Near-Field Microscopy of Plasmonic Waveguides and Tight-Binding Lattices

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Hans-Joachim Schill
aus
Filderstadt

Bonn, 21.01.2026

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer:	Prof. Dr. Stefan Linden
Gutachterin:	Priv.-Doz. Dr. Elisabeth Soergel
Tag der Promotion:	25.03.2026
Erscheinungsjahr:	2026

Abstract

Electronic computing has dominated data processing for decades, but we are reaching a point where fundamental physical limitations motivate the exploration of alternative approaches. Photonic and plasmonic systems provide promising routes toward compact, high-speed, and low-energy information processing by enabling strong light–matter interactions at the nanoscale. In particular, plasmonic waveguides allow optical confinement below the diffraction limit, making them attractive building blocks for integrated nanophotonic circuits. Furthermore, many fundamental condensed matter or quantum effects can be investigated with coupled waveguide arrays. However, their characterization and control require advanced experimental techniques capable of resolving optical fields at subwavelength scales.

This thesis investigates plasmonic waveguides with a focus on their fabrication, near-field characterization, and functional coupling schemes. After an introduction to surface plasmon polaritons and their excitation, the thesis describes the fabrication methods used to realize plasmonic nanostructures and discusses the associated challenges. As a key tool for nanoscale optical characterization, the principles and experimental implementation of scattering-type scanning near-field optical microscopy are then presented.

These techniques are used to investigate three distinct plasmonic waveguide systems. First, metal strip waveguides with chiral couplers are studied, demonstrating spin-dependent directional excitation of surface plasmon polaritons based on the polarization of incident light. Second, dielectric-loaded plasmonic waveguides are investigated, including a detailed analysis of their electromagnetic field components. Building on this, coupled dielectric-loaded surface plasmon polariton waveguides are explored as a platform for studying coupling dynamics and analogies to solid-state systems. By implementing a plasmonic analogue of the Su–Schrieffer–Heeger model, topological edge states are experimentally realized and directly visualized using near-field microscopy.

Contents

1 Introduction **1**

2 Surface Plasmon Polaritons and Plasmonic Waveguides **5**
- 2.1 Optical Properties of Nobel Metals . . . 5
- 2.2 Propagating Surface Plasmon Polaritons . . . 6
- 2.3 Excitation of Surface Plasmon Polaritons . . . 8

3 Fabrication Techniques **11**
- 3.1 Electron Beam Lithography . . . 11
  - 3.1.1 Exposure . . . 12
  - 3.1.2 Spin Coating . . . 14
  - 3.1.3 Positive Tone Lithography . . . 15
  - 3.1.4 Negative Tone Lithography . . . 15
  - 3.1.5 Two-Step Lithography . . . 16
  - 3.1.6 Auxiliary Structures . . . 16
- 3.2 Metal Deposition . . . 17
- 3.3 Focused Ion Beam Milling . . . 18
- 3.4 Incoupling Gratings for DLSPP Waveguides . . . 19
  - 3.4.1 FIB Milling . . . 21
  - 3.4.2 Electron Beam Lithography . . . 22

4 Near-Field Microscopy **27**
- 4.1 Atomic Force Microscopy . . . 28
- 4.2 Scattering-Type Scanning Near-Field Optical Microscopy . . . 31
  - 4.2.1 Setup . . . 32
  - 4.2.2 Laser Source and Detectors . . . 35
  - 4.2.3 Extraction of the Near-Field Signal . . . 35
  - 4.2.4 Tip-Sample Interaction . . . 45
- 4.3 Measurement Geometry Considerations . . . 50
- 4.4 Measurement Corrections . . . 53
  - 4.4.1 Sync-Correction . . . 54
  - 4.4.2 Pixel Scaling . . . 54

4.4.3 Height Leveling . . . 54
4.4.4 Amplitude and Phase Drift Correction . . . 55
4.4.5 Gaussian Filtering . . . 55
4.4.6 Calculating the Real Part Data . . . 56

5 Chiral Couplers **57**
5.1 Photonic Spin-Hall Effect . . . 58
5.2 Design . . . 60
5.3 Simulation . . . 60
5.4 Fabrication . . . 63
5.5 Measurement . . . 63
5.6 Directionality . . . 67
5.7 Conclusion . . . 69
5.8 Outlook . . . 70

6 Single DLSPP Waveguides **73**
6.1 Single Waveguides . . . 74
6.1.1 Simulations . . . 74
6.1.2 Fabrication . . . 77
6.1.3 Measurement . . . 82
6.1.4 Discussion of the Phase Shift . . . 86
6.1.5 Indication of the In-Plane Component . . . 90
6.2 Waveguide with Reflection Grating . . . 92
6.2.1 Simulation . . . 93
6.2.2 Fabrication . . . 94
6.2.3 Measurement . . . 96
6.3 Conclusion . . . 100

7 Coupled DLSPP Waveguides **101**
7.1 Two Coupled Waveguides . . . 102
7.1.1 Theory . . . 102
7.1.2 Fabrication . . . 103
7.1.3 Measurements . . . 104
7.2 Coupled DLSPP Waveguide Arrays . . . 112
7.2.1 Coupled waveguides as quantum simulators . . . 112
7.2.2 Su-Schrieffer-Heeger Model . . . 114
7.2.3 Plasmonic Implementation . . . 120
7.2.4 Fabrication . . . 121
7.2.5 Measurements . . . 121
7.3 Conclusion . . . 126

8 Conclusion **127**

A Additional Information **131**
A.1 s-SNOM Calculations . . . 131
A.1.1 Lock-in Detection . . . 131
A.1.2 Polarization Projection in s-SNOM Setup . . . 132
A.2 DLSPP Waveguides . . . 136
A.2.1 Single DLSPP Waveguide Mode Simulations . . . 136
A.2.2 PDM for the Adaptation of the Simulations . . . 140
A.2.3 Three Dimensional Field Mapping . . . 143
A.2.4 Fabrication Problems . . . 143
A.2.5 Fabrication of coupled waveguides with positive tone EBL . . . 148
A.3 Silver Crystals . . . 150
A.3.1 Growth Recipe . . . 150
A.3.2 Fabrication of Nanogroove Arrays . . . 151
A.4 Incorporated Software . . . 154
A.4.1 Elphy/Raith Software . . . 154
A.4.2 SNOM/AFM Analysis Software . . . 155
A.4.3 Comsol Software . . . 156

Bibliography **157**

List of Figures **169**

Glossary **173**

CHAPTER 1

Introduction

Light plays a central role in modern information technology. Optical communication is the backbone of the internet and telecommunication networks, enabling high-speed data transfer over long distances. Compared to traditional electrical signal transmission using copper wires, optical communication offers several advantages, such as faster speeds, lower signal attenuation and significantly higher bandwidth. However, data processing and computing are still predominantly performed using electronic components.

Electronic processors have seen tremendous advancements over the past decades, following Moore's law, which predicts the doubling of transistor density on integrated circuits approximately every two years. This trend has led to exponential increases in computing power and efficiency. But as transistor sizes approach the atomic scale, fundamental physical limits arise, that challenge further miniaturization and performance improvements. Especially heat dissipation [1] and quantum effects like tunneling [2] become significant obstacles. Therefore, alternative computing paradigms are being explored to overcome these limitations.

Photonic computing, which utilizes coupling of light guiding structures for data processing, has emerged as a promising alternative to electronic computing [3–6]. Photonic components can be integrated on a chip similar to electronic components, allowing for high-speed data processing. Photonic integrated circuits enable non-binary computation and analog simulation paradigms. Combined with their intrinsically high bandwidth and low energy dissipation, these properties make photonic platforms particularly well suited for the rapid execution of non classical operations and for the efficient simulation of neuronal networks [7, 8]. An example are the recently developed photonic chips from *Q.Ant* and *LIGHTMATTER*. However, such photonic chips will likely not replace electronic computing completely [9], but hybrid photo-electronic systems may improve certain aspects [10].

The footprint of the individual structures is a central aspect for the fabrication of chips with a high computation power. The miniaturization of optical components is fundamentally limited by the diffraction limit, which restricts the confinement of light to dimensions smaller than approximately half its wavelength [11, 12]. Over the past decades, the field of plas-

monics has emerged as a promising platform to shrink the size of photonic components [13, 14], by exploiting plasmon polaritons, which arise from the coupling of electromagnetic waves to collective electron density oscillations. This allows for subwavelength confinement and manipulation of light [15]. The high confinement and associated coupling strength are favorable for coupling to plasmonic structures, but also lead to substantially higher losses in metal based plasmonic systems. The losses limit plasmonic systems to applications where a high coupling strength outweighs the losses. Plasmonic systems can for instance serve as a versatile platform for exploring fundamental light-matter interactions at the nanoscale and are used for implementations in sensing [16, 17], nonlinear optics [18, 19], photonic circuitry [20] and quantum optics [21, 22]. One particularly interesting use case are coupled plasmonic waveguides, which can serve as a platform to simulate condensed matter phenomena and quantum mechanical models [23–26]. The basis for the simulations is the mathematical equivalence of coupled waveguides with a one dimensional tight-binding lattice.

While plasmonics is a well-established field, there are still many open questions and ongoing research efforts aimed at addressing challenges such as losses [27], fabrication techniques [28], and characterization methods [29, 30]. Characterizing plasmonic waveguides is particularly challenging due to the diffraction limit of conventional optical microscopy [11, 12]. Additionally, plasmonic modes often exhibit intricate field distributions, thus requiring the full information of the electromagnetic field components and not just the intensity.

To overcome these limitations, a powerful and versatile tool is scattering-type scanning near-field optical microscope (s-SNOM). The s-SNOM technique utilizes a sharp tip to locally probe and scatter the evanescent near-fields above the sample surface [31, 32]. This technique allows for imaging and probing of optical fields at the nanoscale, well below the diffraction limit, providing insights into the behavior of plasmonic waveguides and their interactions with light.

In this thesis, we will explore various aspects of plasmonic waveguides, including their fabrication and characterization using s-SNOM. We start with a short introduction to surface plasmon polaritons (SPPs) and their excitation, followed by the fabrication techniques used to create nanostructures and plasmonic waveguides, highlighting the challenges involved in the process. The following chapter adresses near-field microscopy, explaining the working principles and experimental considerations. These tools are applied in the following three chapters, where we investigate three different plasmonic waveguide systems.

The first waveguide system we investigate are metal strip waveguides with chiral couplers. The photonic spin Hall effect is used here to achieve directional excitation of SPPs based on the polarization state of the incident light [33–36]. This effect is a manifestation of spin-orbit coupling in photonics, where the spin angular momentum of light (its polarization) is coupled to the orbital angular momentum of the waveguide mode propagating on a curved trajectory. This allows for spin selective routing of light in plasmonic circuits, which has potential applications in integrated photonic devices and quantum information processing. Measuring the directionality of the coupling allows to sense the polarization state of the illuminating light.

The second waveguide system we study are dielectric-loaded surface plasmon polariton (DLSPP) waveguides. These waveguides consist of a dielectric ridge on top of a metal film, which supports SPPs that are confined to the dielectric structure. We first investigate the properties of isolated DLSPP waveguides, including their in-plane electric field component, which is often neglected in literature [37–39].

We extend our study to the third system, coupled DLSPP waveguides. Here, we explore the coupling dynamics between adjacent waveguides, described mathematically by the coupled mode theory (CMT). The CMT equation for an array of coupled waveguides is mathematically equivalent to the tight-binding Hamiltonian used to describe electrons in a one-dimensional lattice. This analogy allows us to explore phenomena known from solid state physics in a plasmonic waveguide platform. In particular, we implement the Su-Schrieffer-Heeger (SSH) model, which describes a one-dimensional chain of atoms with alternating strong and weak bonds. This model exhibits topological edge states that are robust against static and dynamic perturbations [40, 41]. By fabricating a plasmonic waveguide array that mimics the SSH model, we can directly observe these topological edge states using s-SNOM. This study not only provides insights into topological photonics but also demonstrates the versatility of plasmonic waveguides as a platform for exploring fundamental concepts in condensed matter physics.

