0.8$ (see figure 6.6 in the appendix), and there found errors between 0.02 (slow rotation) and 0.18 (fast rotation). To put this result into perspective, consider the determination of the quantum yields of Atto 655 and Atto 647N in water and in supported lipid bilayers described in the previous section. There, the fit itself was performed with unconstrained parameters, yielding values for $\Phi$, $\tau_0$ and $\tau_{\text{rot}}$. Independent measurements of $\tau_0$ for the four different samples deviated by $3\%$ or less from the fit results, indicating that the quantum yield results are most likely also within $\pm 0.05$ of the true values.

To conclude, an unconstrained fit of lifetime data acquired in a nanocavity can, in some cases[22], lead to completely wrong results for the quantum yield, see figure 4.43. However, these wrong results are accompanied by wrong values of the free space lifetime and the rotational diffusion time, and can therefore be identified by comparing with independent measurements of $\tau_0$ or $\tau_{\text{rot}}$. If the latter are close to the fit results, then the quantum yield value can be assumed to be quite accurate, too, typically within $5\%$-$10\%$ of the correct value. Alternatively, a restriction of the allowed fit parameters of $\tau_0$ or $\tau_{\text{rot}}$ to independently measured ranges results in the same accuracy of the quantum yield values.

---

[22]Since the number of data points is not so large, this depends mostly on the random variations of the measured lifetimes. For example, if, by chance, all lifetimes at small cavity heights are higher than their expectation value, this already drastically changes the shape of the curve, shifting the global minimum of the SSR and thus leading to wrong fit results.

# 5 Discussion and outlook

In recent years, significant progress has been made in the field of optical microscopy. This process was driven by experimental advances (such as the development of fluorophores with tuned properties or detectors with higher sensitivity), new theoretical approaches, as well as improved data analysis. In this work, it was shown how a profound theoretical understanding of a seemingly exotic phenomenon, namely the influence of metal structures on fluorescence lifetimes, has lead to the development of a microscopy technique with outstanding $z$-resolution that is based on a completely different physical concept than other optical super-resolution methods.

The first part of this work was dedicated to classical electrodynamics. In particular, the radiation properties of oscillating electric dipoles – such as the angular distribution of radiation or the total energy flux – were derived in both planar and spherical geometries. Remarkably, these results can be used to explain a quantum mechanical phenomenon, namely the change of the fluorescence lifetime of single fluorophores in the vicinity of metal structures. The power of this semi-classical approach was demonstrated in the second part of this work, where metal-induced energy transfer (MIET) microscopy was employed to study several biological systems, and the fluorescence quantum yield of dyes in different environments was determined via lifetime measurements in a nanocavity.

As one application of MIET, a complex cellular adhesion process (the epithelial-to-mesenchymal transition in NMuMG cells) was studied with unprecedented axial resolution, revealing a reduction of focal adhesion sites and a lift-off of the cell membrane between these sites during the initial stages of the process. By carefully taking into account possible Förster resonance energy transfer (FRET) between different fluorophores, MIET was equipped with dual-colour imaging capabilities. This new technique allowed to monitor the maturation of focal adhesion complexes in human mesenchymal stem cells and to elucidate their three-dimensional structure, confirming the standard picture of a partial overlap of actin stress fibers and focal adhesion complexes (represented by vinculin), and showing that stress fibers emerge from focal adhesion complexes under a very shallow angle of below one degree. In a second dual-colour MIET experiment, three-dimensional profiles of the basal part of the inner and outer nuclear membrane of HeLa cells were obtained, providing the first study of the nuclear envelope structure along the vertical axis using optical microscopy. Even though these intracellular structures were localized at the limit of the MIET operating range, the obtained results of the average distance between inner and outer nuclear membrane agreed well with previous electron microscopy and horizontally acquired optical microscopy measurements, proving the applicability of MIET even in challenging samples.

For the evaluation of these various experiments, a versatile software package equipped with a user-friendly graphical interface was developed and constantly improved. It implements the non-trivial calculation of the height profile from given lifetime values for virtually arbitrary stratified samples. To encourage the application of MIET in the general biological community, the software is freely available. From personal communi-

cation, we know about a recent increase in the usage of the technique.

It can be expected that MIET will soon be applied to answer a broader range of biological questions when it becomes possible to conduct live-cell MIET imaging of intracellular structures[1]. Typically, fluorescence microscopy of living cells is performed with fluorescent proteins that are expressed by the cell itself. However, most fluorescent proteins exhibit a multiexponential fluorescent decay due to the presence of different conformations of the chromophore or the amino acid residues in its immediate surroundings (see e.g. [75, 152–154]). If these species have different quantum yields that are not known individually, calculating an accurate MIET calibration curve is not possible. Similar problems have been identified in the context of lifetime-based FRET measurements, which lead to the development of some truly monoexponential fluorescent proteins, e.g. the cyan variant mTurquoise2 [155] and the red variant mScarlet [156]. With the help of these new labels, live-cell MIET measurements are now a realistic option.

The range of possible MIET applications also depends on the achievable localization accuracy. While an axial resolution of 2.5 nm to 3 nm was shown in previous single-molecule and ensemble MIET experiments [38–40], those measurements were performed in well-defined environments. Thus, the resolution was determined solely by the accuracy of the lifetime values and the steepness of the MIET curve. In this work, the additional error introduced by incorrectly modeling the sample structure during the calculation of the MIET curve was studied. It was found that the refractive index of the sample itself has a relatively small impact, for example, replacing a correct value of $n = 1.37$ with an incorrect value of $n = 1.33$ or $n = 1.40$ only leads to a bias of less than 7 nm in the calculated axial position of the fluorophore in a typical sample. The refractive index of dielectric structures which are not in direct contact with the fluorophore, such as the plasma membrane in experiments targeting intracellular structures, influences the MIET curve even less. In general, the error due to simplified or incorrectly guessed refractive indices is $z$-dependent, with smaller errors at small $z$. The same trend was found when assuming an incorrect quantum yield in the calculation of the MIET curve. These results are encouraging in two respects. Firstly, while of course an effort should be made to determine the refractive index structure of the sample as accurately as possible, $z$-positions obtained by MIET are quite accurate even in complex cellular environments that have to be simplified for the calculations. Secondly, a future application of MIET in structural biology (by labeling and localizing two distinct sites within a biological macromolecule in order to determine its three-dimensional structure) would mainly require to determine heights close to the surface, where the accuracy is highest.

The latter example leads us from the discussion of MIET experiments on densely labeled samples, which were the focus of this work, to potential applications of single-molecule MIET (smMIET). The basic physical principle of smMIET is the same as for MIET of densely labeled samples, however, the orientation dependence of the MIET curve becomes more important when observing single molecules. In dense samples, the fluorophores were either assumed to rotate quickly (e.g. in solution measurements) or to be oriented randomly, thus leading to an averaging of the different MIET curves when many fluorophores are present within the focal spot. The former case is also possible for single molecules, for example if they are attached to the structure of interest

---

[1] To the present date, MIET measurements of living cells always targeted the plasma membrane because it can be labeled by organic fluorophores without a lethal permeabilization of the membrane.

via a flexible linker. Contrarily, if the molecules are fixed, their orientation has to be determined in order to choose the correct MIET curve. Some pioneering work in this direction was done in [40], where the intensity patterns observed in defocused images of single fluorophores were used to determine the orientation of the dipole moment. More recently [9], an extensive study containing theoretical considerations, simulations and experimental data treated the three-dimensional localization of fixed dipoles by performing smMIET-scans with a radially polarized excitation laser or by alternatively combining smMIET with defocused imaging. It was demonstrated that lateral and horizontal localization accuracies of 2 nm to 4 nm and 0.5 nm to 1.5 nm, respectively, can be achieved under realistic experimental conditions. In the same work, a similar theoretical resolution was derived for rapidly rotating dipoles studied with a simple confocal FLIM setup. The latter situation was recently realized in [8], where rapidly rotating fluorophores were used to label rigid DNA origami pillar structures with well-defined geometry and orientation. By imaging single pillars labeled with two dyes, two bleaching steps could be observed in the intensity time-traces, allowing to fit two lifetime components to the corresponding TCSPC curves and thus to colocalize the two dyes along the optical axis. There, axial localization accuracies ranged from 3 nm to 5 nm. These examples demonstrate the potential of smMIET for the field of structural biology. Potentially, this can be improved even further by screening for dyes that exhibit both photoswitching suitable for approaches such as PALM or dSTORM, and stable, environment-independent fluorescence lifetimes.

The second important application of electrodynamics simulations presented in the present work is the determination of fluorescence quantum yield (QY) values using a metallic nanocavity. This technique has the great advantage of yielding absolute QY values without need of calibration or comparison with a known sample. Furthermore, it does not underestimate the QY of samples with a non-fluorescent population (such as photoactivatable fluorescent proteins) because it is solely based on the cavity-induced fluorescence lifetime modulation of the fluorescent population. Finally, the method allows to determine the QY of fluorophores in a large variety of environments, from solution measurements to molecules embedded in a solid, thin layers of dyes deposited on a surface, or even single molecules [157]. This versatility was demonstrated in the present work by measuring and comparing the QY of two dyes both in aqueous solution and embedded in a supported lipid bilayer. The results strongly suggest that both the empty-cavity model and the Strickler-Berg equation correctly predict the refractive index dependence of the radiative deexcitation rate of a dipole emitter. This finding allows to deduce the QY of a fluorophore in any medium based on a lifetime measurement in said medium, and the knowledge of the QY and lifetime of the emitter in a second medium. Such calculations are especially important for MIET in cells, where it is hard to measure the QY but straightforward to determine the lifetime.

To sum up, the present work has treated an unusual application of classical electrodynamics by modeling the fluorescence lifetime changes of dipole emitters in the vicinity of metal nanostructures. Two applications of this theory were demonstrated: MIET microscopy, which provides axial super-resolution in fluorescence microscopy, and lifetime-based quantum yield measurements in a metal nanocavity. Based on our own results and the recent developments outlined in this discussion, exciting advances in lifetime-based localization microscopy can be expected in the near future.

# 6 Appendix

## 6.1 Proofs, additional calculations and code snippets

### 6.1.1 Orthogonality of vector spherical harmonics

We want to show that

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \ \boldsymbol{M}_{\ell m}^f \cdot \overline{\boldsymbol{N}_{\ell' m'}^g} \sin\theta = 0 \ , \tag{6.1}$$

$$\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \ \boldsymbol{M}_{\ell m}^f \cdot \overline{\boldsymbol{M}_{\ell' m'}^g} \sin\theta \propto \delta_{l,l'}\delta_{m,m'} \ , \tag{6.2}$$

$$\text{and} \ \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \ \boldsymbol{N}_{\ell m}^f \cdot \overline{\boldsymbol{N}_{\ell' m'}^g} \sin\theta \propto \delta_{l,l'}\delta_{m,m'} \ . \tag{6.3}$$

It is easy to see from the orthogonality of the complex exponential functions $e^{im\phi}$ (equation 2.24) that all three integrals are zero for $m \neq m'$. The $\ell$-dependency is a little harder to prove, the following is based heavily on [44]. We will start with equation (6.1), it requires us to determine the following integral:

$$\int_0^{\pi} d\theta \sin\theta \left( i\pi_{\ell m}(\theta)\tau_{\ell' m}(\theta) + i\pi_{\ell' m}\tau_{\ell m} \right) = imc_{\ell m}c_{\ell' m} \int_0^{\pi} \left( P_{\ell}^m \frac{\partial P_{\ell'}^m}{\partial \theta} + P_{\ell'}^m \frac{\partial P_{\ell}^m}{\partial \theta} \right) d\theta$$

$$= imc_{\ell m}c_{\ell' m} \left[ P_{\ell}^m P_{\ell'}^m \right]_0^{\pi} \ . \tag{6.4}$$