CHAPTER 2

Surface Plasmon Polaritons and Plasmonic Waveguides

This chapter provides a brief introduction to propagating surface plasmon polaritons (SPPs) and their excitation mechanisms, which form the basis for the plasmonic waveguide systems investigated in this thesis.

SPPs arise from the strong coupling between electromagnetic waves and collective oscillations of free electrons at metal-dielectric interfaces. These bosonic quasiparticles exhibit unique properties, such as subwavelength confinement and field enhancement. They can be categorized into different types, such as localized surface plasmon resonances and propagating surface modes. In this thesis, we focus on propagating SPPs.

2.1 Optical Properties of Nobel Metals

The optical properties of the metal and dielectric materials play a crucial role in determining the behavior of SPPs. Metals are characterized by a negative real part of their permittivity at optical frequencies, which is essential for supporting SPPs. The dielectric material, on the other hand, has a positive permittivity. In this thesis, we rely on gold for the realization of SPP waveguides. Gold is a noble metal with a high density of free electrons, which makes it an excellent conductor of electricity and supports strong plasmonic resonances in the visible and near-infrared spectral range. Additionally, gold is chemically stable and resistant to oxidation, making it suitable for various applications. The optical properties of gold can be described using the Drude model, which treats the conduction electrons as a free electron gas. According to the Drude model, the dielectric function of gold can be approximated as [42]

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_\mathrm{p}^2}{\omega^2 + i\gamma\omega}, \tag{2.1}$$

where ε_∞ is a high frequency correction accounting for a background polarizability caused by the ion cores. ω_p is the plasma frequency, which describes the natural frequency of oscillation of the free electron gas, and γ is the damping constant, which accounts for losses due to electron scattering. A comparison of the dielectric function of gold calculated with the Drude model and experimental data from Johnson and Christy [43] is shown in Fig. 2.1 (a).

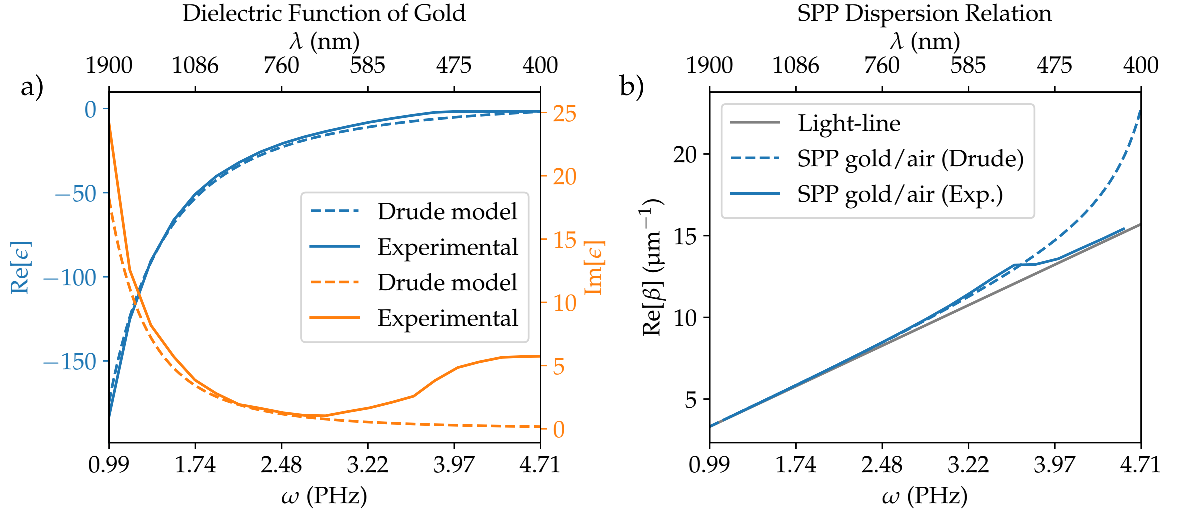


Figure 2.1: (a) Dielectric function of gold of the experimental data from Johnson and Christy [43] and a fit of the Drude model. The fit parameters used for the Drude model are $\varepsilon_\infty = 6.21$, $\omega_\mathrm{p} = 13.4\,\mathrm{PHz}$ and $\gamma = 0.1\,\mathrm{PHz}$, taken from [44]. The deviations of the experimental data of the imaginary part for the high frequency range stem from interband transitions in gold [42]. (b) Dispersion relation of SPPs at a gold/air interface calculated using the Drude model and experimental data from Johnson and Christy [43]. The light-line in air is shown for reference.

2.2 Propagating Surface Plasmon Polaritons

Propagating SPPs are surface waves that propagate along metal-dielectric interfaces and are exponentially confined perpendicular to the interface. Just like plane waves in free space, they can be characterized by their dispersion relation, which describes the relation between their frequency and wavevector. The derivation of the dispersion relation of SPPs is omitted here for brevity, but can be found in standard textbooks on plasmonics [42]. The dispersion relation of SPPs can be expressed as [42]

$$\beta_\mathrm{SPP}(\omega) = k_0(\omega)\sqrt{\frac{\varepsilon_1(\omega)\varepsilon_2(\omega)}{\varepsilon_1(\omega)+\varepsilon_2(\omega)}}, \tag{2.2}$$

where β_SPP is the complex-valued SPPs wavevector, k_0 is the free-space wavevector, and ε_1 and ε_2 are the permittivities of the metal and dielectric, respectively. This relation is

displayed in Fig. 2.1 (b) and shows that the wavevector of SPPs is always larger than that of light in the dielectric, which is why direct excitation from free space is in general not possible. The excitation mechanisms are explained in the next section. One can define two seperate quantities from β_{SPP}. The SPP wavelength

$$\lambda_{\text{SPP}} = \frac{2\pi}{\text{Re}(\beta_{\text{SPP}})}, \tag{2.3}$$

and the propagation length

$$L_{\text{SPP}} = \frac{1}{2\text{Im}(\beta_{\text{SPP}})}, \tag{2.4}$$

which describes the distance over which the intensity of the SPPs decays to $1/e$ of its initial value. Both quantities are commonly used to characterize SPPs.

The fields associated with SPPs are transverse magnetic (TM), meaning that the magnetic field is perpendicular to the direction of propagation, while the electric field has both perpendicular and parallel components to the interface [42]. The electric field vector rotates in the x-z-plane, perpendicular to the magnetic field [1], leading to a transverse spin that is locked to the propagation direction [45–47]. This spin-momentum locking gives rise to interesting phenomena such as the photonic spin Hall effect, where the direction of propagation of SPPs can be controlled by the polarization state of the incident light [48]. This will become important for Chapter 5. A sketch of a propagating SPP is shown in Fig. 2.2.

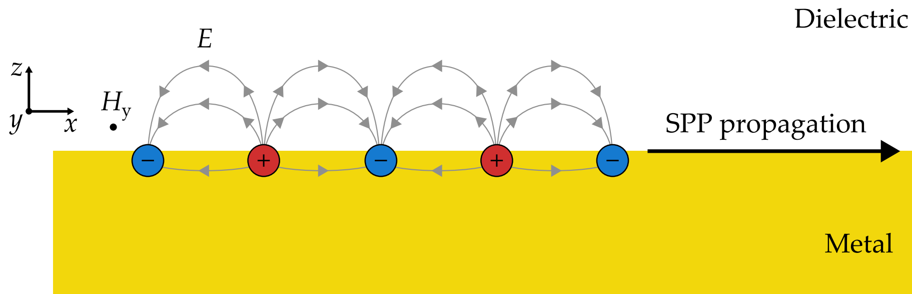


Figure 2.2: Snapshot sketch of a propagating SPP on a metal/dielectric interface at. The blue and red circles represent the chare density oscillations in the metal, while the arrows indicate the electric field vectors. The magnetic field is perpendicular to the wave propagation and the electric field vectors, and thus perpendicular to the plane of the figure.

The propagation of SPPs can be controlled and directed using plasmonic waveguides. A waveguide can be realized by a local variation in the dielectric environment or the metal geometry, which modifies the dispersion relation of the SPPs and allows for confinement and guidance of the surface waves along a specific path. Over the years, various types

[1] The rotation of the electric field vector can be understood by following the vector orientation in a point above the interface, while moving the observation point against the propagation direction.

of plasmonic waveguides have been developed. In this thesis we employ two types of plasmonic waveguides, strip metallic waveguides and dielectric-loaded surface plasmon polariton (DLSPP) waveguides, as sketched in Fig. 2.3. Depending on the size of the waveguide and the wavelength, it can support one or multiple modes. All measurements in this thesis are performed using single mode waveguides.

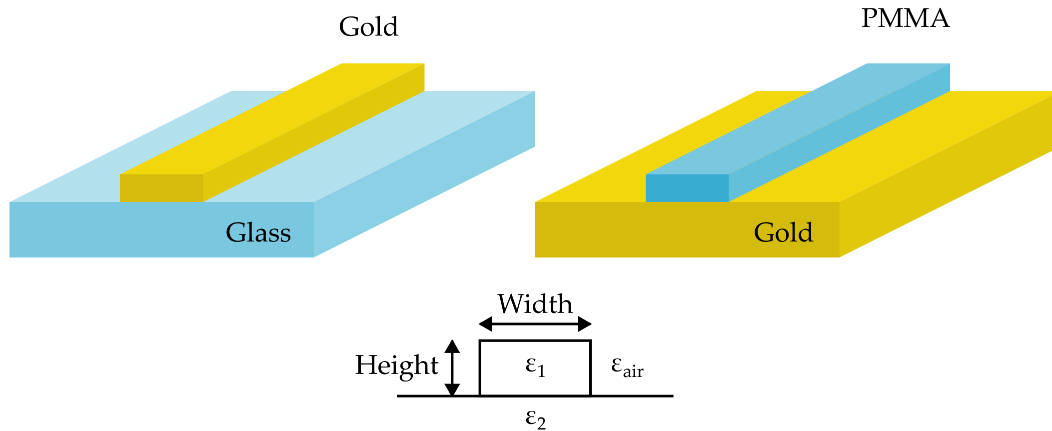


Figure 2.3: Sketch of the two waveguide types used in this thesis. Depicted on the left is a strip metal waveguide on a glass substrate surrounded by air, and on the right a DLSPP waveguide consisting of a thin dielectric ridge on top of a gold film surrounded by air.

2.3 Excitation of Surface Plasmon Polaritons

As discussed in the previous section, the dispersion relation of SPPs dictates that their wavevector is larger than that of light in the dielectric. Therefore, SPPs cannot be excited directly by light in free space due to the momentum mismatch. To overcome this challenge, evanescents fields are required to provide the additional momentum needed for excitation. Various excitation schemes have been developed to achieve this momentum matching, including prism coupling (Kretschmann and Otto configurations), grating coupling or near-field coupling using a sharp tip or defect. In principle even a sharp subwavelength feature, like the end of a waveguide or a nanoparticle, can provide the necessary momentum components to excite SPPs. When incoming light interacts with such a feature, it scatters and generates evanescent field components that can match the momentum of the SPPs, enabling their excitation at the metal-dielectric interface. In this thesis, we primarily utilize grating couplers to couple light into SPP waveguides, because it is a reliable method to excite waveguide modes. But also the excitation from sharp edges and from the probing tip in scattering-type scanning near-field optical microscope (s-SNOM) measurements are employed.

A grating coupler consists of a periodic structure, such as a series of grooves or ridges, fabricated on or in the metal surface. When light is incident on the grating at a specific angle, the periodicity of the grating provides additional momentum components through diffraction. If the grating period is choosen correctly, the superposition of all individual diffractions leads to constructive interference along the direction of the SPP propagation. The momentum matching condition for grating coupling can be expressed as [42]

$$\beta_{\text{SPP}} = k_0 \sin(\theta) + mg\,, \tag{2.5}$$

where β_{SPP} is the wavevector of the SPPs, k_0 is the free-space wavevector, θ is the angle of incidence, m is an integer representing the diffraction order and $g = \frac{2\pi}{a}$ is the reciprocal grating vector with the grating period a. Since the grating structures in this thesis are always excited under normal incidence, the grating period has to roughly match the SPP wavelength to achieve efficient coupling. The use of a focused Gaussian beam for excitation introduces a range of incident angles, which relaxes the strict momentum matching condition. The coupling efficiency depends on various factors, including the grating geometry (period, depth, duty cycle), the incident light polarization [2] and overlap with the SPP mode, and the material properties of the metal and dielectric. A sketch of the grating coupling scheme is shown in Fig. 2.4.

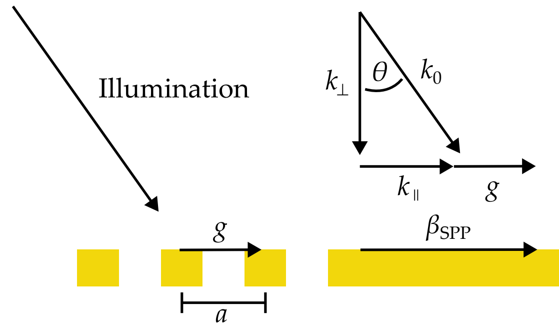


Figure 2.4: Sketch of grating coupling of free space light, incident under the angle θ, to a SPP mode.

[2] Only TM polarized light can excite SPPs.

CHAPTER 3

Fabrication Techniques

This thesis relies heavily on the fabrication of nanophotonic structures. In this context, the word "nano" means, that at least one of the dimensions of the structures is on the order of nanometers, resulting in sub-wavelength features.

In the following sections the main fabrication techniques used in this thesis are introduced. These are electron beam lithography (EBL), focused ion beam (FIB) milling and thermal evaporation. The individual fabrication recipes of the measured samples are described in the respective chapters. The grating couplers used in Chapter 6 and Chapter 7 are fabricated with the same process, so instead these will be described at the end of this chapter. Additional to the fabrication methods used in this thesis, we also discuss the chemical growth and FIB structuring of silver crystals, which have been fabricated in collaboration with Yuhao Zhang. The silver crystal fabrication is discussed in the appendix A.3, since they are not used for measurements in this thesis.

3.1 Electron Beam Lithography

EBL is a technique, where a focused electron beam is used to pattern an electron sensitive resist. Depending on its type, the resist is either strenghend or weakened, allowing for positive or negative tone lithography. In a positive tone procedure the exposed areas of the resist become soluble in the developer solution, while in a negative tone procedure the exposed areas become insoluble. In Fig. 3.1 a schematic of both processes is shown. Some resists, such as poly-methyl-methacrylate (PMMA), can be used for both procedures, depending on the electron dose used [1]. This is because, in a certrain electron dose range the molecules in the resist are broken down and start to form crosslinks, making the resist more resilient to solvents [49, 50]. Additionally, negative tone resists can be used for gray scale lithography, since a higher electron dose leads to a larger volume of properly exposed resist.

[1] The electron dose corresponds to the deposited energy.

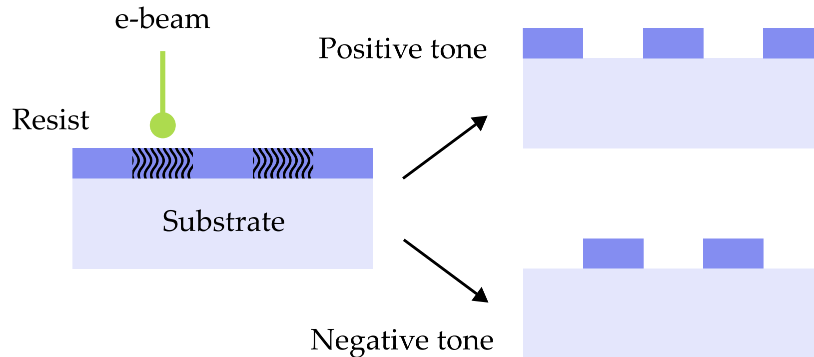


Figure 3.1: Schematic illustration of positive and negative tone lithography. In positive tone lithography the exposed areas of the resist become soluble in the developer solution, while in negative tone lithography the exposed areas become insoluble.

3.1.1 Exposure

The deposited dose is the central parameter in EBL. It depends on many factors, such as the electron energy, the beam current, the beam size, the dwell time and the pattern that is written. In the following, these factors are explained in more detail.

The electron energy is defined by the acceleration voltage in the scanning electron microscope (SEM). Typically, the electrons are accelerated to energies between 10 to 20 keV. The electron energy is important for the deposited dose, as the electrons do not lose their energy in the resist linearly. For example the deposited electron dose increases by a factor of approximately 4-5, when changing the acceleration voltage from 20 keV to 10 keV. The exact energy loss of low energy electrons in resists such as PMMA is difficult to calculate, but can be modeled using Monte Carlo simulations [51, 52] if need be. For sufficently low acceleration voltages and thick enough resist layers the electrons can be stopped in the resist, allowing to write suspended structrues in the resist [53]. Additionally, the deposited dose varies within the resist layer, as the electrons can scatter in the resist and substrate. The electron scattering in the resist and substrate leads to an effect called the proximity effect [54]. This means the effective dose can strongly depend on surrounding already exposed regions of the resist. This effect is not time independent. It is stronger in areas that have already been exposed above threshold and areas that have not been exposed above threshold show a partial healing or distribution effect. When exposing structures multiple times with reduced dose a higher total dose is required compared to exposing them directly above threshold. For example, exposing an area twice with half the dose, leads to different results as exposing the area once. When exposing areas close to each other, the proximity effect needs to be taken into account. This effect is even stronger for negative tone lithography, due to the higher overall electron dose. Structures that are written without a sufficient correction of the proximity effect often show different height, width and shape. For positive tone lithography this effect is less pronounced, and leads to widening and rounding of the structures.

The beam current and size are influenced by the aperture size in the SEM. A larger aperture leads to a larger current but also a larger beam size. Therefore, we typically use a small aperture of 10 to 30 μm. The beam current influences the deposited dose directly, as a higher current leads to a higher dose for the same dwell time.

The dwell time describes the time the electron beam spends on a single point during the exposure. It is scaled to the beam current to achieve a reproducible dose. A higher beam current allows to use a shorter dwell time, which can speed up the writing process. However, if the beam currents are too high the beam size increases, decreasing the resolution.

The pattern that is written is a 2d map of point exposures with the corresponding dwell time. It is adapted with a dose factor using dose tests for each new structure. The dose factor allows to scale the dose for different structures, while keeping all other parameters constant.

The EBL system in our group is controlled by the *Elphy Plus* interface from *Raith*. The interface allows to create or load design files, which can then be used to write the patterns into the resist. All design files used in this thesis are created with a custom python script detailed in the appendix A.4.1. This script allows to place geometric objects, such as rectangles and circles, within a writefield. This writefield is the total area that the electron beam can scan without moving the stage. The available writefield size depends on the magnification chosen in the SEM. For fine structures, usually a writefield size of 100 μm by 100 μm is used, corresponding to a magnification of 600 X. The scan resolution of the electron beam is discretized into pixels, which define the smallest step size of the beam. A smaller writefield and larger magnification allows to use a finer step size, which is favorable for high resolution structures. The design software allows to write the geometric objects as areas or lines. The beam patterning software translates the design files into a series of points, which the electron beam then exposes sequentially with an according dwell time. Therefore, writing the geometric objects as areas or lines does not matter in general, but especially for fine structures a more precise control over the writing process is desired. Thus, most structures in this thesis are written as lines. The respective dose per area is 1 400 $\mu C\,cm^{-2}$ and 1 500 $pC\,cm^{-1}$ per line. All following desing doses are defined relative to these doses by their dose factor.

In order to achieve a high quality of the written structures, stitching of connected areas within a writefield should be avoided [2]. The reason is that the stage is not perfectly stable and slight drifts over time can lead to misalignments within the structures. Therefore, all fine connected structures, like the chiral couplers in Chapter 5, should be written in one stroke.

Electrons are charged particles and as such susceptible to electric and magnetic fields. Therefore, the electron beam can be deflected by external fields. Additionally, the electrons collecting in the sample during exposure can lead to a charging of the sample, which can deflect the incoming electrons and lead to distorted structures. For this reason the sample has to be conductive. For isolating substrates, a conductive layer can be added on top of the resist. We use Electra 92 (*Allresist*) for this purpose, which is a thin layer of polyanilin that can

[2] Stitching of areas across multiple writefields is even harder, since the stage motor movement has an uncertainty of several microns.

be removed easily by rinsing the sample in water. It has however also some disadvantages. It has to be applied for writing and imaging, and it can burn for high electron doses. This renders it unsuitable for negative tone lithography with PMMA, as the required electron dose is too high.

3.1.2 Spin Coating

Spin coating is a commonly used technique, were a fast rotation of the substrate is used to evenly apply a thin layer of a resist. The resists are dissolved in a solvent, which helps to prevent degradation of the resist and reduces the final film thickness. The process usually includes two steps, firstly the slow distribution phase typically at 500 rpm for 5 s, where the aim is to distribute the resist evenly accross the substrate, and secondly the fast spinning phase around 4000 rpm for 90 s, where excess resist and solvent are removed. The final thickness of the resist layer depends on the spin speed and time, as well as the viscosity and concentration of the resist/solvent mixture. After spin coating, the sample should be baked on a hot plate for several minutes to evaporate the remaining solvent. This step is also crucial for the later use in a vacuum environment, as slowly evaporating solvent can lead to a pressure increase in the SEM chamber.

If thicknesses are required which are not achievable with a single layer of resist, the process can be repeated. Here, one should be careful since the application of subsequent layers could dissolve the previous ones. It was found that adding the resist to the sample on the last second of the slow distribution phase leads to no noticable dissolution of the previous layer. The spin curves can be found on the respective product information sheets of the resist manufacturers. In Table 3.1 some example achievable thicknesses for PMMA resists with different concentrations are shown. The thicknesses are measured with the atomic force microscopy (AFM), but the final thicknesses may vary by ± 10 nm, depending on process details and the substrate. The values given here are just ment as a guideline, the actual thicknesses vary and require a regular adjustment of the spin speed.

Spin Speed (rpm)	Thickness for 2 % (nm)	Thickness for 4.5 % (nm)
1000	120	410
2000	90	320
4000	70	230
6000	60	190

Table 3.1: Expected layer thicknesses of two different concentrations of PMMA (in anisol) achieved with different spin speeds, taken from the product information of *Allresist*.

3.1.3 Positive Tone Lithography

Positive tone lithography is the most common type of electron beam lithography. It is used to create a mask, which can then be used to either etch the substrate or to deposit a material on top of the substrate. In this thesis we utilize solely the latter. We use mainly PMMA (AR-P 672, 950k, *Allresist*) in two different concentrations. For the fabrication of evaporation masks we use a default concentration of 4.5%. For the fabrication of dielectric-loaded surface plasmon polariton waveguides (DLSPPWs) with the positive tone technique we use a lower concentration of 2%, which allows to achieve thinner resist layers. The resist, solved in anisol, is spin coated onto the substrate and then baked on a hot plate for 120 s at 170 °C. After exposure with the electron beam, the sample is developed in a methyl isobutyl ketone (MIBK) mixture (65.7% isopropanol, 32.8% MIBK, 1.5% butanone) for 60 s, then rinsed in isopropanol and dried with nitrogen. The developement of resists is quite complicated on a molecular level [55], but for PMMA a ratio of MIBK to isopropanol of 1:3 is usually recommended [56]. After the metal deposition (Section 3.2) the sample is placed in a solvent to remove the resist and with it the metal on top. For this one can either use acetone, which is more agressive and faster but should be avoided for fine structrues, or n-methyl-2-pyrrolidone (NMP) heated to 80 °C. The latter takes longer, around two hours, but is typically preffered for more delicate structures. After the resist is removed, the sample is rinsed in isopropanol and dried with nitrogen. The electron dose factor for PMMA is typically around 0.25 when exposing a large area with lines spaced by 5 nm or 0.46 when exposing as an area.