This is always zero because $P_{\ell}^m(\cos\theta) \propto \sin^m(\theta)$. Next, we consider equation (6.2), which is proportional to

$$\frac{1}{c_{\ell m}c_{\ell' m}} \int_0^{\pi} d\theta \sin\theta \left\{ \pi_{\ell m}\pi_{\ell' m} + \tau_{\ell m}\tau_{\ell' m} \right\}$$

$$= \int_0^{\pi} d\theta \sin\theta \left\{ m^2 \frac{P_{\ell}^m(\cos\theta) \cdot P_{\ell'}^m(\cos\theta)}{\sin^2\theta} + \frac{\partial P_{\ell}^m}{\partial \theta} \cdot \frac{\partial P_{\ell'}^m}{\partial \theta} \right\}$$

$$= \frac{1}{2} \int_0^{\pi} d\theta \left[ \frac{m^2}{\sin^2\theta} P_{\ell}^m - \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P_{\ell}^m}{\partial \theta} \right) \right] P_{\ell'}^m \sin\theta + \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P_{\ell}^m}{\partial \theta} \right) P_{\ell'}^m$$

$$+ \left[ \frac{m^2}{\sin^2\theta} P_{\ell'}^m - \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P_{\ell'}^m}{\partial \theta} \right) \right] P_{\ell}^m \sin\theta + \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P_{\ell'}^m}{\partial \theta} \right) P_{\ell}^m$$

$$+ 2 \frac{\partial P_{\ell}^m}{\partial \theta} \cdot \frac{\partial P_{\ell'}^m}{\partial \theta} \sin\theta$$

$$= \frac{1}{2} \int_0^{\pi} d\theta \ \ell(\ell+1)P_{\ell}^m P_{\ell'}^m + \ell'(\ell'+1)P_{\ell}^m P_{\ell'}^m + \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial P_{\ell}^m}{\partial \theta} P_{\ell'}^m + \sin\theta \frac{\partial P_{\ell'}^m}{\partial \theta} P_{\ell}^m \right)$$

$$= \frac{2\ell(\ell+1)}{(2\ell+1)} \frac{(\ell+m)!}{(\ell-m)!} \delta_{\ell,\ell'} + \left[ \sin\theta P_{\ell}^{m+1} P_{\ell'}^m - m\cos\theta P_{\ell}^m P_{\ell'}^m \right]_0^{\pi}$$

$$+ \left[ \sin\theta P_{\ell'}^{m+1} P_{\ell}^{m} - m\cos\theta P_{\ell'}^{m} P_{\ell}^{m} \right]_{0}^{\pi}$$
$$= \frac{2\ell(\ell+1)}{(2\ell+1)} \frac{(\ell+m)!}{(\ell-m)!} \, \delta_{\ell,\ell'} \; , \tag{6.5}$$

where the second step uses the fact that $P_{\ell}^{m}$ solves equation (2.25) and the third step employs the orthogonality of the $P_{\ell}^{m}$. In the fourth step we used again that $P_{\ell}^{m}(\cos\theta)$ is zero at $\theta = \{0, \pi\}$ for $m \neq 0$ and finite for $m = 0$. Thus,

$$\int_{0}^{2\pi} \mathrm{d}\phi \int_{0}^{\pi} \mathrm{d}\theta \; \boldsymbol{M}_{\ell m}^{f} \cdot \overline{\boldsymbol{M}_{\ell' m'}^{g}} \; \sin\theta = \delta_{l,l'} \delta_{m,m'} \; \ell(\ell+1) \; f_{\ell}(kr) \overline{g}_{\ell}(kr) \; . \tag{6.6}$$

The proof of equation (6.3) simply requires us to combine some previous results:

$$\int_{0}^{2\pi} \mathrm{d}\phi \int_{0}^{\pi} \mathrm{d}\theta \; \boldsymbol{N}_{\ell m}^{f} \cdot \overline{\boldsymbol{N}_{\ell' m'}^{g}} \sin\theta$$
$$= \delta_{m,m'} 2\pi \frac{c_{\ell m}^{2}}{(kr)^{2}} \left\{ \ell(\ell+1)\ell'(\ell'+1) f_{\ell}\overline{g_{\ell'}} \int_{0}^{\pi} P_{\ell}^{m} P_{\ell'}^{m} \sin\theta \mathrm{d}\theta \right.$$
$$\left. + \frac{\partial}{\partial r}(r f_{\ell}) \frac{\partial}{\partial r} \overline{(r g_{\ell})} \int_{0}^{\pi} \left( m^{2} \frac{P_{\ell}^{m}(\cos\theta) \cdot P_{\ell'}^{m}(\cos\theta)}{\sin^{2}\theta} + \frac{\partial P_{\ell}^{m}}{\partial\theta} \cdot \frac{\partial P_{\ell'}^{m}}{\partial\theta} \right) \sin\theta \mathrm{d}\theta \right\}$$
$$= \delta_{m,m'} \delta_{\ell,\ell'} \frac{1}{(kr)^{2}} \left\{ \ell^{2}(\ell+1)^{2} f_{\ell}\overline{g_{\ell}} + \ell(\ell+1) \frac{\partial}{\partial r}(r f_{\ell}) \frac{\partial}{\partial r} \overline{(r g_{\ell})} \right\} \tag{6.7}$$

## 6.1.2 Radius-independence of spherical flux integrals

We want to proof that $\mathfrak{Im}[f(\rho)] = -1/\rho^{2}$ for all $\rho \in \mathbb{R}$, as used in section 2.4.6. The function is defined as

$$f_{\ell}(\rho) := j_{\ell} \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho h_{\ell})}{\mathrm{d}\rho}} + h_{\ell} \overline{\frac{1}{\rho} \frac{\mathrm{d}(\rho j_{\ell})}{\mathrm{d}\rho}}, \tag{6.8}$$

with the spherical Bessel function $j_{\ell}$ and the spherical Hankel function $h_{\ell}$, and where a bar denotes complex conjugation. We make use of the fact that the spherical Hankel function can be expressed as a sum of $j_{\ell}$ and the spherical Neumann function $y_{\ell}$, $h_{\ell} = j_{\ell} + i y_{\ell}$, to find an expression of real functions:

$$\mathfrak{Im}[f_{\ell}(\rho)] = j_{\ell} \frac{1}{\rho} \frac{\mathrm{d}(-\rho y_{\ell})}{\mathrm{d}\rho} + y_{\ell} \frac{1}{\rho} \frac{\mathrm{d}(\rho j_{\ell})}{\mathrm{d}\rho}$$
$$= -j_{\ell} \left( \frac{y_{\ell}}{\rho} + \frac{\mathrm{d}y_{\ell}}{\mathrm{d}\rho} \right) + y_{\ell} \left( \frac{j_{\ell}}{\rho} + \frac{\mathrm{d}j_{\ell}}{\mathrm{d}\rho} \right)$$
$$= -j_{\ell} \frac{\mathrm{d}y_{\ell}}{\mathrm{d}\rho} + y_{\ell} \frac{\mathrm{d}j_{\ell}}{\mathrm{d}\rho}. \tag{6.9}$$

As the induction basis, we evaluated this expression numerically for $\ell = 1$ and $\ell = 2$, and found the result to be $-1/\rho^2$ within the numerical accuracy. For the induction step, we need the two recursion formulas ([46], §19):

$$Z_{\ell+1} + Z_{\ell-1} = \frac{2n}{\rho} Z_\ell$$

$$Z_{\ell+1} - Z_{\ell-1} = -2\frac{\mathrm{d}Z_\ell}{\mathrm{d}\rho}, \tag{6.10}$$

which can be combined to give

$$Z_{\ell-1} = \frac{n}{\rho}Z_\ell + \frac{\mathrm{d}Z_\ell}{\mathrm{d}\rho},$$

$$Z_{\ell+1} = \frac{n}{\rho}Z_\ell - \frac{\mathrm{d}Z_\ell}{\mathrm{d}\rho}, \tag{6.11}$$

where $Z$ can be either $j$ or $y$. This also has the useful implication that

$$j_\ell y_{\ell-1} - y_\ell j_{\ell-1} = \frac{n}{\rho}j_\ell y_\ell + j_\ell\frac{\mathrm{d}y_\ell}{\mathrm{d}\rho} - \frac{n}{\rho}y_\ell j_\ell - y_\ell\frac{\mathrm{d}j_\ell}{\mathrm{d}\rho} = -\mathfrak{Im}[f_\ell]. \tag{6.12}$$

With the induction hypothesis that

$$\mathfrak{Im}[f_\ell(\rho)] = -\frac{1}{\rho^2}, \tag{6.13}$$

we can then prove the inductive step:

$$\begin{aligned}
\mathfrak{Im}[f_{\ell+1}] =& y_{\ell+1}\frac{\mathrm{d}j_{\ell+1}}{\mathrm{d}\rho} - j_{\ell+1}\frac{\mathrm{d}y_{\ell+1}}{\mathrm{d}\rho}\\
=& \left(\frac{2\ell}{\rho}y_\ell - y_{\ell-1}\right)\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{2\ell}{\rho}j_\ell - j_{\ell-1}\right) - \left(\frac{2\ell}{\rho}j_\ell - j_{\ell-1}\right)\frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{2\ell}{\rho}y_\ell - y_{\ell-1}\right)\\
=& \left(\frac{2\ell}{\rho}\right)^2 \underbrace{\left(y_\ell\frac{\mathrm{d}j_\ell}{\mathrm{d}\rho} - j_\ell\frac{\mathrm{d}y_\ell}{\mathrm{d}\rho}\right)}_{\text{I.H.: }-1/\rho^2} - \frac{4\ell^2}{\rho^3}\underbrace{(y_\ell j_\ell - j_\ell y_\ell)}_{=0} - \frac{2\ell}{\rho}\left(y_\ell\frac{\mathrm{d}j_{\ell-1}}{\mathrm{d}\rho} - j_\ell\frac{\mathrm{d}y_{\ell-1}}{\mathrm{d}\rho}\right)\\
& - \frac{2\ell}{\rho}\left(y_{\ell-1}\frac{\mathrm{d}j_\ell}{\mathrm{d}\rho} - j_{\ell-1}\frac{\mathrm{d}y_\ell}{\mathrm{d}\rho}\right) + \frac{2\ell}{\rho^2}\underbrace{(y_{\ell-1}j_\ell - j_{\ell-1}y_\ell)}_{(6.12):\ 1/\rho^2} + \underbrace{\left(y_{\ell-1}\frac{\mathrm{d}j_{\ell-1}}{\mathrm{d}\rho} - j_{\ell-1}\frac{\mathrm{d}y_{\ell-1}}{\mathrm{d}\rho}\right)}_{\text{I.H.: }-1/\rho^2}\\
=& -\frac{1}{\rho^2}\left(\frac{4\ell^2}{\rho^2} - \frac{2\ell}{\rho^2} + 1\right) - \frac{2\ell}{\rho}\left[y_\ell\left(\frac{\ell-1}{\rho}j_{\ell-1} - j_\ell\right) - j_\ell\left(\frac{\ell-1}{\rho}y_{\ell-1} - y_\ell\right)\right]\\
& - \frac{2\ell}{\rho}\left[y_{\ell-1}\left(j_{\ell-1} - \frac{\ell}{\rho}j_\ell\right) - j_{\ell-1}\left(y_{\ell-1} - \frac{\ell}{\rho}y_\ell\right)\right]\\
=& -\frac{1}{\rho^2}\left(\frac{4\ell^2}{\rho^2} - \frac{2\ell}{\rho^2} + 1\right) - \frac{2\ell}{\rho^2}\left[(\ell-1)\underbrace{(y_\ell j_{\ell-1} - j_\ell y_{\ell-1})}_{(6.12):\ -1/\rho^2} + \ell\underbrace{(j_{\ell-1}y_\ell - y_{\ell-1}j_\ell)}_{(6.12):\ -1/\rho^2}\right]\\
=& -\frac{1}{\rho^2} \qquad \text{q.e.d.} \tag{6.14}
\end{aligned}$$

### 6.1.3 Excitation intensity in a nanocavity

Continuing the discussion from the section *Excitation probability* in 4.3.1, the electrical field inside a nanocavity that is illuminated with a confocal microscope is given by [107]:

$$
\begin{aligned}
E_x(\rho, \phi, z) &= f_0(\rho, z) + f_2(\rho, z) \cdot \cos(2\phi) \\
E_y(\rho, \phi, z) &= f_2(\rho, z) \cdot \sin(2\phi) \\
E_z(\rho, \phi, z) &= f_1(\rho, z),
\end{aligned}
\tag{6.15}
$$

where it was assumed that the electric field in the back focal plane is polarized along the $x$-axis, and that the point $\boldsymbol{r} = (\rho, \phi, z)$ in the nanocavity is given by cylindrical coordinates. The functions $f_{0,1,2}(\rho, z)$ are independent of the angle $\phi$, they are obtained by integrating over all plane wave components up to the maximum angle permitted by the numerical aperture of the objective, $\chi_{\max} = \arcsin(n/NA)$:

$$
f_j = \int_0^{\chi_{\max}} \mathrm{d}\chi \ \sin\chi\sqrt{\cos\chi}J_j(k_m\sin\chi_m\rho)\left[\kappa_j^+(\chi_m)\mathrm{e}^{izk_m\cos\chi_m} + \kappa_j^-(\chi_m)\mathrm{e}^{-izk_m\cos\chi_m}\right].
\tag{6.16}
$$

Here, $k_m = n_m k_v$ is the wave vector in the fluorophore's medium, $J_j$ is a Bessel function of the first kind, and the two summands in the square brackets describe upwards (+) and downwards (-) traveling waves:

$$
\begin{aligned}
\kappa_0^\pm(\chi_m) &= \quad\ T_s^\pm(\chi_m)\cos(\chi_m) + T_p^\pm(\chi_m) \\
\kappa_1^\pm(\chi_m) &= -2iT_s^\pm(\chi_m)\sin(\chi_m) \\
\kappa_2^\pm(\chi_m) &= -\ T_s^\pm(\chi_m)\cos(\chi_m) + T_p^\pm(\chi_m).
\end{aligned}
\tag{6.17}
$$

The coefficients $T_{p,s}^\pm$ are the effective transmission coefficients (from the objective to the cavity) for plane $p$- and $s$-waves, which already include the multiple reflections of the waves between the two cavity mirrors:

$$
\begin{aligned}
T_{p,s}^+ &= \frac{t_{p,s}^+}{1 - r_{p,s}^+ r_{p,s}^- \mathrm{e}^{2ihk_m\cos(\chi_m)}} \\
T_{p,s}^- &= \frac{t_{p,s}^+ r_{p,s}^+ \mathrm{e}^{2ihk_m\cos(\chi_m)}}{1 - r_{p,s}^+ r_{p,s}^- \mathrm{e}^{2ihk_m\cos(\chi_m)}}.
\end{aligned}
\tag{6.18}
$$

The Fresnel reflection and transmission coefficients $r_{p,s}^\pm$ and $t_{p,s}^+$ at the bottom or top of the cavity are calculated as described in section 3.2. Here, $t_{p,s}^+$ describes the transmission from the objective's immersion medium to the inside of the cavity (wave traveling in positive $z$-direction), $r_{p,s}^+$ describes the reflection at the top of the cavity (incident wave traveling in positive $z$-direction) and $r_{p,s}^-$ describes the reflection at the bottom of the cavity (incoming wave traveling in negative $z$-direction).

For the determination of quantum yield values from lifetime measurements in a nanocavity, the $\rho$- and $\phi$-coordinates are not important, since the lifetime only depends on $z$. Therefore, the intensity is averaged over $\rho$ and $\phi$, see also section 4.3.1.

## 6.1.4 Simulation of a rotating fluorophore inside a nanocavity

The following code is a Matlab-implementation of the algorithm introduced in section 4.3.1.

```matlab
function [R,tau] = DecayRotDiffusion(iniDistr)
% sample parameters
nruns       = 1000;                     % number of simulation runs
ncurves     = 5000;                     % number of "molecules" per run
load('metals.mat','wavelength','silver'); % dispersion of refr. index
lamem = 650;                            % emission wavelength [nm]
n0      = [1.52 silver(wavelength==lamem)]; % bottom mirror on glass
n       = 1.48;                         % refractive index of DMSO
n1      = [silver(wavelength==lamem) 1.52]; % top mirror on glass
d0      = 30;                           % thickness bottom mirror [nm]
d       = 200;                          % height of the cavity [nm]
d1      = 60;                           % thickness top mirror [nm]
QY      = 0.5;                          % quantum yield of the fluorophore
k0      = 2*pi/lamem;                   % wave vector in vacuum

% simulation parameters
steplengths = [0.01 0.01 0.01 0.01 0.1 0.1 0.1 0.1 0.3 0.3 0.3 0.3];
rates       = [1.5*10.^[-1 -2 -3 -4 -1 -2 -3 -4] 1.35*10.^[-1 -2 -3 -4]];

% MIET
[~,~,~,~,qvd,qvu,qpd,qpu] = LifetimeLSimpsExp(k0*d/2,n0,n,n1,k0*d0,k0*d,k0*d1);
S0 = n*4/3;     % total energy S_tot emitted per time in free space
Sv = qvd+qvu;   % S_tot for vertical dipole
Sp = qpd+qpu;   % S_tot for horizontal dipole (i.e. parallel to mirrors)

% results: mean and standard deviation over all "nruns" simulation runs
tauMean = zeros(1,numel(rates));
tauStd  = zeros(1,numel(rates));

for Dcounter=1:numel(rates)
    steplength = steplengths(Dcounter); % steplength of rotational diffusion
    rate       = rates(Dcounter);   % total decay rate in free space
    kr         = QY*rate;           % radiative rate in free space
    knr        = (1-QY)*rate;       % non-radiative rate in free space
    nt         = ceil(8/rate);      % maximum number of time steps for one "molecule"
    dt         = 1;                 % length of one time step [a.u.]
    tauV       = zeros(nruns,1);    % fitted lifetimes for all runs

    % start values for theta, depending on polarization of excitation light
    thetaV = (0:100)/100*pi;
    if iniDistr == 0          % random excitation
        prob   = sin(thetaV);
    elseif iniDistr == 1      % parallel excitation
        prob   = sin(thetaV).^3;
    else                      % vertical excitation
        prob   = cos(thetaV).^2.*sin(thetaV);
    end
    cumprob= cumsum(prob); cumprob=cumprob/cumprob(end);
    % interp1 cannot deal with non-unique x-values:
    % average thetas with same value of cumulative probability
    cumprobU=unique(cumprob); thetaVU=thetaV; index=false(size(thetaV));
    for counter=2:numel(cumprob)
        if cumprob(counter)==cumprob(counter-1)
            thetaVU(counter-1)=(thetaVU(counter)+thetaVU(counter-1))/2;
            index(counter)=true;
        end
    end
    thetaVU(index)=[];

    parfor rc = 1:nruns
        % variable for storing decay times
        time = NaN(ncurves,1);

        % draw random start positions
```

```
                 theta0 = interp1(cumprobU,thetaVU,rand(1,ncurves));
66               phi0   = 2*pi*rand(1,ncurves);
                 xstart = sin(theta0).*cos(phi0);
                 ystart = sin(theta0).*sin(phi0);
                 zstart = cos(theta0);

71           % "propagate"
             for k=1:ncurves
                 % position of the molecule, r will be updated with each step
                 r = [xstart(k); ystart(k); zstart(k)];
                 % random steps from the "north pole"
76               angles= 2*pi*rand(nt,1);
                 xstep = sin(steplength)*cos(angles);
                 ystep = sin(steplength)*sin(angles);
                 zstep = cos(steplength);
                 % random numbers for fluorescent decay
81               decayNumbers = rand(nt,1);
                 % decay already at time zero?
                 krTheta = kr*(Sv*cos(theta0(k))^2+Sp*sin(theta0(k))^2)/S0;
                 if decayNumbers(1)<=(krTheta+knr)*dt          % decay, rad. or non-rad.
                    if decayNumbers(1)<=krTheta/(krTheta+knr) % decay is radiative
86                       time(k)=1;
                    end
                    continue
                 end
                 for j=2:nt
91                   theta = acos(r(3));           % on unit sphere: theta=acos(z)
                     phi   = atan2(r(2),r(1)); % phi=atan(y/x)
                     M     = [cos(theta)*cos(phi) -sin(phi) sin(theta)*cos(phi);...
                              cos(theta)*sin(phi) cos(phi) sin(theta)*sin(phi); -sin(theta) 0
                                 cos(theta)]; % rotation matrix
                     r     = M*[xstep(j); ystep(j); zstep];          % rotate North-pole step
96                   krTheta = kr*(Sv*cos(theta)^2+Sp*sin(theta)^2)/S0; %theta-dep.rad.rate
                     if decayNumbers(j)<=(krTheta+knr)*dt          % decay, rad. or non-rad
                        if decayNumbers(j)<=krTheta/(krTheta+knr) % decay is radiative
                           time(k)=j;
                        end
101                     break
                     end
                 end
             end

106         % calculate average photon arrival time
            tauV(rc) = MeanNaN(time)*dt; % MeanNaN ignores all entries that are NaN
        end

        tauMean(Dcounter) = mean(tauV);
111     tauStd(Dcounter)  = std(tauV);
    end
    R   = (3*steplengths.^2/2)./rates;     % R=tau0/taurot=tau0*6D=tau0*3l^2/2
    tau = [tauMean.*rates; tauStd.*rates]; % relative time tau/tau0
```

# 6.2 Additional figures
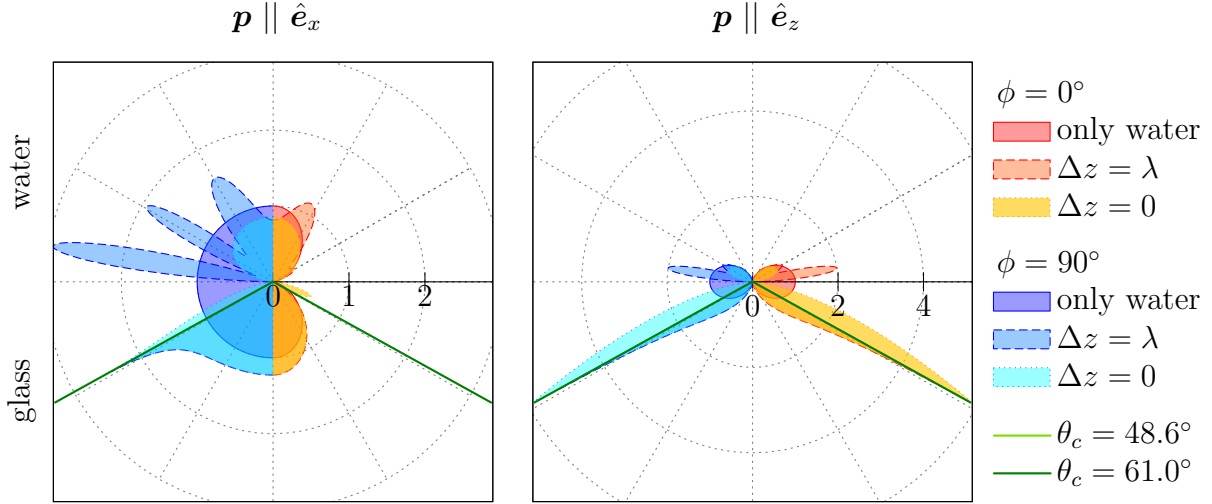


**Figure 6.1:** Angular distribution of radiation (ADR) of an oscillating dipole emitter oriented parallel to the $x$- or $z$-axis (left or right image) and placed in water near a 5 nm thick silver film on glass. The distance $\Delta z$ between the surface of the silver layer and the dipole is zero (dotted line) or $\lambda$ (dashed line). As a reference, a dipole in water without interface is shown with a solid line. All values are normalized to the maximum value of the water only case.
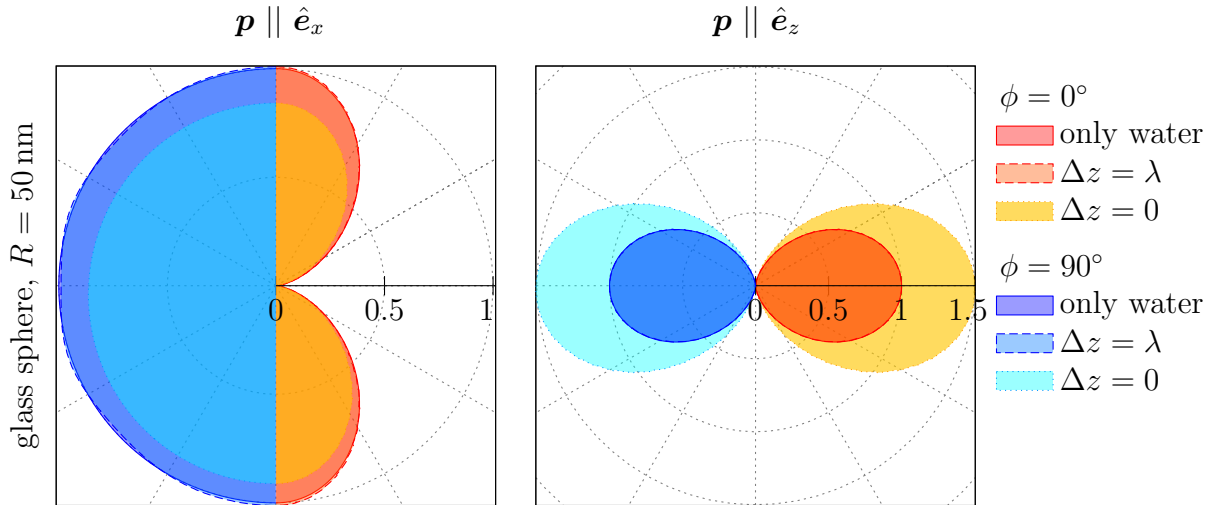