3.1.4 Negative Tone Lithography

In negative tone lithography the sample is spin coated with a negative tone resist, exposed with the electron beam and then developed to remove the unexposed areas. The exposed resist is then the material that remains after the development step. This has the advantage that the height of the structures can be controlled by the electron dose, in case of a gray scale resist. But the material of the structures is limited to the available resists. Negative tone lithography is utilized in this thesis to create the DLSPPWs in Chapter 6 and Chapter 7. The required dose for the crosslinking of PMMA is significatly higher than for positive tone lightography. In order to reduce the writing time, an acceleration voltage of 10 kV is used. A variation of the acceleration voltage, shown in the appendix in Fig. A.15, showed that for an even lower acceleration voltage the electrons can not penetrate the resist sufficiently and for a much higher voltage the dose needs to be increased significantly.

After the exposure the sample is developed in acetone for 60 s, then rinsed in isopropanol and dried with nitrogen. The choice of the solvent is not critical, as the crosslinked resist becomes harder [57] and extremely resilient to solvents, even highly concentrated nitric acid did not show any effect. The solvent should however be strong enough to remove the unexposed resist without leaving too much residue. The dose factor for PMMA negative tone lithography is typically around 1, when exposing a large area with lines spaced by 5 nm or 1.8 when exposing as an area.

During this thesis we also encountered numerous issues with PMMA negative tone lithography, such as bubbles forming in the resist, variations in dose sensitivity and grainy structures. Overall, structures written with negative tone EBL often show less sharp features compared to positive tone lithography. Thus, at some point the fabrication process was changed to PMMA positive tone lithography, but with an inverted design. While this adds some complexity to the desing and removes the freedom of gray scale lightography, the process leads more reliably to high quality structures.

3.1.5 Two-Step Lithography

In this thesis, we make extensive use of two-step lithography, which is a technique that allows to align structures written in different lithography steps with high precision. This enables the creation of complex designs, unachievable with one lightography step alone. For example the DLSPPW structures in Chapter 6 are created with two-step lithography. In this case the first step is used to create the incoupling grating and alignment markers, while the second step is used to create the DLSPPW itself. Here, it is important to align the two steps with high precision, as the DLSPPW has to be placed on top of the incoupling grating to ensure an efficient excitation. This is done by combining large and small alignment markers. The large markers, typically 100 µm in size, are used to align the two steps with a precision of a few µm and are essential to find the small alignment markers within the writefield. The small markers, typically 15 µm in size, are used to align the writefield of the second step with the first step with a precision of less than 200 nm. An example optical image of a small alignment markers is shown in Fig. 3.2 (left panel).

3.1.6 Auxiliary Structures

In addition to the main structures, auxiliary structures are often created as well. These can be text to properly identify the written structures, markers for alignment in a two step process and other structures aimed at helping the measurement process in the scattering-type scanning near-field optical microscope (s-SNOM). For the alignment of the reflection arm in the s-SNOM setup, a highly reflective reference area is needed. For the alignment of the lower parabolic mirror in the transmission mode, small nano-antenna structures like gold discs can be helpful to find the right focus and position. Especially for the adjustment of the interferometer a known reference structure is very useful. Here, an array of gold discs with varying radius from 200 nm to 500 nm in steps of 50 nm is used. The discs have a height of 40 nm and are arranged in a grid with a period of 2 µm. An example image of the disc array is shown in Fig. 3.2 (right panel).

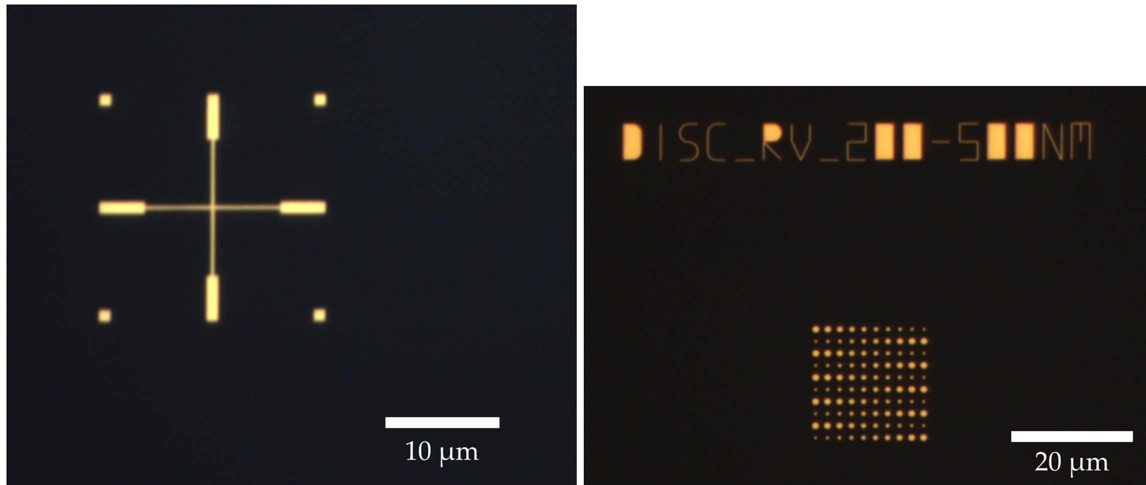


Figure 3.2: (left panel) Example of a small alignment marker (gold on glass), which is used to align the writefield of the second lithography step with the first step. The marker is 15 μm in size and consists of a fine cross with larger patches at the ends for visibility. The corners have a small square to ensure the marker is easy to find during alignment, even if it is shifted diagonally. The marker allows to align the writefield with a precision of less than 200 nm. (right panel) A gold disc array with different disc radii is shown, which is used in s-SNOM measurements as a reference structure.

3.2 Metal Deposition

Metal films can be deposited with a variety of techniques, such as thermal evaporation, electron beam evaporation, sputtering, or even electroplating [58]. In this thesis we rely exclusively on thermal evaporation with a *Leybold Univex 250* system, because it reliably results in high quality metal films. In this system the substrate is placed in a vacuum chamber, which is evacuated to a pressure of around 1×10^{-6} to 5×10^{-7} mbar. The metal is placed in a tungsten boat clamped between two high current electrodes, which are connected to a power supply. During the evaporation process, a current of approximately 80 A is passed through the boat, which heats it and the metal inside up to high temperatures, leading to a slow evaporation of the metal. The evaporated metal then condenses on the substrate, forming a thin film. The evaporation is controlled by a set of shutters and the evaporation rate is monitored by a quartz crystal (750-1000-G10, *Kurt J. Lesker*), which measures the evaporation rate, via the change in frequency caused by the added mass of the evaporated metal. For gold films we use a constant rate of $2\,\text{Å}\,\text{s}^{-1}$ and adjust the evaporation time according to the desired thickness. To increase the adhesion of the metal film to the substrate a few nano meters of chromium are evaporated prior to the gold film.

3.3 Focused Ion Beam Milling

In the context of this thesis, the FIB is mainly used to mill grating couplers into metal substrates and for the creation of nanogroove arrays in silver crystals, as described in Appendix A.3.2. The FIB is a system in which a focused beam of ions is used to mill structures into a substrate. Gallium ions are often used as the source. The extracted ions are accelerated with voltages of around 30 kV. Some more modern systems also use other sources, such as helium, which allows to achieve a higher resolution. Our group has access to a *Zeiss 1540XB Crossbeam* system, which is stationed at the *Max-Planck-Institute for neurobiology of behavior* (*caesar*) in Bonn. This system features both, a FIB and a SEM column, which allows to use the SEM for imaging and alignment, while the FIB is used for milling. Similarly to the SEM, the FIB utilizes different apertures to control the ion current and the resolution. For nanofabrication, a small aperture of 10 to 30 µm is used, resulting in a resolution in the range of tens of nanometer. The larger the aperture, the larger the ion current and the milling rate. During this thesis only the 30 µm aperture was available, which leads to currents of around 10 to 100 pA, and a minimal milling width of around 80 nm. Similar to the SEM, the FIB can also be controlled with the *Elphy Plus* software interface from *Raith*. Thus, the same design files as for the EBL can be used, but with a different dose factor. The individual lines, which make up an area, are separated by a distance of 5 nm, as this works well in most cases. The dose has to be recalibrated each time the beam current is changed. Similar to the EBL, the designs are created by filling areas with lines. Whilst the ordering of the lines plays only a minor role in EBL, it is crucial for the FIB. Here, the ion bombardment leads to a weakening of the material and its structure making it easier to mill over time. For example, if one mills a square made of lines from the inside out, the resulting square hole will be deeper on the outside than in the middle. This effect is ususally compensated by milling the same area multiple times with a reduced dose. This so called looping is sometimes repeated thousands of times. However, in our case this also leads to a significant decrease in the feature sharpness, so it is in general not used. The milling also leads to a redeposition of the material, which can lead to a pattern when milling larger areas, rounder features and even buildup of material on edges. To get the best results one should mill from the inside out, such that the important features are milled last, while the depth differences can be compensated by a dose gradient, if needed. The discussed behavior can for instance be observed in the images of an alignment marker in Fig. 3.3 and the later discussed image of a dosetest milled into a silver crystal in Fig. A.21. In the image of the alignment marker one can see that the middle of the milled area is not as deep as the edges, this is caused by the milling from inside out. The material in the middle has not been weakened by previous milling passes and the milling on the outside leads to a redeposition of material in the middle. The same can be seen in the second image of the dosetest in Fig. A.21, where the center of the horizontal milled area is substantially deeper on the outside than in the middle. While this is hard to see due to the persepective the widening of the grooves is clearly visible. Due to the structural weakening around the grooves, the silver is milled much easier than in the unexposed areas between the grooves.

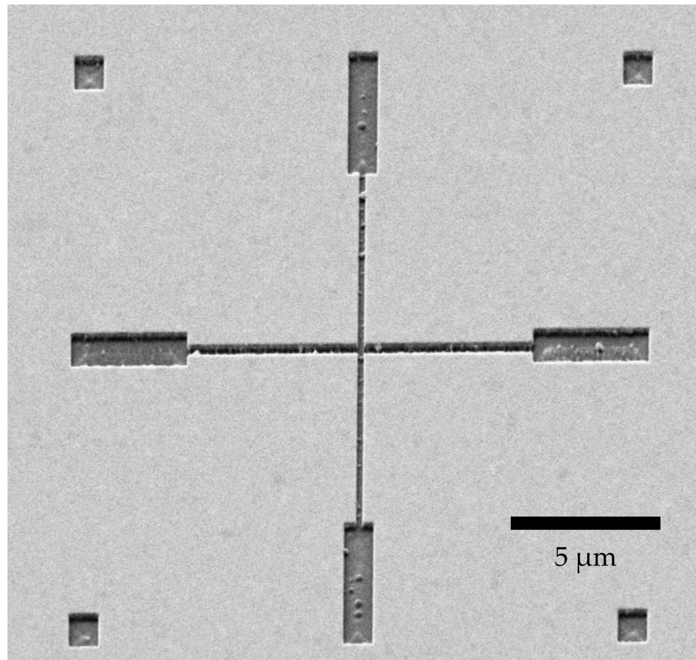


Figure 3.3: Example of a 20 μm alignment marker milled with the FIB into a gold film. The image is taken with the SEM inside the FIB system. The bright spots in the center of the milled areas are remains of the metal film caused by the milling procedure and redeposition of the milled material. The glass substrate below has a much darker shade than the thin metal film above.

The FIB is a versatile tool for nanofabrication, but it is also a time consuming and difficult to use technique. Especially the older device, which we have access to, suffers from strong current drifts over time and stage drifts, which can make it difficult to achieve consistent results. Additionally, the gallium ions can contaminate the sample, which can be an issue for certain applications.

3.4 Incoupling Gratings for DLSPP Waveguides

Incoupling gratings are necessary to excite the dielectric-loaded surface plasmon polariton (DLSPP) modes in the waveguides. As they will be used in two seperate projects, their fabrication is explained here and the fabrication of the accompanying waveguides in the respective chapters. The gratings are designed to match the momentum of the incident light to the momentum of the DLSPP mode, as explained in Section 2.3. The grating lines are oriented perpendicular to the waveguide axis and the polarization of the incident light is perpendicular to the grating lines, as sketched in Fig. 3.4. Samples used in the transmission mode s-SNOM setup are illuminated from the backside. Therefore, the incoupling gratings need to be created in the metal film below the waveguides. This requires a two step fabrication process, where first the grating in the metal film and then the waveguide itself

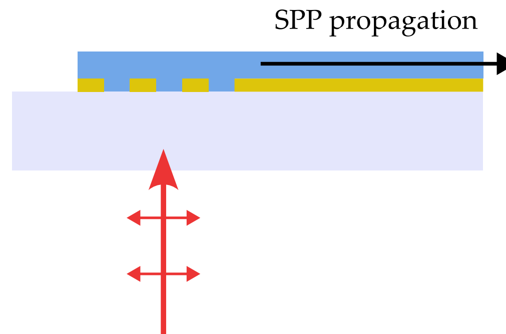


Figure 3.4: Sketch of the excitation of a DLSPPW mode with a grating coupler illuminated with normal incidence from the backside of the substrate.

is fabricated. Initially we used standard microscope glass coverslips (22 cm x 22 cm) as the substrate, because they allowed measurements in the leakage radiation microscopy (LRM) setup. Later they were switched for indium tin oxide (ITO) coated coverslips (1 cm x 1 cm, CEC500S, *Praezision Glas & Optik GmbH*) because they are smaller and more durable. Additionally, they are conductive, which makes the fabrication process easier because no extra conductive layer needs to be applied. The metal films are created via thermal evaporation of a layer with a thickness of 60 nm, as described in Section 3.2. Depending on the process, the metal film is either added before or after the patterning of the incoupling gratings. In this thesis, FIB milling and EBL have been used. The two methods are displayed in Fig. 3.5.

The advantage of FIB milling is that it is a direct write process, which means that the gratings can be created in the metal film without the need of a mask. This makes the process very flexible, since the design can be changed easily and the gratings are created after the evaporation of the metal film. The disadvantage of FIB milling is that the process is time consuming and not as reliable. An additional problem is that the second step needs to be aligned to the first step, which can be quite challenging because the FIB stage tends to drift much more than our SEM. There is, however, a big advantage of using the FIB milling process for DLSPPWs. Using the FIB only a small portion of the metal film is carved out, leaving most of the surface covered. As a result the illumination light is mostly blocked by the metal film, leading to a better contrast in LRM measurements.

The SEM on the other hand is readily available and the process is much faster. Samples created with the EBL process usually have much larger uncovered areas, as each covered area needs to be exposed and the creation of thin grating lines is more difficult.

3.4.1 FIB Milling

The general FIB milling process is described in Section 3.3. The gratings are created by milling parallel lines into a gold film on a glass substrate. The lines are spaced by 5 nm and the number of lines and dose is varied for the best possible result. The used parameters varied quite significantly over the different fabrication sessions, as the FIB current was not very stable. Often the dose had to be adjusted while writing a set of writefields. But typical parameters are an extractor current of 2 to 4 µA, a 20 pA current (measured on the faraday cup), a 30 µm aperture (the smallest available aperture) and a line dose of 1.6 nC cm^{-1}. The individual milled grating slots have a design length of 1 µm and a width of 200 nm at a line distance of 5 nm. This ensures sufficient looping of the beam to create a smooth grating but not too much as this would lead to rounder edges. As described in Section 3.3, we mill the outside edges last for the best possible edge quality. The actual width of the milled gratings is usually closer to 240 nm. The gratings are intentionally short and small to minimize the transmission of the laser light through the grating to make them more suitable for LRM measurements. Due to their size, short gratings can potentially lead to a lower coupling efficiency and decrease the error margin during alignment. A second version of these gratings with a cone shape to improve the couling efficiency and decrease unwanted excitation of free surface plasmon polaritons (SPPs) has also been tried. The stage drift of the FIB was too strong for the aligment of the second step and thus these gratings could not be used. An example image of both gratings is shown in Fig. 3.6. The image was taken with the SEM inside the FIB system. The SEM has a tilt of 54°, which is corrected by stretching the image, and allows to see into the milled structures. The polycrystalline nature of the gold film is visible as well, and can lead to uneven milling and the formation of small gold islands. Residual gold at the bottom of the grating would decrease its performance. To avoid this, the grating is overexposed leaving an imprint of the structure in the glass substrate.

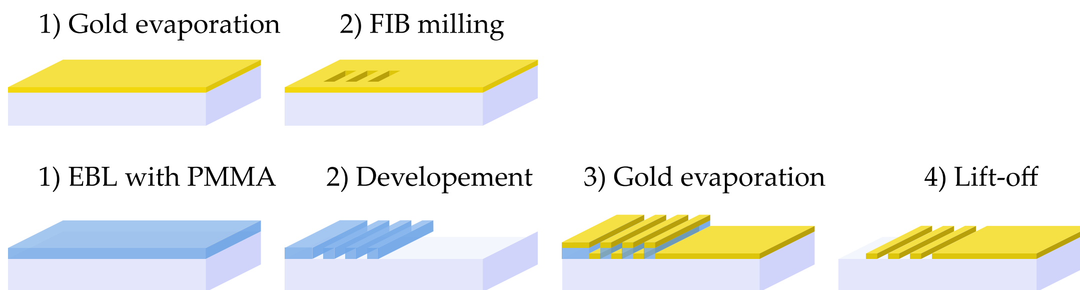


Figure 3.5: Sketch of the two employed fabrication methods of the grating couplers used for the excitation of DLSPPW modes.

As discussed before, the FIB milling process can be quite unreliable. An alternative process using EBL was therefore developed, which is described in the next section.

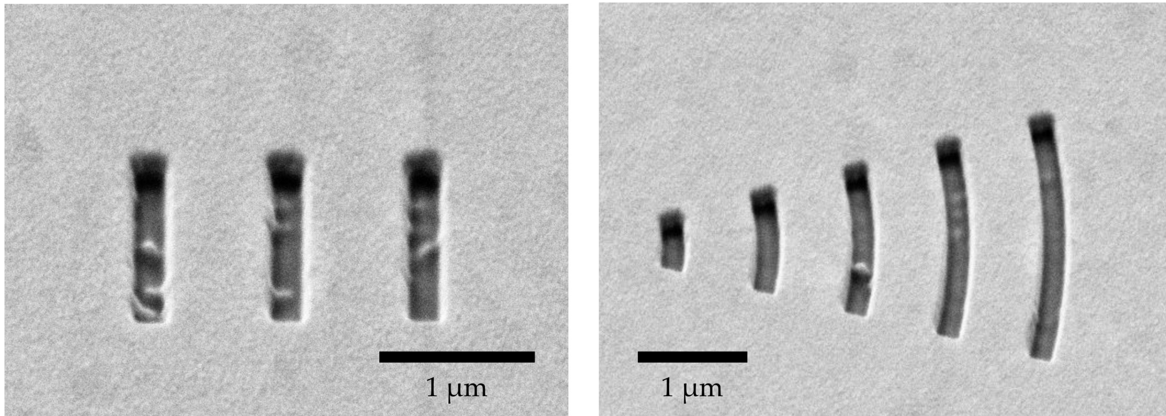


Figure 3.6: Example of two FIB milled gratings on a gold film. The grating lines are approximately 240 nm wide and have a periodicity of 880 nm. The image was taken with the SEM inside the FIB system. The left incoupler is an example of a typical coupler used for s-SNOM measurements. The right grating is an alternative design aimed at improving coupling efficiency, but it was not used since the vertical alignment precision was not sufficient.

3.4.2 Electron Beam Lithography

The creation of incoupling gratings with EBL can be done in different ways. Two approaches, which have been tried are based on negative and positive tone lithography. Negative tone lightography would seem like the obvious choice, because one only needs to expose the areas that should not be covered by the metal film. The processes of removing the exposed resist after the evaporation of the metal film can be quite tricky and unreliable. For positive tone lightography the design needs to be inverted and much more area needs to be exposed. But because of the sharper features, the lift-off process works more reliable. The dose values given in this section are given relative to the base dose defined in Section 3.1.

Negative tone lithography

Negative tone lithography was performed using the photoresist Medusa 82 (*Allresist*). This resist has a high sensitivity and can be removed with a solvent such as potassium hydroxide. After the lithography the developed structures are covered with a metal film, which acts as the substrate for the second step. However, in our case the removal of the Medusa 82 structures after evaporation did not work consistently enough. PMMA can be used alternatively to Medusa 82, but exposed PMMA can not be removed easily. Therefore, one can try to fabricate high and narrow gratings, which can cause discontinuities in the metal film, enabling coupling without the need to remove the exposed resist. We were not able to couple light into the waveguides in this way, thus we switched to positive tone lithography.

Positive tone lithography

Positive tone lithography with PMMA was implemented as an alternative to negative tone lithography and foucused ion beam milling. The main advantage are more flexibility, faster sample fabrication and better alignment precision. Due to the inverted design, not only the incoupler but also the areas the waveguides are written on in the second step need to be patterned. The design of the grating is shown in Fig. 3.7. Due to the large areas of substrate

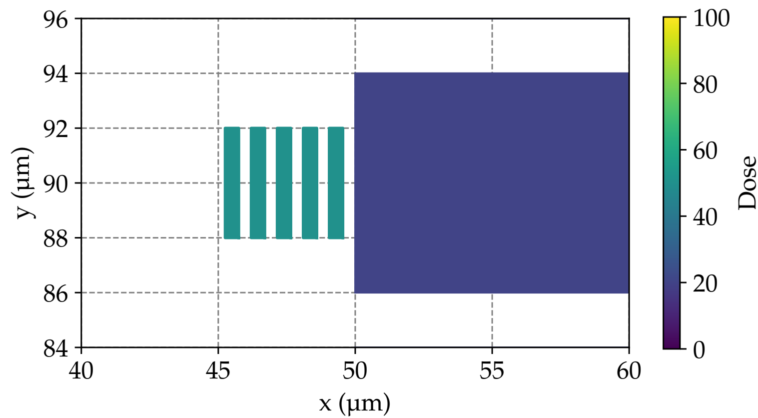


Figure 3.7: Design of the first step of the two step EBL process for the fabrication of DLSPPWs. The incoupling gratings are written as individual lines while the bigger areas are written as areas. The alignment markers and text are also included in this step.

without a metal film these samples can not be used in the LRM setup. The grating is made up of individual lines and the big area is written as an area. The used parameters were varied over time, but a typical set of parameters, which yield reliable incoupling gratings for the respective wavelengths, is shown in Table 3.2. The period is chosen to match a mode

Wavelength (nm)	Period (nm)	Design width (nm)	Dose factor
800	727	364	0.4
900	818	409	0.4
970	882	441	0.4
1400	1273	796	0.4
1500	1364	852	0.4
1600	1455	909	0.4
1700	1546	966	0.4

Table 3.2: Parameters for the incoupling gratings created with EBL for different wavelengths. The third column is the width of the exposed area, which is typically smaller than the actual width of the grating lines after liftoff. The fourth column contains the dose used for the individual grating lines, which are spaced 20 nm apart. The area in between the gratings and the waveguide is written as an area with a dose factor of 0.163.

index of 1.1. A reasonable approximation, as the mode index does not change much in the investigated parameter range. Additionally, plasmonic resonances are broad and are not very sensitive to a slight mismatch in the grating period. The gratings are then created like the default introduced in Section 3.1.3 for PMMA positive tone lithography. A dose test of incoupling gratings fabricated in this manner is shown in Fig. 3.8. The gratings on the left side are too thin and the ones on the right end start to overlap due to overexposure. In this case we chose the parameters from the grating in row 5 column 2, as indicated by the white dashed box. This grating is also imaged with the SEM in Fig. 3.9. A complete

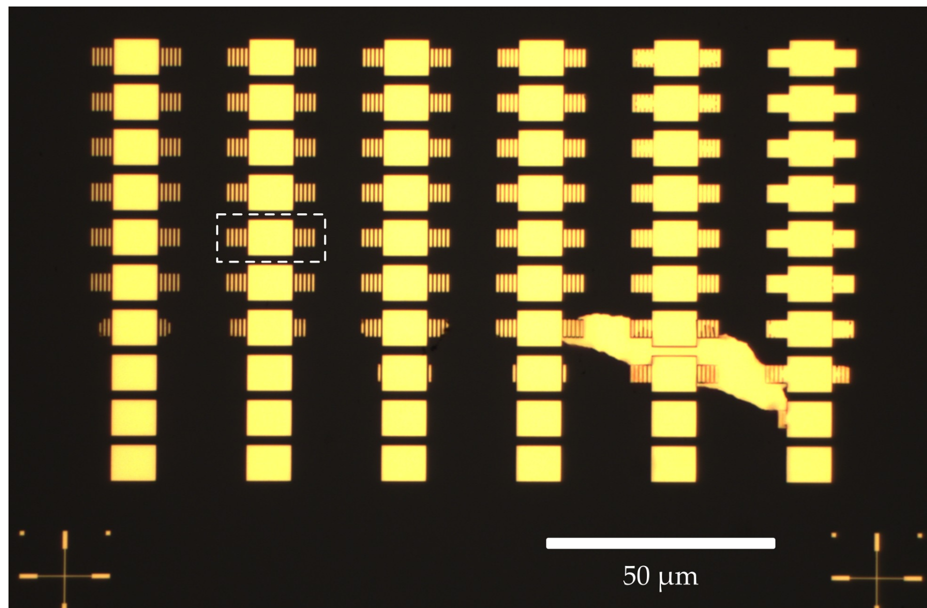


Figure 3.8: Example of a dose test of incoupling gratings for a wavelength of 970 nm written with PMMA positive tone and imaged with an optical microscope after the liftoff of the gold film. On the horizontal axis the width of the grating increases from left to right while on the vertical axis the dose of the grating is increased from bottom to top. The connecting area in between is exposed to emulate the influence of the proximity effect. The highlighted grating has a period of 890 nm, a width of 500 nm and a dose factor of 0.4.

writefield would look like in Fig. 3.10, where the area connecting the two incoupling gratings is increased to allow for a 100 µm long waveguide to be written on top in the second step.