**Figure 6.2:** ADR of a dipole emitter oriented parallel to the $x$- or $z$-axis (left or right image) and placed in water at different height on the $z$-axis close to a sphere made of glass. The wavelength is $\lambda = 500$ nm, the radius of the sphere is $R = 50$ nm, the distance $\Delta z$ between the surface of the sphere and the dipole is zero (dotted line) or $\lambda$ (dashed line). As a reference, a dipole in water without sphere is shown with a solid line. All values are normalized to the maximum value of the water only case.
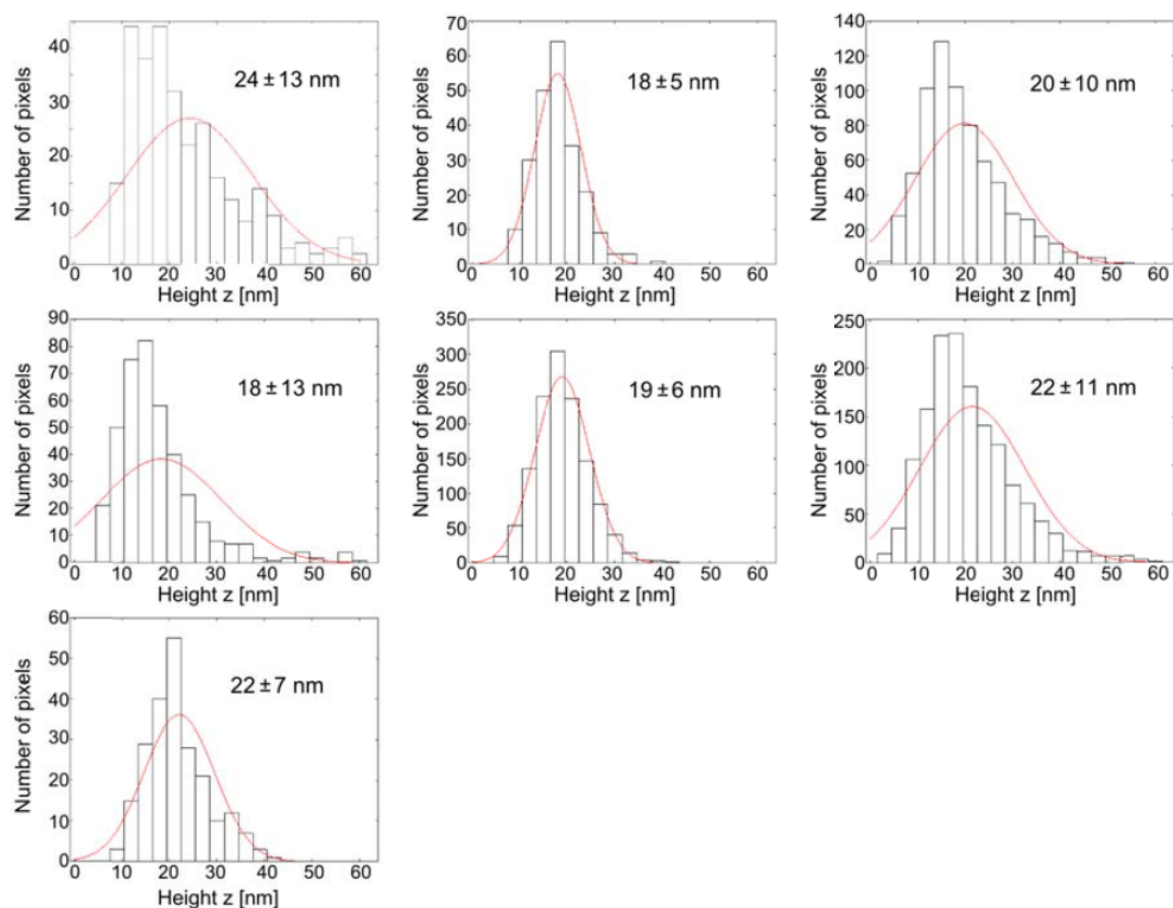
**Figure 6.3:** Height distributions of vinculin in seven different human mesenchymal stem cells 24 h after seeding. Vinculin was labeled with the fluorescent dye Atto 488, fluorescence lifetimes were acquired and converted to height values using a MIET calibration curve. Vinculin clusters were identified as explained in section 4.1.3, and the height values of all pixels belonging to such a cluster were histogrammed. Red solid lines correspond to Gaussian distributions with the same mean and standard deviation as the histograms. The overall mean value of $(20 \pm 8)$ nm agrees well with the value for a cell fixed at 24 h in figure 4.15(d), $(19 \pm 6)$ nm. Data recorded and analyzed by Anna Chizhik, image published in the supplementary information of [4].
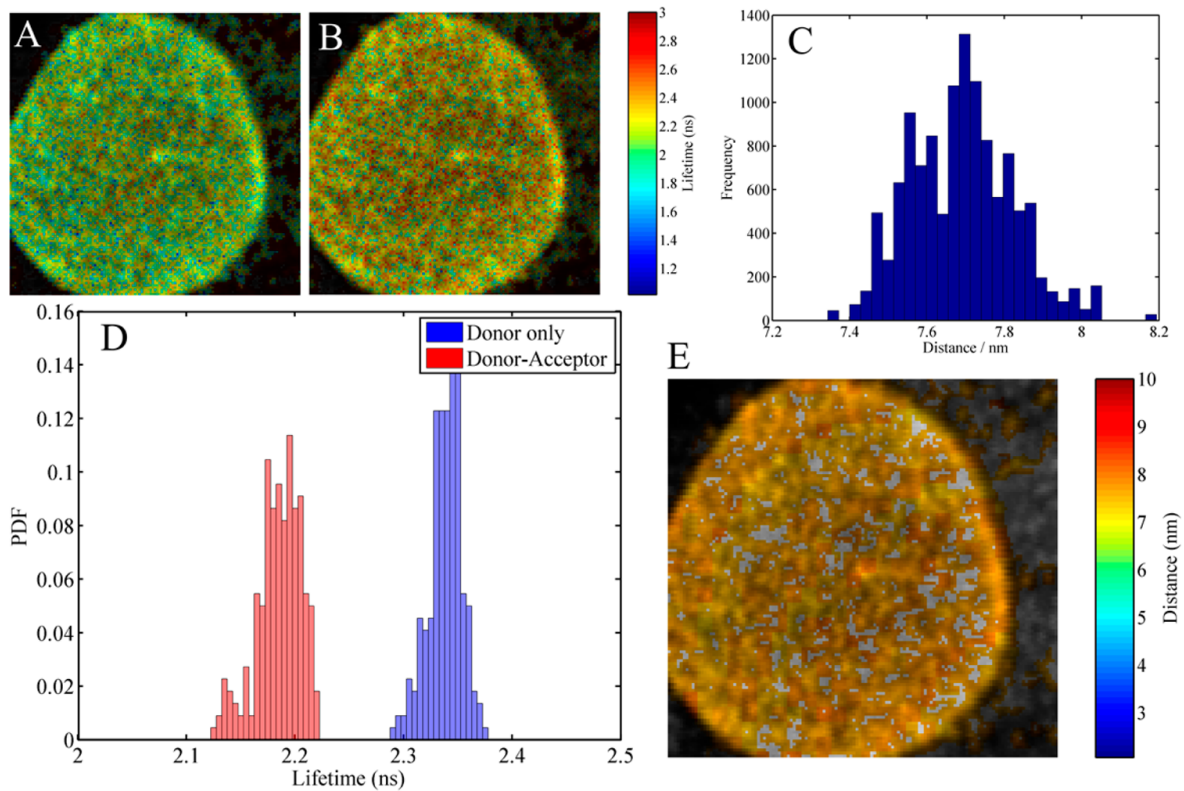
**Figure 6.4:** FRET between Lap2$\beta$-A488 and Nup358-A633, see section 4.1.4. (A, B) Fluorescence lifetime of A488 before and after bleaching of the acceptor A633, respectively. For clusters of 4×4 pixels, average lifetimes $\tau_{DA}$ and $\tau_D$ were determined as described in section 4.1.3 and converted to donor-acceptor distances assuming a one-on-one interaction of donor and acceptor. The resulting distances are histogrammed in C and visualized in the false-colour image E. For comparison, $\tau_{DA}$ and $\tau_D$ are histogrammed in D. Data recorded by Anna Chizhik and analyzed by Narain Karedla, image published in [3].
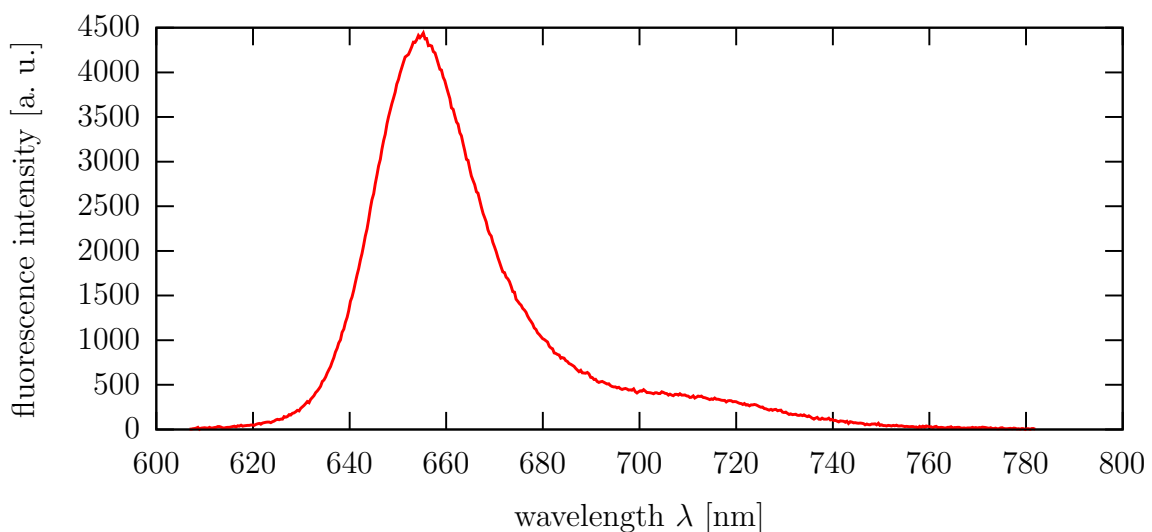


**Figure 6.5:** Fluorescence spectrum of the membrane-staining dye dSQ12S whose fluorescence quantum yield was determined in section 4.3.1, see also figure 4.37. Data provided by Alexey Chizhik.
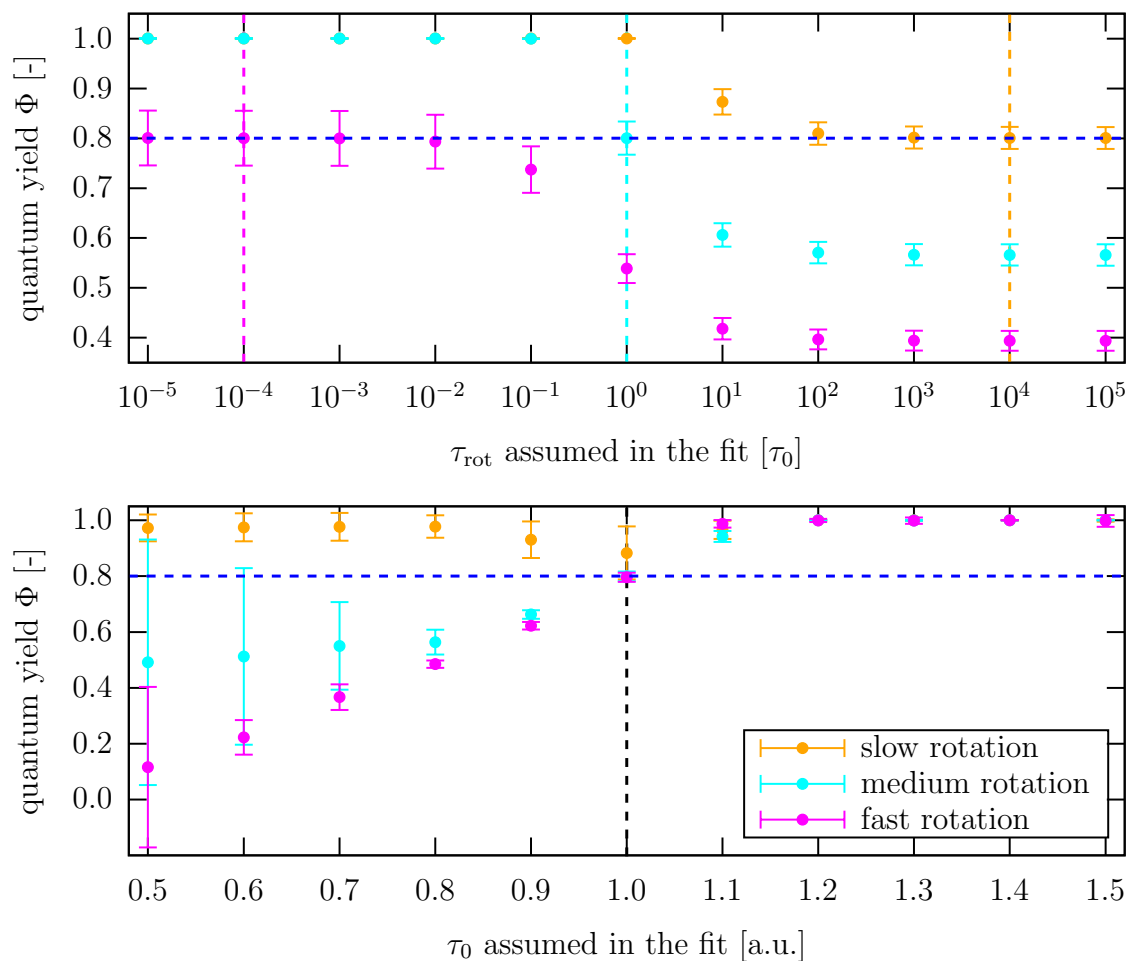
**Figure 6.6:** Fit parameters obtained by evaluating 100 data sets such as the one shown in figure 4.42, but now with quantum yield $\Phi = 0.8$ instead of 0.5. Either $\tau_{\text{rot}}$ (top) or $\tau_0$ (bottom) were fixed to the values indicated on the axis. The simulations were carried out for three ratios $R = \tau_0/\tau_{\text{rot}}$, namely $R = 10^{-4}$ ('slow rotation', orange), $R = 1$ ('medium rotation', cyan) and $R = 10^4$ ('fast rotation', pink). Vertical and horizontal lines indicate the underlying true parameter values. The legend is valid for both plots.