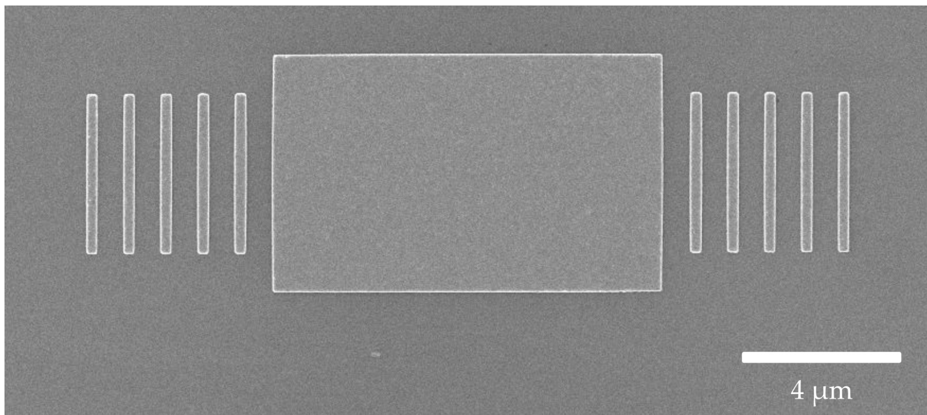


Figure 3.9: Example of an incoupling grating for a wavelength of 970 nm written with PMMA positive and imaged with the SEM after the liftoff of the gold film.

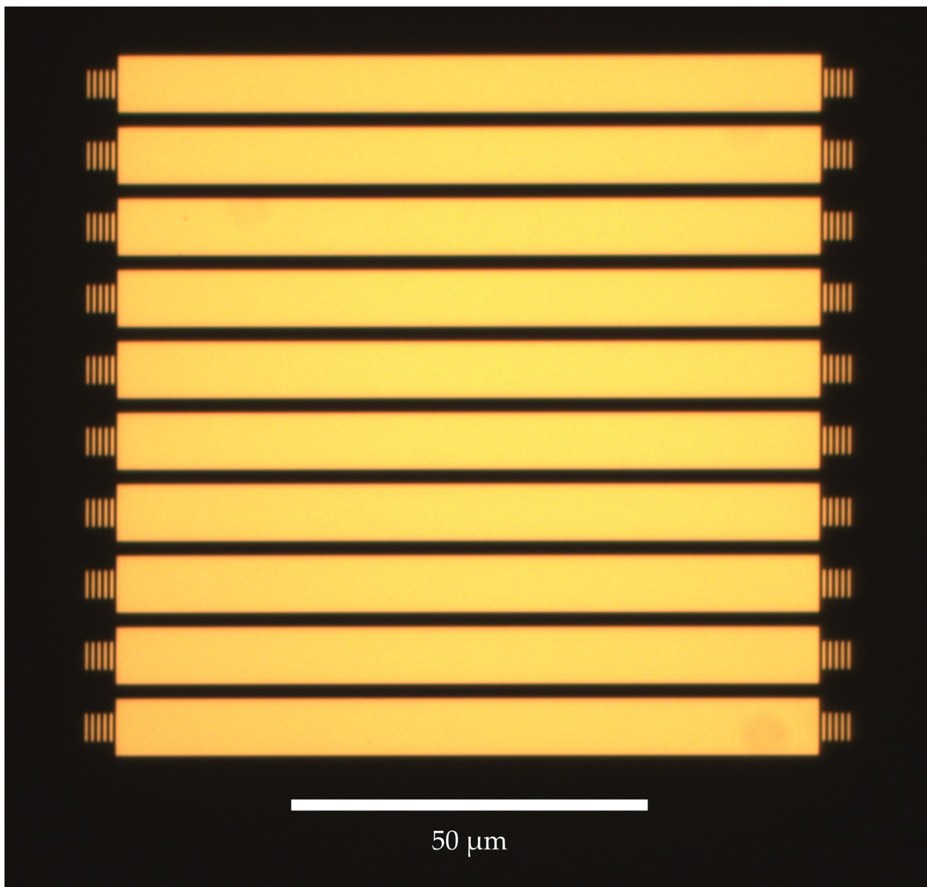


Figure 3.10: Example of multiple incoupling gratings for a wavelength of 970 nm written with PMMA positive tone and imaged with an optical microscope after the liftoff of the gold film. These incouplers together with the long gold areas form the basis for the second lithography step where the waveguides are written on top.

Final Fabrication Process

1. Spin coating of PMMA 950k A4 at 4000 rpm for 90 s on a ITO covered glass slide (1 cm^2).
2. Baking at 170 °C for 120 s. This yields a resist thickness of approximately 200 nm.
3. EBL of the incoupling gratings and the large gold areas with 20 kV and a 10 µm aperture. The gratings are written as individual lines with a line distance of 20 nm and a dose factor of 0.4. The period and width are chosen to match the illumination wavelength as described in Table 3.2. The large areas are written as areas with a dose factor of 0.163. The alignment markers and text are also included in this step.
4. Development of the resist in MIBK for 55 s followed by a rinse in isopropanol and drying with nitrogen.
5. Thermal evaporation of 60 nm gold with a 3 nm chromium adhesion layer.
6. Liftoff of the resist in NMP for one to two hours at 80 °C followed by rinsing in isopropanol, deionized water and drying with nitrogen.

CHAPTER 4

Near-Field Microscopy

Conventional far-field imaging techniques, such as optical microscopy, are fundamentally limited in their resolution by the diffraction limit. This limit arises from the wave-like nature of light and restricts the smallest resolvable feature size to approximately half the wavelength of the impinging light [11, 12]. For near-infrared (NIR) light wavelengths in the range of around 900 nm to 1 700 nm, this corresponds to a resolution limit of roughly 450 nm to 850 nm. However, many modern applications in nanophotonics, plasmonics, and material science require imaging and characterization at the nanoscale, well below the diffraction limit. One of the most prominent techniques is scanning near-field optical microscopy (SNOM). Similar to atomic force microscopy (AFM) or scanning tunneling microscopy (STM), it uses a sharp probe to locally interact with the sample allowing to push beyond the classical diffraction limit. Two different types of SNOM were initially developed: aperture-based SNOM [59, 60] and scattering-type (or aperture-less/tip-enhanced) SNOM (so called scattering-type scanning near-field optical microscope (s-SNOM)) [31, 61]. In this thesis, we employ a commercial s-SNOM setup (neaSNOM, *Neaspec*, *Attocube*). This system is based on an AFM to control the tip-sample distance and position and is paired with a tunable laser source and interferometric detection to obtain amplitude and phase resolved near-field data. The s-SNOM setup can be operated in both reflection and transmission mode, allowing for versatile measurements of various samples.

In the following sections the principles of s-SNOM are explained. Starting with a brief introduction to the AFM, which forms the basis of the s-SNOM setup. Then the s-SNOM itself is described, including the setup, laser source, detectors, and background suppression techniques. After that, the tip-sample interaction is discussed, followed by some practical aspects of operating a s-SNOM, including measurement corrections.

4.1 Atomic Force Microscopy

The AFM belongs to the family of scanning probe microscopy techniques, which were introduced first in the 1980s. The first scanning probe microscope, the STM, was invented by Gerd Binnig and Heinrich Rohrer and demonstrated for the first time in 1982 [62]. “For their design of the scanning tunneling microscope” they were awarded with the Nobel Prize for Physics, which they shared with Ernst Ruska “for his fundamental work in electron optics, and for the design of the first electron microscope” [63]. In the same year, Binning, Quate, and Gerber invented the AFM [64], which is the most widely used scanning probe microscopy technique today. In this thesis, the role of the AFM is to provide topography information of the fabricated samples, mainly the height and width of the waveguides, and it forms the foundation of the s-SNOM setup. The following description of the AFM is based on many sources and common knowledge, but a good overview can be found in [65].

The AFM uses a sharp tip attached to a cantilever to raster scan the surface of a sample. The tip is typically made of a sharp silicon or silicon nitride material, with a radius of curvature in the nanometer range. Close to the sample the tip is exposed to multiple forces, which can be described qualitatively by a Lennard-Jones potential [66]. For large distances, the potential is attractive, due to the van der Waals forces, and for small distances, the potential becomes repulsive, due to the Pauli exclusion principle. The resulting attractive and repulsive forces on the probing tip are sketched in Fig. 4.1. Depending on the sample many more forces like electrostatic, magnetic, and capillary forces can play a role, but can commonly be neglected for hard samples. An AFM can be used in a variety of modes such as the contact mode, tapping mode, or non-contact mode. In the contact mode, the tip is in constant contact with the sample surface, and the cantilever is deflected by the repulsive forces between the tip and the sample. In the non-contact mode, the tip is purely in the attractive regime, which is usually difficult to control. Therefore, the most commonly, and in this thesis exclusively, used mode is the intermittent contact mode, also known as tapping mode. In this mode, the cantilever oscillates at a frequency Ω close to its resonance frequency, and the tip intermittently contacts the surface. This means, that during each oscillation cycle, the tip hits the sample surface at its lower turning point. During the oscillation cycle, the tip is pulled towards the sample by the attractive forces, and then pushed away by the repulsive forces. The driving frequency is usually placed on the steepest side of the resonance at a tapping amplitude of 80-90 % of the maximum. This allows for a fast feedback of changes in the oscillation frequency of the tip. The attractive and repulsive forces blue and red shift the resonance frequency of the cantilever which is then translated into a change of the tapping amplitude. To monitor the bending amplitude of the cantilever, a laser beam is focused on the backside of the cantilever, and the reflected beam is detected by a photodiode as sketched in Fig. 4.2. In order to protect the photodiode from stray light, especially important when using additional laser sources to probe the sample, a narrow bandpass filter is used (FBH785-10, *Thorlabs*). A slight change in the cantilever’s bending causes a change in the laser beam’s position on the photodiode. In order to monitor the position of the deflection laser, one either uses a knife-edge or a quadrant photodiode.