## 6.3 Additional contributions

Besides the MIET and nanocavity measurements presented in chapter 4, the electromagnetic models described in this work were applied to a number of other projects. These mainly include other MIET studies, for example the axial colocalization of single fluorophores attached to DNA origami structures with an accuracy of 5 nm [8], or a joint theoretical and experimental study of the three-dimensional localization of single molecules [9]. One very different example is the calculation of an electric field close to silica nanospheres, where a decomposition of the field into vector spherical harmonics was used [5]. This work will be briefly presented in the following section.

Additionally, I was involved in the development of a technique which significantly improves TCSPC-based FLIM by correcting dead-time effects. This work is summarized in section 6.3.2.

### 6.3.1 Photoactivation of luminescent centers in single $SiO_2$ nanoparticles

Since the localization accuracy of all single molecule localization microscopy techniques is improved when more photons are detected from the fluorophore, increasing the photon yield of fluorescent labels is an important subject of current investigations. One approach is to find labels which can be re-activated after photobleaching. Silicon dioxide ($SiO_2$) nanospheres were identified as possible candidates. These particles can possess luminescent defects, similar to nanodiamonds, but because of the weaker chemical bonds these defects can be generated by UV light. The fluorescence of $SiO_2$ nanoparticles of varying sizes during the simultaneous illumination with a pulsed 488 nm laser beam for exciting fluorescence and UV light (378 nm) for photogeneration of luminescent centers was studied in:

> [5] L. Tarpani, **D. Ruhlandt**, L. Latterini, D. Hähnel, I. Gregor, J. Enderlein, and A. I. Chizhik, "Photoactivation of luminescent centers in single $SiO_2$ nanoparticles," *Nano Letters*, vol. 16, no. 7, pp. 4312–4316, 2016. **D. R.** calculated the excitation intensity distributions for various particle sizes.

Based on the fluorescence time traces, it was found that some particles exhibited two or more periods of fluorescence, separated by dark times of several seconds. Since the fluorescence lifetime of each bright period was stable, but different from the lifetimes of other bright periods, this observation was interpreted as the creation and subsequent photobleaching of new luminescent defects. When comparing the relative number of reactivated $SiO_2$ nanoparticles, it was found that small particles showed significantly more new defects than larger spheres (blue bars in figure 6.7).

One possible explanation was that the small spheres strongly amplified the local intensity of the UV light. In order to test this hypothesis, I modeled the electric field strength in the air around the nanospheres, see figure 6.8. To this end, both the spherical interface between a $SiO_2$ nanoparticle and the surrounding air, as well as the planar interface between the air and the glass cover slip on which the particle rests had to be taken into account. This was accomplished iteratively: Firstly, the field in the absence of the nanosphere was calculated using the approach of Wolf and Richards [142, 143], in
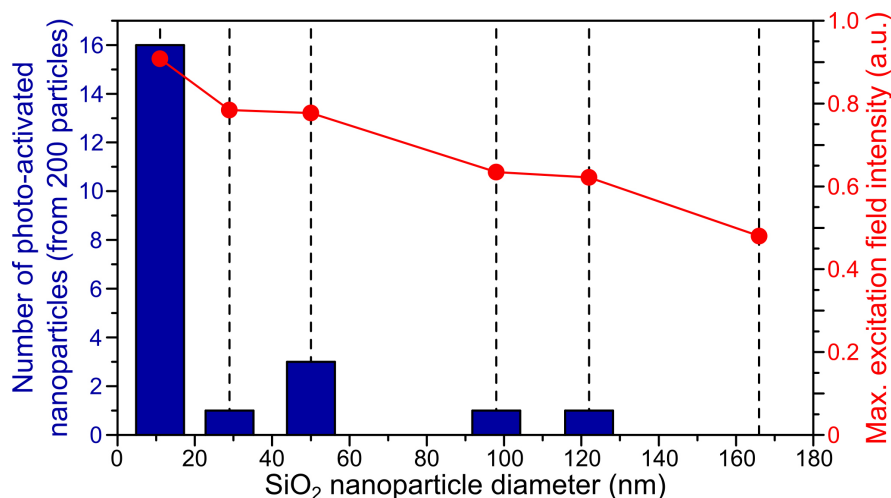
**Figure 6.7:** Number of reactivated SiO$_2$ nanoparticles out of 200 observed particles per particle size (blue histogram) and maximum excitation field intensity near the surface of the particles (red dots, see figure 6.8 for more details). Experimental data recorded by Alexey Chizhik, image published in [5].

the same way as described in section 4.3.1 when calculating the excitation intensity inside a nanocavity. Secondly, the resulting field was decomposed into vector spherical harmonics (VSH), which were then scattered at the nanosphere. This step made use of the theory described in sections 2.1.4 (*Conversion between PW and VSH*) and 2.3.3 (*VSH at a spherical interface*). The scattered field was then again decomposed in plane waves, which were allowed to interact with the planar interface, the reflected field converted to VSH and so on. The magnitudes of the scattered or reflected fields descrease quickly, leading to a convergence of the sum of all fields after approximately ten iterations.

Based on these results, the maximum excitation field intensity (in arbitrary but consistent units) at the particle surface was compared for all particle sizes, see red dots in figure 6.7. Since the maximum excitation intensity decreases much less dramatically with increasing particle size than the number of reactivations, it was ruled out as the dominant effect. Instead, it was proposed that the previously shown [158] distortion of the chemical bonds on the surface of ultrasmall SiO$_2$ particles increases the chemical reactivity and can thus facilitate the formation of structural defects.
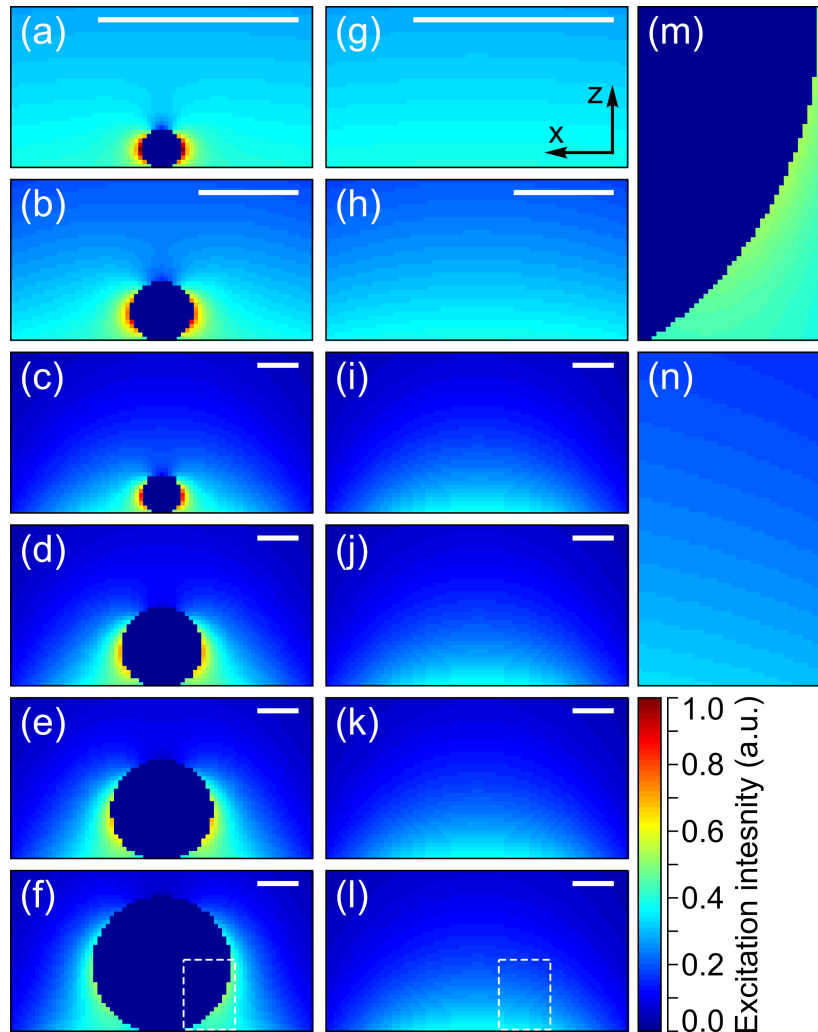
**Figure 6.8:** Calculated excitation field intensity distributions for a linearly polarized laser beam ($\lambda = 378$ nm) focused at the interface between air and a cover slip (objective's NA = 1.49). Shown is only the air halfspace, the glass halfspace is below the field of view. (a-f) $SiO_2$ nanoparticles of different sizes are aligned with the optical axis. (g-l) Calculations in the absence of the nanoparticles. Scale bars 50 nm. (m-n) Excitation field intensity distribution within the dashed rectangles in (f) and (l), respectively, calculated with the same spatial resolution as the one in image (a). All of the images are normalized to the same intensity scale. Figure published in [5].

## 6.3.2 Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging
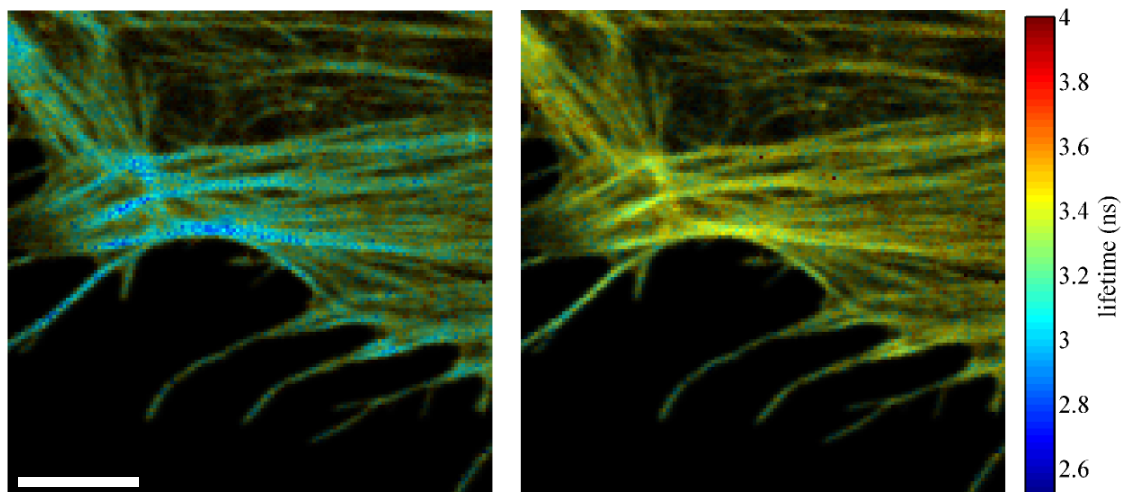


**Figure 6.9:** Effect of dead-time correction on exemplary FLIM data: Human mesenchymal stem cell whose actin filaments were labeled with Atto 647N, and imaged with a confocal scanning TCSPC microscope. Fluorescence lifetime before (left) and after (right) dead-time correction. The highest number of photon hits per excitation cycle was 0.5, the scale bar corresponds to 5 μm. Note that the corrected lifetimes are more homogeneous and overall higher than the uncorrected values. Sample prepared by Carina Wollnik, data recorded by Anna Chizhik and analyzed by Narain Karedla, published in [8].