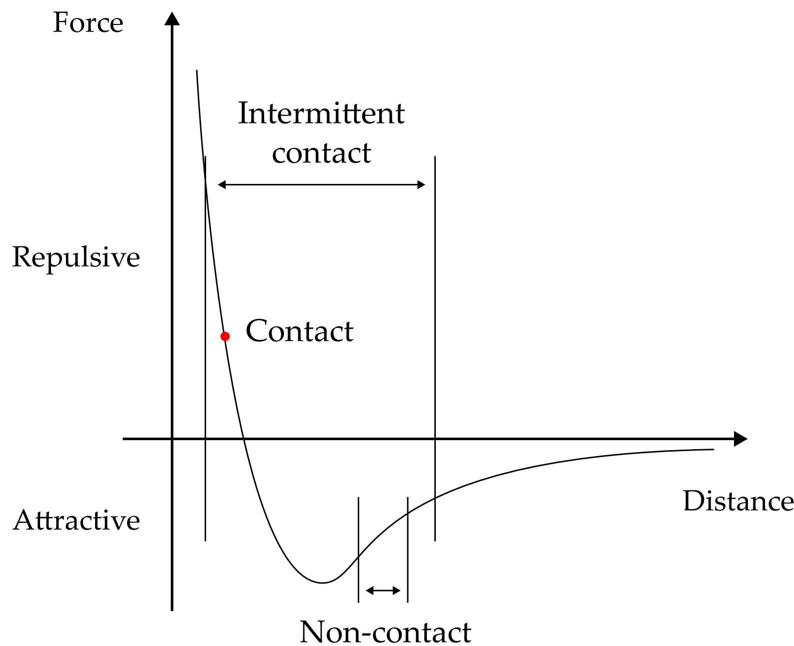


Figure 4.1: Schematic of the forces acting on the tip in the different modes of operation of an AFM.

The photodiode signal demodulated at the tip drive frequency [1] is used to control the tip-sample distance in a feedback loop with a proportional-integral-differential (PID) regulator. The distance is changed by moving the sample table closer to or further away from the tip using the piezo motor of the sample table. The piezo range in z-direction is approximately 3 µm. The height information can directly be extracted from the voltage signal of the piezo stage. The tip oscillation amplitude is typically in the tens of nanometer range. The contact between the tip and the sample is defined by how much the tapping amplitude should decrease, which is typically 75 to 90 % of the maximum tapping amplitude. A higher value corresponds to a smaller change and therefore leads to a weaker contact and a smaller force on the sample, while an unneccessary low value can lead to a faster tip wear. This value needs to be adjusted for each tip, and in general should be chosen as high as possible, while still keeping the tip in good contact with the sample.

The AFM can be used to measure topography, mechanical properties, and other surface characteristics. There are over 30 different modes of operation for AFM, each with its own advantages and disadvantages, but here we will focus on the topography measurement in tapping mode. To receive a topography image, the sample is raster scanned on a pixel by pixel basis. The lateral x-y-scan is done by two seperate piezo motors moving the sample

[1] The demodulation is discussed in more detail in Section 4.2.3.

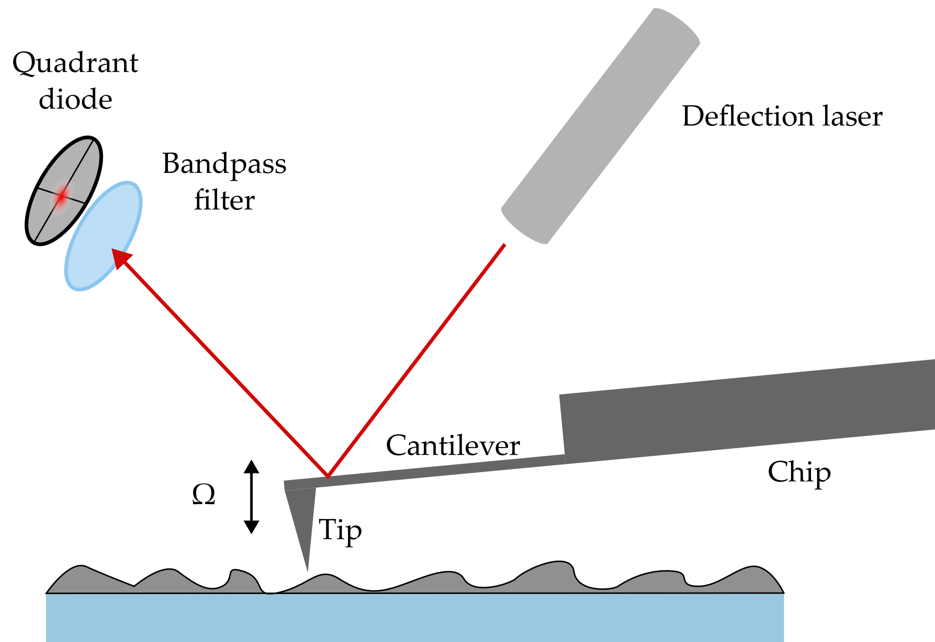


Figure 4.2: Schematic of the AFM setup. The tip is attached to a cantilever, which is oscillated at a frequency Ω close to its resonance. The deflection of the cantilever is measured by a laser beam reflected from the backside of the cantilever. The reflected laser beam is focused onto a photodiode, which detects the position of the laser beam on the diode. The photodiode signal is used to control the tip-sample distance in a feedback loop, which keeps the average tip-sample distance constant.

table, with a scan range of $100\,\mu\mathrm{m} \times 100\,\mu\mathrm{m}$ in x- and y-direction. This has the simple reason, that for s-SNOM measurements the tip is part of the interferometer, and therefore needs to be stationary during the scan. The scan resolution defines the distance between adjacent measurement points, where the signal is integrated for a short time. The two dimensional scan is done line by line, with a typical speed of 2 to $10\,\mu\mathrm{m}\,\mathrm{s}^{-1}$. Each line is scanned twice, once in the forward direction and one in the backward direction. This is on the one hand important to keep the lines well aligned, and on the other hand it allows to detect directional influences and artifacts in the measurement. The scan along the lines is called the fast axis, while the axis perpendicular is labeled the slow axis. The differentiation is important because slow variations like height drifts are much more visible in the slow axis. The direction of the fast axis can be rotated by 360° in steps of 0.1°. The scan speed is adapted to the sample topography and the desired resolution (or integration time). For AFM measurements, the integration time for each pixel typically does not matter, therefore one can choose a high resolution on the fast axis without increasing the overall measurement time. The number of lines (resolution of the slow axis) however is directly proportional to the overall measurement time. Therefore, the resolution on the fast axis is typically chosen higher than on the slow axis. The pixels created from the differently spaced measurement

points are then not symmetric. To receive a proportionally adequate representation of the sample the measurement should then be scaled or stretched accordingly, as will be explained in Section 4.4.

The piezo voltage to height transformation needs to be calibrated and checked in regular intervals against a reference sample, such as a silicon wafer with a known height. An example scan of the calibration sample (TGQ1, *Nanoandmore*) for this purpose is shown in Fig. 4.3. This ensures that the height information is accurate. Our AFM is only passively vibration isolated, therefore the height signal can be noisy, but typically height changes in the sub-nanometer range can be detected. For example, the presence of transition metal dichalcogenide (TMDC) monolayers can easily be observed, which are only a few Å thick.

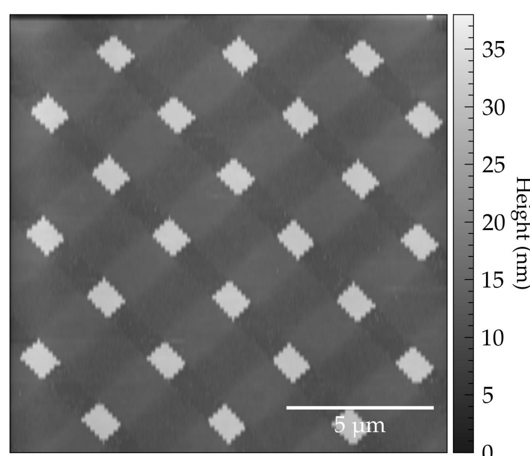


Figure 4.3: Example AFM scan of the TGQ1 (Nanoandmore) test sample.

Since the cantilever is not perfectly rigid, the tip can also bend or sway, which can lead to artifacts in the measurement. Typically, this is visible at the outsides of a scan where the scanning direction changes. Together with a finite tilt of the sample and imperfections in the sample table motors, this can lead to a non-flat surface in the measurement. This often needs to be corrected by a plane fit or a polynomial fit.

4.2 Scattering-Type Scanning Near-Field Optical Microscopy

A s-SNOM is, like the AFM or STM, a scanning probe technique, in which a sharp tip is used to probe the surface of a sample. It is based on an AFM to control the probes location and distance to the sample and adds illumination and detection. The s-SNOM can be used as a platform to measure a variety of nanoscale interactions. Depending on the needs one can perform for example Fourier transform infrared (nano-FTIR) spectroscopy [67, 68], Raman spectroscopy [69, 70], photoluminescence measurements [71], and even amplitude and phase resolved imaging of the optical near-fields [72, 73], all with nanoscale interaction volumes.

In our group we have a *neaSNOM* provided by *Neaspec* (*attocube*) with the nanoscale imaging module for reflection and transmission mode measurements. As the illumination source we use a tunable laser from *Stuttgart Instruments* with a wide tuning range. The tip is usually made out of a dielectric material coated with a thin metallic layer, such as gold or a platinum alloy. The metallic coating is crucial as it boosts the polarizability of the tip, which leads to stronger light scattering and therefore a higher signal. But metallic tips are softer and have in general also a stronger influence on the sample, which can lead to artificial features induced by the tip. The achievable lateral resolution is governed by the tip shape and is typically in the 30 to 50 nm range. Depending on the illumination wavelength the resolution can be multiple orders of magnitude smaller than the diffraction limit. This allows for imaging of local near-field information, which would otherwise be lost in the far-field.

In the following sections, the s-SNOM is described in more detail. First the setup itself and the laser source and detectors are introduced. Then the background suppression techniques are explained, which are crucial to extract the weak near-field signal from the dominating background. The interaction of the probing tip with the sample is discussed and supported by an analytical model. Some practical aspects of operating the s-SNOM are presented, including measurement corrections.

4.2.1 Setup

The s-SNOM setup is based on the previously described AFM setup, with the addition of optical components for illumination and detection. The image in Fig. 4.4 shows a schematic depiction of the setup viewed from the front. The s-SNOM can be operated in two different modes, the reflection mode and the transmission mode, indicated by the blue (reflection) and red (transmission) arrows.

In the reflection mode, the incident light is split by a beam splitter into a reference beam and a beam illuminating the sample. The built in beam splitters are based on zinc selenide (ZnSe) and are designed to work in the NIR with a specified wavelength range of 1 to 12 µm. The beam splitters are not wavelength independent, and especially for wavelengths below 900 nm the transmission is significantly reduced and the splitting ratio is not constant. The reference beam is reflected from a mirror oscillating with frequency M and passes an optional neutral density (ND) filter to reduce the reference beam intensity. The oscillating mirror is important for the background reduction explained in Section 4.2.3. The beam illuminating the sample is focused on the apex of the tip by the upper parabolic mirror. The scattered light from the near-field interaction region is collected by the same parabolic mirror and is then overlaid with the reference beam. Together, these two beams form a Michelson interferometer, which is used to measure amplitude and phase resolved near-field information.