Modern TCSPC systems such as those used for the FLIM measurements presented in the present work measure the arrival times of detected photons independent of the laser pulse times, and only correlate the two retrospectively. However, both the photon detectors as well as the timing electronics have dead times, i.e. they need to recover between two subsequent detection/timing events. Typically, these dead-times are in the order of tens of nanoseconds, and thus longer than standard fluorescence lifetimes ($\sim 1\,\text{ns} - 5\,\text{ns}$) but similar to the inverse of the laser pulse repetition rate ($\sim 20\,\text{MHz} - 80\,\text{MHz}$). Especially at high photon count rates, which are desirable in FLIM scan-images to reduce the required pixel dwell time, dead-time effects can lead to a loss of recorded photons and a distortion of the measured TCSPC curve. Our group has developed a data evaluation method which provides dead-time corrected TCSPC curves that faithfully reproduce the correct fluorescence decay curves even at high count rates. My contribution mainly consisted of the implementation of the algorithm in Matlab, including an optimization of time-critical routines as C++ MEX files. The method, its theoretical derivation and several examples were published in:

[6] S. Isbaner, N. Karedla, **D. Ruhlandt**, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein, "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging," *Optics Express*, vol. 24, no. 9, pp. 9429–9445, 2016. **D. R.** contributed to the implementation of the algorithm in Matlab, including C++ MEX files.

# 7 List of Figures

# List of Figures

# 8 Bibliography

[1] T. Baronsky, **D. Ruhlandt**, B. R. Brückner, J. Schäfer, N. Karedla, S. Isbaner, D. Hähnel, I. Gregor, J. Enderlein, A. Janshoff, and A. I. Chizhik, "Cell-substrate dynamics of the epithelial-to-mesenchymal transition," *Nano Letters*, vol. 17, no. 5, pp. 3320–3326, 2017.

[2] F. Schneider, **D. Ruhlandt**, I. Gregor, J. Enderlein, and A. I. Chizhik, "Quantum yield measurements of fluorophores in lipid bilayers using a plasmonic nanocavity," *Journal of Physical Chemistry Letters*, vol. 8, no. 7, pp. 1472–1475, 2017.

[3] A. M. Chizhik, **D. Ruhlandt**, J. Pfaff, N. Karedla, A. I. Chizhik, I. Gregor, R. H. Kehlenbach, and J. Enderlein, "Three-dimensional reconstruction of nuclear envelope architecture using dual-color metal-induced energy transfer imaging," *ACS Nano*, vol. 11, no. 12, pp. 11839–11846, 2017.

[4] A. M. Chizhik, C. Wollnik, **D. Ruhlandt**, N. Karedla, A. I. Chizhik, L. Hauke, D. Hähnel, I. Gregor, J. Enderlein, and F. Rehfeldt, "Dual-color metal-induced and Förster resonance energy transfer for cell nanoscopy," *Molecular Biology of the Cell*, vol. 29, no. 7, pp. 846–851, 2018.

[5] L. Tarpani, **D. Ruhlandt**, L. Latterini, D. Hähnel, I. Gregor, J. Enderlein, and A. I. Chizhik, "Photoactivation of luminescent centers in single $SiO_2$ nanoparticles," *Nano Letters*, vol. 16, no. 7, pp. 4312–4316, 2016.

[6] S. Isbaner, N. Karedla, **D. Ruhlandt**, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein, "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging," *Optics Express*, vol. 24, no. 9, pp. 9429–9445, 2016.

[7] N. Karedla, **D. Ruhlandt**, A. M. Chizhik, J. Enderlein, and A. I. Chizhik, *Advanced Photon Counting*, ch. Metal-Induced Energy Transfer, pp. 265–281. Springer, 2014.

[8] S. Isbaner, N. Karedla, I. Kaminska, **D. Ruhlandt**, M. Raab, J. Bohlen, A. I. Chizhik, I. Gregor, P. Tinnefeld, J. Enderlein, and R. Tsukanov, "Axial colocalization of single molecules with nanometer accuracy using metal-induced energy transfer," *Nano Letters*, vol. 18, no. 4, pp. 2616–2622, 2018.

[9] N. Karedla, A. M. Chizhik, S. C. Stein, **D. Ruhlandt**, I. Gregor, A. I. Chizhik, and J. Enderlein, "Three-dimensional single-molecule localization with nanometer accuracy using metal-induced energy transfer (MIET) imaging," *The Journal of Chemical Physics*, vol. 148, no. 20, p. 204201, 2018.

[10] C. Huygens, *Traité de la lumière: Où sont expliquées les causes de ce qui luy arrive dans la reflexion, & dans la refraction. Et particulierement dans l'etrange refraction du cristal d'Islande.* Chez Pierre vander Aa, 1690.

[11] I. Newton, *Opticks: or, a treatise of the reflexion, refraction, inflexions and colours of light.* Smith and Walford, 1704.

[12] H. C. Oersted, "Experimenta circa effectum conflictus electrici in acum magneticam." Pamphlet, 1820.

[13] O. Darrigol, *Electrodynamics from Ampère to Einstein.* Oxford University Press, 2000.

[14] J. C. Maxwell, "III. on physical lines of force," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 23, no. 151, pp. 12–24, 1862.

[15] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," *Annalen der Physik*, vol. 322, no. 6, pp. 132–148, 1905.

[16] B. R. Wheaton, *The tiger and the shark. Empirical roots of wave–particle dualism.* Cambridge University Press, 1983.

[17] E. K. Abbe, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv für mikroskopische Anatomie*, vol. 9, pp. 413–468, 1873.

[18] P. Kukura, H. Ewers, C. Müller, A. Renn, A. Helenius, and V. Sandoghdar, "High-speed nanoscopic tracking of the position and orientation of a single virus," *Nature Methods*, vol. 6, no. 12, pp. 923–927, 2009.

[19] L. Limozin and K. Sengupta, "Quantitative reflection interference contrast microscopy (RICM) in soft matter and cell adhesion," *ChemPhysChem*, vol. 10, no. 16, pp. 2752–2768, 2009.

[20] The Royal Swedish Academy of Sciences, "The nobel prize in chemistry 2014." Press release. www.nobelprize.org/prizes/chemistry/2014/press-release.

[21] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Optics Letters*, vol. 19, no. 11, pp. 780–782, 1994.

[22] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 15, pp. 8206–8210, 2000.

[23] K. Stock, R-Sailer, W. S. L. Strauss, M. Lyttek, R. Steiner, and H. Schneckenburger, "Variable-angle total internal reflection fluorescence microscopy (VA-TIRFM): realization and application of a compact illumination device," *Journal of Microscopy*, vol. 211, no. 1, pp. 19–29, 2003.

[24] M. C. D. Santos, R. Déturche, C. Vézy, and R. Jaffiol, "Topography of cells revealed by variable-angle total internal reflection fluorescence microscopy," *Biophysical Journal*, vol. 111, no. 6, pp. 1316–1327, 2016.

[25] J. Bewersdorf, R. Schmidt, and S. W. Hell, "Comparison of I$^5$M and 4Pi-microscopy," *Journal of Microscopy*, vol. 222, no. 2, pp. 105–117, 2006.

[26] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[27] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods*, vol. 3, no. 10, pp. 793–796, 2006.

[28] A. Sharonov and R. M. Hochstrasser, "Wide-field subdiffraction imaging by accumulated binding of diffusing probes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 50, pp. 18911–18916, 2006.

[29] M. Heilemann, S. van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, "Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes," *Angewandte Chemie, International Edition*, vol. 47, no. 33, pp. 6172–6176, 2008.

[30] F. Curdt, S. J. Herr, T. Lutz, R. Schmidt, J. Engelhardt, S. J. Sahl, and S. W. Hell, "isoSTED nanoscopy with intrinsic beam alignment," *Optics Express*, vol. 23, no. 24, pp. 30891–30903, 2015.

[31] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[32] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 9, pp. 2995–2999, 2009.

[33] M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf, "Three-dimensional sub-100 nm resolution fluorescene microscopy of thick samples," *Nature Methods*, vol. 5, no. 6, pp. 527–529, 2008.

[34] G. Shtengel, J. A. Galbraith, C. G. Galbraith, J. Lippincott-Schwartz, J. M. Gillette, S. Manley, R. Sougrat, C. M. Waterman, P. Kanchanawong, M. W. Davidson, R. D. Fetter, and H. F. Hess, "Interferometric fluorescent super-resolution microscopy reveals 3D cellcell ultrastructure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 9, pp. 3125–3130, 2009.

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

# Bibliography

[36] E. M. Purcell, "Spontaneous emission probabilities at radio frequencies," *Physical Review*, vol. 69, no. 11-12, p. 681, 1946.

[37] K.-H. Drexhage, "Influence of a dielectric interface on fluorescence decay time," *Journal of Luminescence*, vol. 1-2, pp. 693–701, 1970.

[38] M. Berndt, M. Lorenz, J. Enderlein, and S. Diez, "Axial nanometer distances measured by fluorescence lifetime imaging microscopy," *Nano Letters*, vol. 10, no. 4, pp. 1497–1500, 2010.

[39] A. I. Chizhik, J. Rother, I. Gregor, A. Janshoff, and J. Enderlein, "Metal-induced energy transfer for live cell nanoscopy," *Nature Photonics*, vol. 8, pp. 124–127, 2014.

[40] N. Karedla, A. I. Chizhik, I. Gregor, A. M. Chizhik, O. Schulz, and J. Enderlein, "Single-molecule metal-induced energy transfer (smMIET): Resolving nanometer distances at the single-molecule level," *ChemPhysChem*, vol. 15, no. 4, pp. 705–711, 2014.

[41] D. J. Griffiths, *Introduction to electrodynamics*. Pearson Education, 3rd ed., 2009.

[42] E. Hecht, *Optics*. Addison Wesley Longman, 4th ed., 2002.

[43] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. John Wiley & Sons, 2007.

[44] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*. Wiley-VCH, 1998.

[45] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*. International Series in Pure and Applied Physics, McGraw-Hill Book Company, Inc., 1953.

[46] A. Sommerfeld, *Partielle Differentialgleichungen der Physik*. Vorlesungen über Theoretische Physik, Verlag Harri Deutsch, 6 ed., 1992.

[47] J. A. Stratton, *Electromagnetic Theory*. McGraw-Hill Book Company, 1941.

[48] J. Nearing, *Mathematical Tools for Physixs*. Dover Publications, Inc., 2010.

[49] X. Yu, Q. Ye, H. Chen, S. Liu, and Z. Lin, "Simple algorithm for partial wave expansion of plasmonic and evanescent fields," *Optics Express*, vol. 25, no. 4, pp. 4201–4215, 2017.

[50] N. Karedla, *Single-Molecule Metal-Induced Energy Transfer: From Basics to Applications*. PhD thesis, Georg-August Universität Göttingen, 2016.

[51] H. Weyl, "Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter," *Annalen der Physik*, vol. 365, no. 21, pp. 481–500, 1919.

[52] M. Böhmer and J. Enderlein, "Orientation imaging of single molecules by wide-field epifluorescence microscopy," *J. Opt. Soc. Am. B*, vol. 20, no. 3, pp. 554–559, 2003.

[53] P. B. Johnson and R. W. Christy, "Optical constants of the noble metals," *Physical Review B*, vol. 6, no. 12, pp. 4370–4379, 1972.

[54] P. B. Johnson and R. W. Christy, "Optical constants of transition metals: Ti, v, cr, mn, fe, co, ni, and pd," *Physical Review B*, vol. 9, no. 12, pp. 5056–5070, 1974.

[55] S. A. Maier, *Plasmonics: Fundamentals and Applications*. Springer, 2010.

[56] K. Joulain, J.-P. Mulet, F. Marquier, R. Carminati, and J.-J. Greffet, "Surface electromagnetic waves thermally excited: Radiative heat transfer, coherence properties and casimir forces revisited in the near field," *Surface Science Reports*, vol. 57, no. 3-4, pp. 59–112, 2005.

[57] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Springer, 2006.

[58] H. C. Ishikawa-Ankerhold, R. Ankerhold, and G. P. C. Drummen, "Advanced fluorescence microscopy techniques—FRAP, FLIP, FLAP, FRET and FLIM," *Molecules*, vol. 17, no. 4, pp. 4047–4132, 2012.

[59] I. N. Levine, *Quantum Chemistry*. Pearson, 7th ed., 2014.

[60] S. A. Prahl, "Perylene." https://omlc.org/spectra/PhotochemCAD/html/023.html, June 2017.

[61] L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent, and L. E. Hood, "Fluorescence detection in automated DNA sequence analysis," *Nature*, vol. 321, no. 6071, pp. 674–679, 1986.

[62] P. Y. Lee, J. Costumbrado, C.-Y. Hsu, and Y. H. Kim, "Agarose gel electrophorasis for the separation of DNA fragments," *Journal of Visualized Experiments*, vol. 62, p. 3923, 2012.

[63] R. Phillips, J. Kondev, and J. Theriot, *Physical Biology of the Cell*. Ney York: Garland Science, 2009.

[64] E. M. Goldys, K. Drozdowicz-Tomsia, G. Zhu, H. Yu, S. Jinjun, M. Motlan, and M. Godlewski, "Fluorescence labelling," *Optica Applicata*, vol. 36, no. 2-3, pp. 217–224, 2006.

[65] U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, and T. Nann, "Quantum dots versus organic dyes as fluorescent labels," *Nature Methods*, vol. 5, no. 9, pp. 763–775, 2008.

[66] M. Baruah, W. Qin, C. Flors, J. Hofkens, R. A. L. Vallée, D. Beljonne, M. Van der Auweraer, W. M. De Borggraeve, and N. Boens, "Solvent and pH dependent fluorescent properties of a dimethylaminostyryl borondipyrromethene dye in solution," *The Journal of Physical Chemistry A*, vol. 110, no. 18, pp. 5998–6009, 2006.

[67] M. M. Martin and L. Lindqvist, "The pH dependence of fluorescein fluorescence," *Journal of Luminescence*, vol. 10, no. 6, pp. 381–390, 1975.

[68] Y. Ando and H. Akiyama, "ph-dependent fluorescence spectra, lifetimes, and quantum yields of firefly-luciferin aqueous solutions studied by selective-excitation fluorescence spectroscopy," *Japanese Journal of Applied Physics*, vol. 49, no. 11R, p. 117002, 2010.

[69] W. Rettig and R. Lapouyade, *Topics in Fluorescence Spectroscopy*, vol. 4, ch. Fluorescence Probes Based on Twisted Intramolecular Charge Transfer (TICT) States and Other Adiabatic Photoreactions, pp. 109–149. Springer, 2002.

[70] A. P. Jackson, M. P. Timmerman, C. R. Bagshaw, and C. C. Ashley, "The kinetics of calcium binding to fura-2 and indo-1," *FEBS Letters*, vol. 216, no. 1, pp. 35–39, 1987.

[71] M. Kneen, J. Farinas, Y. Li, and A. S. Verkman, "Green fluorescent protein as a noninvasive intracellular pH indicator," *Biophysical Journal*, vol. 74, no. 3, pp. 1591–1599, 1998.

[72] I. N. Bronštein, K. A. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 7th ed., 2008.

[73] W. Schottky, "Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern," *Annalen der Physik*, vol. 362, no. 23, pp. 541–567, 1918.

[74] A. Giudice, M. Ghioni, R. Biasi, F. Zappa, S. Cova, P. Maccagnani, and A. Gulinatti, "High-rate photon counting and picosecond timing with silicon-spad based compact detector modules," *Journal of Modern Optics*, vol. 54, no. 2-3, pp. 225–237, 2007.

[75] A. Ghosh, S. Isbaner, M. Veiga-Gutiérrez, I. Gregor, J. Enderlein, and N. Karedla, "Quantifying microsecond transition times using fluorescence lifetime correlation spectroscopy," *Journal of Physical Chemistry Letters*, vol. 8, no. 24, pp. 6022–6028, 2017.

[76] K. B. Blodgett, "Films built by depositing successive monomolecular layers on a solid surface," *Jorunal of the American Chemical Society*, vol. 57, no. 6, pp. 1007–1022, 1935.

[77] D. J. Griffiths, *Introduction to Quantum Mechanics*. Cambridge University Press, 2nd ed., 2017.

[78] J. Enderlein, "Probing vavuum electromagnetic field modes in complex environments with an electric dipole emitter." unpublished.

[79] K. Joulain, R. Carminati, J.-P. Mulet, and J.-J. Greffet, "Definition and measurement of the local density of electromagnetic states close to an interface," *Physical Review B*, vol. 68, no. 24, p. 245405, 2003.

[80] H. Kuhn, "Classical aspects of energy transfer in molecular systems," *The Journal of Chemical Physics*, vol. 53, no. 1, pp. 101–108, 1970.

[81] R. R. Chance, A. Prock, and R. Silbey, "Lifetime of an emitting molecule near a partially reflecting surface," *The Journal of Chemical Physics*, vol. 60, no. 7, pp. 2744–2748, 1974.

[82] R. R. Chance, A. Prock., and R. Silbey, "Comments on the classical theory of energy transfer," *The Journal of Chemical Physics*, vol. 62, no. 6, pp. 2245–2253, 1975.

[83] R. R. Chance., A. H. Miller, A. Prock, and R. Silbey, "Luminescent lifetimes near multiple interfaces: a quantitative comparison of theory and experiment," *Chemical Physics Letters*, vol. 33, no. 3, pp. 590–592, 1975.

[84] W. Lukosz and R. E. Kunz, "Light emission by magnetic and electric dipoles close to a plane interface. i. total radiated power," *Journal of the Optical Society of America*, vol. 67, no. 12, pp. 1607–1615, 1977.

[85] R. M. Amos and W. L. Barnes, "Modification of the spontaneous emission rate of $Eu^{3+}$ ions close to a thin metal mirror," *Physical Review B*, vol. 55, no. 11, pp. 7249–7254, 1997.

[86] S. Haroche and D. Kleppner, "Cavity quantum electrodynamics," *Physics Today*, vol. 42, no. 1, pp. 24–30, 1989.

[87] E. A. Hinds, "Cavity quantum electrodynamics," *Advances in Atomic, Molecular, and Optical Physics*, vol. 28, pp. 237–289, 1990.

[88] T. Tsutsui, C. Adachi, S. Saito, M. Watanabe, and M. Koishi, "Effect of confined radiation field on spontaneous-emission lifetime in vacuum-deposited fluorescent dye films," *Chemical Physics Letters*, vol. 182, no. 2, pp. 143–146, 1991.

[89] J. Kümmerlen, A. Leitner, H. Brunner, F. R. Aussenegg, and A. Wokaun, "Enhanced dye fluorescence over silver island films: analysis of the distance dependence," *Molecular Physics*, vol. 80, no. 5, pp. 1031–1046, 1993.

[90] M. J. A. de Dood, L. H. Slooff, A. Polman, A. Moroz, and A. van Blaaderen, "Local optical density of states in $SiO_2$ spherical microcavities: Theory and experiment," *Physical Review A*, vol. 64, no. 3, p. 033807, 2001.

[91] H. Schniepp and V. Sandoghdar, "Spontaneous emission of europium ions embedded in dielectric nanospheres," *Physical Review Letters*, vol. 89, no. 25, p. 257403, 2002.

[92] W. L. Vos and A. Polman, "Optical probes inside photonic crystals," *MRS Bulletin*, vol. 26, no. 8, pp. 642–646, 2001.

[93] N. Tomczak, S. Gu, M. Han, N. F. van Hulst, and G. J. Vansco, "Single light emitters in electrospun polymer nanofibers: Effect of local confinement on radiative decay," *European Polymer Journal*, 2006.

## Bibliography

[94] L. Rogobete, H. Schniepp, V. Sandoghdar, and C. Henkel, "Spontaneous emission in nanoscopic dielectric particles," *Optics Letters*, vol. 28, no. 19, pp. 1736–1738, 2003.

[95] A. Moroz, "Spectroscopic properties of a two-level atom interacting with a complex spherical nanoshell," *Chemical Physics*, vol. 317, no. 1, pp. 1–15, 2005.

[96] P. Török and T. Wilson, "Rigorous theory for axial resolution in confocal microscopes," *Optics Communications*, vol. 137, no. 1-3, pp. 127–135, 1997.

[97] C. M. Ajo-Franklin, P. V. Ganesan, and S. G. Boxer, "Variable incidence angle fluorescence interference contrast microscopy for z-imaging single objects," *Biophysical Journal*, vol. 89, no. 4, pp. 2759–2769, 2005.

[98] A. Lambacher and P. Fromherz, "Fluorescence interference-contrast microscopy on oxidized silicon using a monomolecular dye layer," *Applied Physics A*, vol. 63, no. 3, pp. 207–216, 1996.

[99] D. M. Mattox, "Thin film metallization of oxides in microelectronics," *Thin Solid Films*, vol. 18, no. 2, pp. 173–186, 1973.

[100] "IEEE standard for floating-point arithmetic," *IEEE Std 754-2008*, 2008.

[101] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*. Cambridge University Press, 3rd ed., 2007.

[102] M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions*. Dover Publications, Inc., 1964, reprinted 1970.

[103] W. H. Richardson, "Bayesian-based iterative method of image restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55–59, 1972.

[104] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomical Journal*, vol. 79, p. 745, 1974.

[105] Željko Bajzer, T. M. Therneau, J. C. Sharp, and F. G. Prendergast, "Maximum likelihood method for the analysis of time-resolved fluorescence decay curves," *European Biophysics Journal*, vol. 20, no. 5, pp. 247–262, 1991.

[106] M. I. Rowley, A. C. C. Coolen, B. Vojnovic, and P. R. Barber, "Robust bayesian fluorescence lifetime estimation, decay model selection and instrument response determination for low-intensity FLIM imaging," *PLoS ONE*, vol. 11, no. 6, p. e0158404, 2016.

[107] A. I. Chizhik, I. Gregor, B. Ernst, and J. Enderlein, "Nanocavity-based determination of absolute values of photoluminescence quantum yields," *ChemPhysChem*, vol. 14, no. 3, pp. 505–513, 2013.

[108] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. Garland Science, 5th ed., 2007.

[109] D. Schneider, M. Tarantola, and A. Janshoff, "Dynamics of TGF-$\beta$ induced epithelial-to-mesenchymal transition monitored by electric cell-substrate impedance sensing," *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1813, no. 12, pp. 2099–2107, 2011.

[110] A. D. Rakić, A. B. Djurišić, J. M. Elazar, and M. L. Majewski, "Optical properties of metallic films for vertical-cavity optoelectronic devices," *Applied Optics*, vol. 37, no. 22, pp. 5271–5283, 1998.

[111] I. Kaverina, O. Krylyshkina, and J. V. Small, "Regulation of substrate adhesion dynamics during cell motility," *International Journal of Biochemistry & Cell Biology*, vol. 34, no. 7, pp. 746–761, 2002.

[112] D. Schneider, T. Baronsky, A. Pietuch, J. Rother, M. Oelkers, D. Fichtner, D. Wedlich, and A. Janshoff, "Tension monitoring during epithelial-to-mesenchymal transition links the switch of phenotype to expression of moesin and cadherins in NMuMG cells," *PLoS One*, vol. 8, no. 12, p. e80068, 2013.

[113] P. Kanchanawong, G. Shtengel, A. M. Pasapera, E. B. Ramko, M. W. Davidson, H. F. Hess, and C. M. Waterman, "Nanoscale architecture of integrin-based cell adhesions," *Nature*, vol. 468, no. 7323, pp. 580–584, 2010.

[114] A. J. Engler, S. Sen, H. L. Sweeney, and D. E. Discher, "Matrix elasticity directs stem cell lineage specification," *Cell*, vol. 126, no. 4, pp. 677–689, 2006.

[115] A. Pietraszewska-Bogiel and T. W. J. Gadella, "FRET microscopy: from principle to routine technology in cell biology," *Journal of Microscopy*, vol. 241, no. 2, pp. 111–118, 2011.

[116] E. A. Jares-Erijman and T. M. Jovin, "FRET imaging," *Nature Biotechnology*, vol. 21, no. 11, pp. 1387–1395, 2003.

[117] M. J. Paszek, C. C. DuFort, M. G. Rubashkin, M. W. Davidson, K. S. Thorn, J. T. Liphardt, and V. M. Weaver, "Scanning angle interference microscopy reveals cell dynamics at the nanoscale," *Nature Methods*, vol. 9, no. 8, pp. 825–827, 2012.

[118] W. W. Franke, U. Scheer, G. Krohne, and E.-D. Jarasch, "The nuclear envelope and the architecture of the nuclear periphery," *Journal of Cell Biology*, vol. 91, no. 3, pp. 39s–50s, 1981.

[119] K. C. Mudumbi, E. C. Schirmer, and W. Yang, "Single-point single-moledule FRAP distinguishes inner and outer nuclear membrane protein distribution," *Nature Communications*, vol. 7, p. 12562, 2016.

[120] A. Löschberger, S. van de Linde, M.-C. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, and M. Sauer, "Super-resolution imaging visualizes the eightfold symmetry of gp210 protein around the nuclear pore complex and resolves the central channel with nanometer resolution," *Journal of Cell Science*, vol. 125, no. 3, pp. 570–575, 2012.

# Bibliography

[121] J. Enderlein, *Comprehensive Biomedical Physics*, ch. Advanced Fluorescence Microscopy, pp. 111–151. Elsevier, 2014.

[122] K. Monier, J. C. G. Armas, S. Etteldorf, P. Ghazal, and K. F. Sullivan, "Annexation of the interchromosomal space during viral infection," *Nature Cell Biology*, vol. 2, no. 9, pp. 661–665, 2000.

[123] N. E. Cain and D. A. Starr, "SUN proteins and nuclear envelope spacing," *Nucleus*, vol. 6, no. 1, pp. 2–7, 2015.

[124] H.-J. van Manen, P. Verkuijlen, P. Wittendorp, V. Subramaniam, T. K. van den Berg, D. Roos, and C. Otto, "Refractive index sensing of green fluorescenct proteins in living cells using fluorescence lifetime imaging microscopy," *Biophysical Journal*, vol. 94, no. 8, pp. L67–L69, 2008.

[125] J. Beuthan, O. Minet, J. Helfmann, M. Herrig, and G. Müller, "The spatial variation of the refractive index in biological cells," *Physics in Medicine and Biology*, vol. 41, no. 3, pp. 369–382, 1996.

[126] M. Schürmann, J. Scholze, P. Müller, J. Guck, and C. J. Chan, "Cell nuclei have lower refractive index and mass density than cytoplasm," *Journal of Biophotonics*, vol. 9, no. 10, pp. 1068–1076, 2016.

[127] W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, and M. S. Feld, "Tomographic phase microscopy," *Nature Methods*, vol. 4, no. 9, pp. 717–719, 2007.

[128] C. L. Curl, C. J. Bellair, T. Harris, B. E. Allman, P. J. Harris, A. G. Stewart, A. Roberts, K. A. Nugent, and L. M. D. Delbridge, "Refractive index measurement in viable cells using quantitative phase-amplitude microscopy and confocal microscopy," *Cytometry Part A*, vol. 65A, no. 1, pp. 88–92, 2005.

[129] N. Lue, W. Choi, G. Popescu, Z. Yaqoob, K. Badizadegan, R. R. Dasari, and M. S. Feld, "Live cell refractometry using hilbert phase microscopy and confocal reflectance microscopy," *Journal of Physical Chemistry A*, vol. 113, no. 47, pp. 13327–13330, 2009.

[130] G. Nienhuis and C. T. J. Alkemade, "Atomic radiative transition probabilities in a continuous medium," *Physica (Amsterdam)*, vol. 81C, no. 1, pp. 181–188, 1976.

[131] G. L. J. A. Rikken and Y. A. R. R. Kessener, "Local field effects and electric and magnetic dipole transitions in dielectrics," *Physical Review Letters*, vol. 74, no. 6, pp. 880–883, 1995.

[132] F. J. P. Schuurmans, D. T. N. de Lang, G. H. Wegdam, R. Sprik, and A. Lagendijk, "Local-field effects on spontaneous emission in a dense supercritical gas," *Physical Review Letters*, vol. 80, no. 23, pp. 5077–5080, 1998.

[133] S. J. Strickler and R. A. Berg, "Relationship between absorption intensity and fluorescence lifetime of molecules," *Journal of Chemical Physics*, vol. 37, no. 4, pp. 814–822, 1962.

[134] K. Suhling, J. Siegel, D. Phillips, P. M. W. French, S. Lévêque-Fort, S. E. D. Webb, and D. M. Davis, "Imaging the environment of green fluorescent protein," *Biophysical Journal*, vol. 83, no. 6, pp. 3589–3595, 2002.

[135] J. N. Demas and G. A. Crosby, "The measurement of photoluminescence quantum yields. a review," *Journal of Physical Chemistry*, vol. 75, no. 8, pp. 991–1023, 1971.

[136] S. J. Wawilow, "Die Fluoreszenzausbeute von Farbstofflösungen," *Zeitschrift für Physik*, vol. 22, no. 1, pp. 266–272, 1924.

[137] C. Würth, M. Grabolle, J. Pauli, M. Spieles, and U. Resch-Genger, "Relative and absolute determination of fluorescence quantum yields of transparent samles," *Nature Protocols*, vol. 8, no. 8, pp. 1535–1550, 2013.

[138] L. Porrès, A. Holland, L.-O. Påsson, A. P. Monkman, C. Kemp, and A. Beeby, "Absolute measurements of photoluminescence quantum yields of solutions using an integrating sphere," *Journal of Fluorescence*, vol. 16, no. 2, pp. 267–272, 2006.

[139] K. Suzuki, A. Kobayashi, S. Kaneko, K. Takehira, T. Yoshihara, H. Ishida, Y. Shiina, S. Oishi, and S. Tobita, "Reevaluation of absolute luminescence quantum yields of standard solutions using a spectrometer with an integrating sphere and a back-thinned CCD detector," *Physical Chemistry Chemical Physics*, vol. 11, no. 42, pp. 9850–9860, 2009.

[140] G. H. Patterson and J. Lippincott-Schwartz, "A photoactivatable GFP for selective photolabeling of proteins and cells," *Science*, vol. 297, no. 5588, pp. 1873–1877, 2002.

[141] K. A. Lukyanov, D. M. Chudakov, S. Lukyanov, and V. V. Verkhusha, "Photoactivatable fluorescent proteins," *Nature Reviews Molecular Cell Biology*, vol. 6, pp. 885–890, 2005.

[142] E. Wolf, "Electromagnetic diffraction in optical systems I. An integral representation of the image field," *Proceedings of the Royal Society of London A*, vol. 253, no. 1274, pp. 349–357, 1959.

[143] B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems II. Structure of the image field in an aplanatic system," *Proceedings of the Royal Society of London A*, vol. 253, no. 1274, pp. 358–379, 1959.

[144] M. Collot, R. Kreder, A. L. Tatarets, L. D. Patsenker, Y. Mely, and A. S. Klymchenko, "Bright fluorogenic squaraines with tuned cell entry for selective imaging of plasma membrane vs. endoplasmatic reticulum," *Chemical Communications*, vol. 51, no. 96, pp. 17136–17139, 2015.

[145] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

[146] C. Steinem, A. Janshoff, W.-P. Ulrich, M. Sieber, and H.-J. Galla, "Impedance analysis of supported lipid bilayer membranes: a scrutiny of different preparation techniques," *Biochimica et Biophysica Acta - Biomembranes*, vol. 1279, no. 2, pp. 169–180, 1996.

[147] R. P. Richter and A. R. Brisson, "Following the formation of supported lipid bilayers on mica: A study combining AFM, QCM-D, and ellipsometry," *Biophysical Journal*, vol. 88, no. 5, pp. 3422–3433, 2005.

[148] R. P. Richter, R. Bérat, and A. R. Brisson, "Formation of solid-supported lipid bilayers: An integrated view," *Langmuir*, vol. 22, no. 8, pp. 3497–3505, 2006.

[149] R. Tero, "Substrate effects on the formation process, structure and physicochemical properties of supported lipid bilayers," *Materials*, vol. 5, no. 12, pp. 2658–2680, 2012.

[150] L. D. Hughes, R. J. Rawle, and S. G. Boxer, "Choose your label wisely: Water-soluble fluorophores often interact with lipid bilayers," *PLoS ONE*, vol. 9, no. 2, p. e87649, 2014.

[151] T. Weise, *Global Optimization Algorithms – Theory and Application*. Self-published, second ed., 2009. Available online at http://www.it-weise.de.

[152] P. D. Adams, Y. Chen, K. Ma, M. G. Zagorski, F. D. Sönnichsen, M. L. McLaughlin, and M. D. Barkley, "Intramolecular quenching of tryptophan fluorescence by the peptide bond in cyclic hexapeptides," *Journal of the American Chemical Society*, vol. 124, no. 31, pp. 9278–9286, 2002.

[153] J. W. Borst, M. A. Hink, A. van Hoek, and A. J. W. G. Visser, "Effects of refractive index and viscosity on fluorescence and anisotropy decays of enhanced cyan and yellow fluorescent proteins," *Journal of Fluorescence*, vol. 15, no. 2, pp. 153–160, 2005.

[154] M. Lelimousin, M. Noirclerc-Savoye, C. Lazareno-Saez, B. Paetzold, S. L. Vot, R. Chazal, P. Macheboeuf, M. J. Field, D. Bourgeois, and A. Royant, "Intrinsic dynamics in ECFP and cerulean control fluorescence quantum yield," *Biochemistry*, vol. 48, no. 42, pp. 10038–10046, 2009.

[155] J. Goedhart, D. von Stetten, M. Noirclerc-Savoye, M. Lelimousin, L. Joosen, M. A. Hink, L. van Weeren, T. W. J. G. Jr., and A. Royant, "Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93 %," *Nature Communications*, vol. 3, p. 751, 2012.

[156] D. S. Bindels, L. Haarbosch, L. van Weeren, M. Postma, K. E. Wiese, M. Mastop, S. Aumonier, G. Gotthard, A. Royant, M. A. Hink, and T. W. J. G. Jr., "mScarlet: a bright monomeric red fluorescent protein for cellular imaging," *Nature Methods*, vol. 14, no. 1, pp. 53–56, 2017.

[157] A. I. Chizhik, A. M. Chizhik, D. Khoptyar, S. Bär, A. J. Meixner, and J. Enderlein, "Probing the radiative transition of single molecules with a tunable microresonater," *Nano Letters*, vol. 11, no. 4, pp. 1700–1703, 2011.

[158] I. A. Rahman, P. Vejayakumaran, C. S. Sipaut, J. Ismail, and C. K. Chee, "Size-dependent physicochemical and optical properties of silica nanoparticles," *Materials Chemistry and Physics*, vol. 114, no. 1, pp. 328–332, 2009.

# Acknowledgements

# Curriculum vitae

**Daja Ruhlandt**

Date of birth:    20.10.1988

Place of birth:   Walsrode, Germany

Nationality:      German

# Academical background

| | |
|---|---|
| 01.2014 - 12.2018 | Doctoral dissertation<br>Thesis: *Modeling electrodynamics in the vicinity of metal nanostructures*<br>Georg-August-Universität, Göttingen |
| 10.2011 - 10.2013 | Master of Science *with distinction* in Physics<br>Thesis: *Determination of the stoichiometry of ion channels via fluorescence microscopy*<br>Georg-August-Universität, Göttingen |
| 10.2008 - 08.2011 | Bachelor of Science in Physics<br>Thesis: *Beobachtung einzelner, mit Kohlenstoffnanoröhrchen markierter Kinesin-Motorproteine in lebenden Zellen*<br>Georg-August-Universität, Göttingen |
| 08.2001 - 06.2008 | Abitur<br>Gymnasium Walsrode, Walrode |

# Publications

2018 | N. Karedla, A. M. Chizhik, S. C. Stein, **D. Ruhlandt**, I. Gregor, A. I. Chizhik, and J. Enderlein, "Three-dimensional single-molecule localization with nanometer accuracy using metal-induced energy transfer (MIET) imaging," *The Journal of Chemical Physics*, vol. 148, no. 20, p. 204201, 2018.

A. M. Chizhik, C. Wollnik, **D. Ruhlandt**, N. Karedla, A. I. Chizhik, L. Hauke, D. Hähnel, I. Gregor, J. Enderlein, and F. Rehfeldt, "Dual-color metal-induced and Förster resonance energy transfer for cell nanoscopy," *Molecular Biology of the Cell*, vol. 29, no. 7, pp. 846–851, 2018.

S. Isbaner, N. Karedla, I. Kaminska, **D. Ruhlandt**, M. Raab, J. Bohlen, A. I. Chizhik, I. Gregor, P. Tinnefeld, J. Enderlein, and R. Tsukanov, "Axial colocalization of single molecules with nanometer accuracy using metal-induced energy transfer," *Nano Letters*, vol. 18, no. 4, pp. 2616–2622, 2018.

2017 | A. M. Chizhik, **D. Ruhlandt**, J. Pfaff, N. Karedla, A. I. Chizhik, I. Gregor, R. H. Kehlenbach, and J. Enderlein, "Three-dimensional reconstruction of nuclear envelope architecture using dual-color metal-induced energy transfer imaging," *ACS Nano*, vol. 11, no. 12, pp. 11839–11846, 2017.

F. Schneider, **D. Ruhlandt**, I. Gregor, J. Enderlein, and A. I. Chizhik, "Quantum yield measurements of fluorophores in lipid bilayers using a plasmonic nanocavity," *Journal of Physical Chemistry Letters*, vol. 8, no. 7, pp. 1472–1475, 2017.

T. Baronsky, **D. Ruhlandt**, B. R. Brückner, J. Schäfer, N. Karedla, S. Isbaner, D. Hähnel, I. Gregor, J. Enderlein, A. Janshoff, and A. I. Chizhik, "Cell-substrate dynamics of the epithelial-to-mesenchymal transition," *Nano Letters*, vol. 17, no. 5, pp. 3320–3326, 2017.

2016 | L. Tarpani, **D. Ruhlandt**, L. Latterini, D. Hähnel, I. Gregor, J. Enderlein, and A. I. Chizhik, "Photoactivation of luminescent centers in single $SiO_2$ nanoparticles," *Nano Letters*, vol. 16, no. 7, pp. 4312–4316, 2016.

S. Isbaner, N. Karedla, **D. Ruhlandt**, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein, "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging," *Optics Express*, vol. 24, no. 9, pp. 9429–9445, 2016.

2014 | N. Karedla, **D. Ruhlandt**, A. M. Chizhik, J. Enderlein, and A. I. Chizhik, *Advanced Photon Counting*, ch. Metal-Induced Energy Transfer, pp. 265–281. Springer, 2014.