In the transmission mode, the light is focused on the sample from the bottom using the lower parabola. A flip mirror can be used to illuminate the sample from above, in order to align the light collection beam path with the tip in transmission mode. A reference beam is created with a beam splitter and guided through a Mach-Zehnder interferometer and

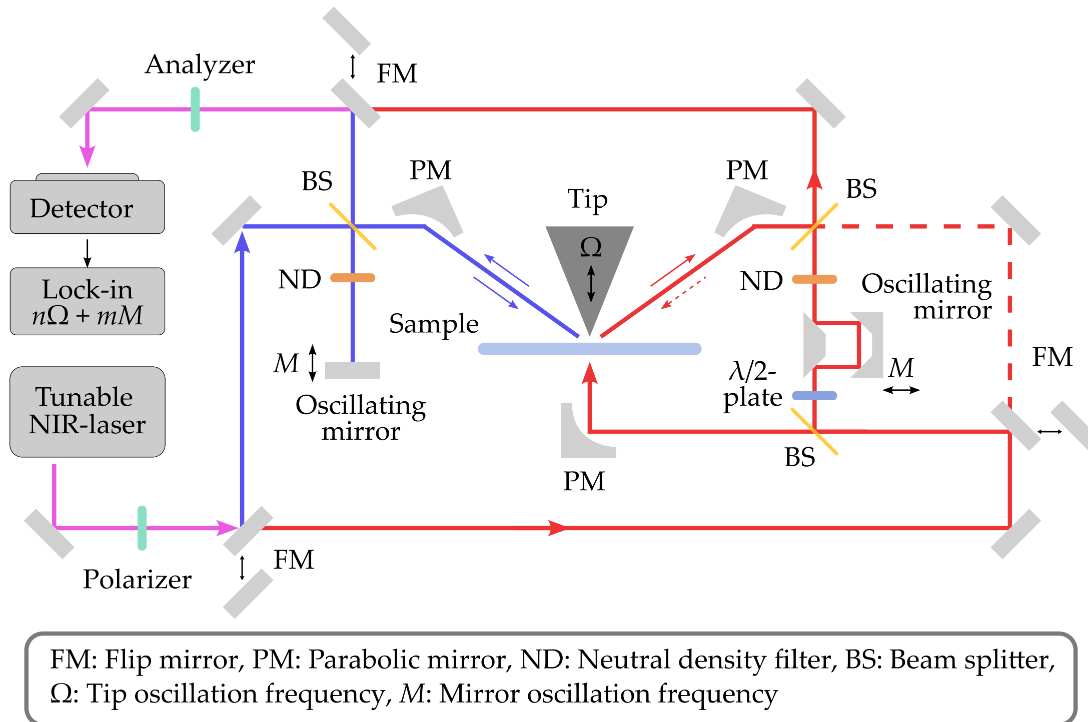


Figure 4.4: Schematic of the s-SNOM setup viewed from the front. The two different imaging modes, reflection (blue) and transmission (red), are indicated by the colored arrows. The purple arrows indicate the beam path shared by both modes. In the transmission mode, the light can also be focused from the top using a flip mirror for alignment purposes.

then overlaid with the collected light. The mirror of the Mach-Zehnder interferometer is oscillating similar to the end mirror of the Michelson interferometer. The reference beam can also be adjusted by a ND filter to reduce the intensity. The transmission mode has a significant advantage over the reflection mode, as the illumination can be moved synchronously with the sample table. This is especially helpful for measurements of travelling waves and waveguides. The movement of the lower parabolic mirror along the x- and z-axis leads to a phase gradient in the measurement, since this changes the optical path length relative to the reference beam. This needs to be corrected in the data processing, which is explained in Section 4.4.

Both modes have an input polarizer and an output analyzer (WP25M-UB, *Thorlabs*). The input polarizer is used to set the polarization of the incident light, while the output analyzer is used to select the polarization of the light collected from the tip. For the excitation of the tip in the reflection mode, a polarization along the tip axis is typically used, as this leads to the strongest near-field via the lightning rod effect [74]. Similar to the excitation, the tip scattered light is primarily polarized along the tip axis, but other polarization components can also be present, these are often undesirable and therefore suppressed by the analyzer. In transmission measurements the tip is in general not directly excited by the incident light, as

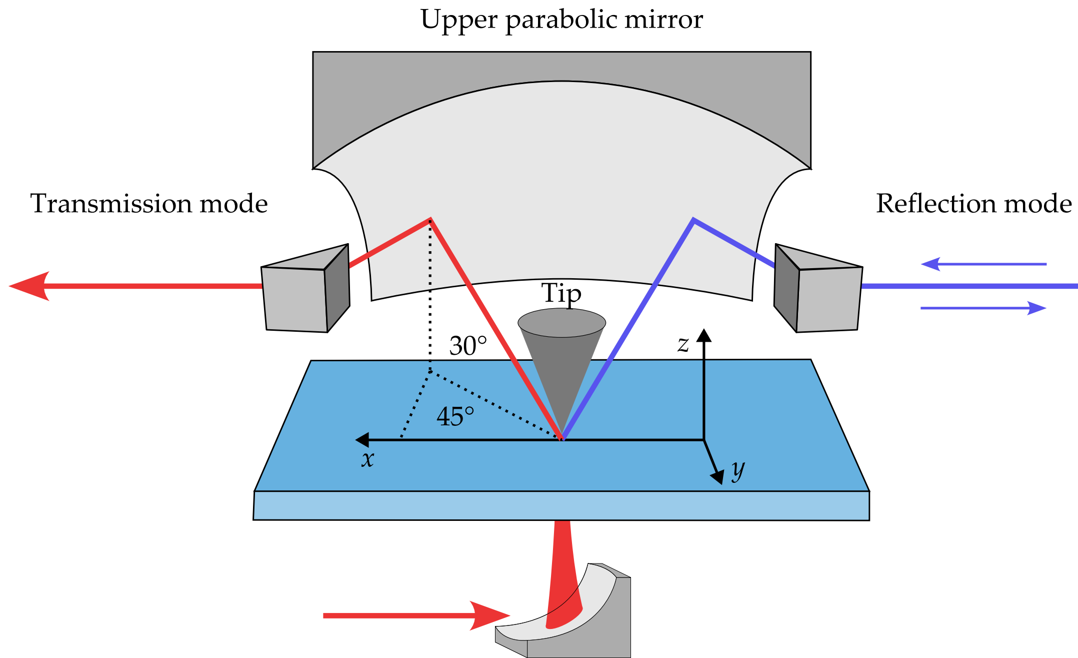


Figure 4.5: Schematic of the tip emission directions in s-SNOM setup, viewed from the back to show the collecting parabolic mirror. The coordinate system defined here is used throughout this thesis. The tip illumination and emission have an inclination angle of 30° with respect to the $x-y$-plane and an azimuthal angle of 45° with respect to the x-axis.

the tip axis is parallel to the incident light propagation direction. Depending on the sample and its rotation, arbitrary input polarizations can be used to excite the sample. Since the space is limited, the polarization optics often have to be placed before the beam splitter. It is important to use a $\lambda/2$-plate (700 to 2 500 nm achr., *B. Halle*) to rotate the polarization of the reference beam, such that it matches the orientation of the analyzer. Otherwise the tip scattered light and the reference cannot interfere on the detector.

As it is shown in Fig. 4.5, the tip illumination and scattering have an inclination angle of 30° with respect to the x-y-plane and an azimuthal angle of 45° with respect to the x-axis. This leads to a rotation of the polarization vector of approximately 12.7° to the z-axis, a detailed derivation can be found in the appendix A.1.2. This effect is quite small, but can be important for measurements where the in-plane polarization component is important. It should be noted that the angle is reversed for the reflection and transmission mode respectively, due to the mirrored collection geometry.

4.2.2 Laser Source and Detectors

The laser source used in the s-SNOM setup is a tunable femtosecond laser from *Stuttgart Instruments*. It offers pulsed excitation with approximately 400 fs long pulses at a repetition rate of 40 MHz. The laser can be tuned from 700 to 980 nm and 1 400 to 1 960 nm with the option to use the pump laser at 1 064 nm and the idler signal with wavelengths of up to 4 µm. The laser passes through one of two polarization maintaining single mode fibers, (P1-1064PM-FC-2, *Thorlabs*) and (P1-1310PM-FC-2, *Thorlabs*). The laser beam has a diameter of approximately 4 mm after the single mode fiber, which is increased to roughly 8 mm by a telescope to improve the usable numerical aperture (NA) of the parabolic mirrors and thus improve the focusing of the laser light on the tip or sample.

The s-SNOM has two built in detectors for different wavelength ranges. The first detector (2051-FS, *Newport*) is a silicon based photodiode, which can detect wavelengths from 300 to 1 050 nm. The second detector (PDA20CS-EC, *Thorlabs*) is an InGaAs based photodiode, which can detect wavelengths from 900 to 1 700 nm. The gain of the detectors can be adjusted to optimize the signal for different measurement conditions. It is optimized to guarantee a detection of signals above 1 MHz bandwidth. This is important, as higher harmonics of the tip oscillation frequency can easily reach frequencies in the MHz range. The background supression in s-SNOM is based on higher harmonic demodulation, as will be explained in the next section.

4.2.3 Extraction of the Near-Field Signal

One of the key difficulties in s-SNOM measurements is the detection of the weak near-field signal scattered by the tip, as the background signal typically dominates the detected signal by several orders of magnitude. To overcome this issue, two background suppression techniques are employed. The most common technique is to use a lock-in amplifier, which allows for demodulation at a higher harmonic of the tip oscillation frequency. As we will see later, this is not sufficient for an effective background suppression. Thus, a second technique is required. Here, we use an interferometric detection scheme, specifically a pseudo-heterodyne detection scheme.

Before we dive into the details of the background suppression techniques, we first need to understand the origin of the background electric fields. Typically, the background is caused by light from the illumination source, which is scattered by the tip and the sample. The light scattered by the sample or other stationary objects can be filtered via lock-in detection. However, the background light scattered by the tip is modulated with its oscillation frequency, which makes it difficult to differentiate from the near-field scattering. Additionally, there are many possible pathways for the light reflected from the tip to the detector, which can lead to interference effects and therefore a complex background signal. Since part of the modulated background light is also reflected from the sample, it can also vary whitin the scan if the sample reflectivity is not constant. The two modes in which our s-SNOM can be operated, the reflection mode and the transmission mode, differ substantially in the amount

of background light. In the transmission mode the tip is often not directly illuminated. The fields scattered by the tip are then primarily created by the sample's interaction with the incident light and not via reflections of the incident light from the tip. Especially for samples with weak light transmission the background is significantly reduced. This is the case for the samples used in Chapter 6 and Chapter 7. Here, the substrate is mostly covered by a metallic film. This allows to use lower demodulation orders compared to the reflection mode, as less background light reaches the detector.

In the following, the two background suppression techniques are explained in more detail. Both techniques are described analytically and are also visualized by a numerical example. The analytical derivations are based on the works of Nenad Ocelić et al. [75, 76].

Lock-in Detection

A lock-in amplifier is often used to extract the amplitude and phase of a weak alternating current (AC) signal with known frequency, even if the signal strength is orders of magnitude below the noise floor. The technique is based on the multiplication of the input signal with a reference signal at the same frequency, followed by a low-pass filter to extract the direct current (DC) component. By using not only a single reference signal but two 90° phase shifted signals, one can extract both the in-phase and out-of-phase components of the input signal. This allows to reconstruct the amplitude and phase of the original signal. More details about the use of lock-in amplifiers in s-SNOM can be found in [67].

As mentioned before, the lock-in detection is a first approach to filter out background light reaching the detector. In the following, we will discuss how the fields reaching the detector can be expressed and how the signals are processed in order to achieve an effective background suppression. During operation the scanning tip oscillates with a frequency Ω and an amplitude A_T. The tip oscillation amplitude, also called the tapping amplitude, is usually in the tens of nanometer range. The oscillation frequency is in the range of 50 to 300 kHz. Due to the small tip oscillation amplitude relative to the cantilever length (several hundred micrometers), the restoring force can be approximated as linear. Thus, the tip-sample distance $H_T(t)$ follows a sinusoidal oscillation

$$H_\mathrm{T}(t) = H_0 + A_\mathrm{T}\left(1 + \cos\left(\Omega t\right)\right) , \tag{4.1}$$

where H_0 is the minimum tip-sample distance. If the tip is illuminated and in close proximity to a sample it will scatter light, which eventually reaches the detector. If the light is scattered or reflected from the tip it will be modulated with the tip frequency. The light impinging on the detector is then constituted of multiple contributions. The main contribution is unmodulated stray light, followed by light reflected from the tip shaft, which does not contain any near-field information. The weakest contribution is light scattered from the tip originating from the tip-sample interaction. One exploits the fact that the unmodulated stray light is filtered out completely by the lock-in detection. Secondly, the background signal is only weakly modulated compared to the near-field signal. We can demonstrate this by

writing out the amplitudes of the electric fields interacting with the tip as

$$A_{\mathrm{NF}}(t) = e^{-\frac{H_{\mathrm{T}}(t)}{\Lambda}} E_0\,, \tag{4.2}$$

$$A_{\mathrm{BG}}(t) = (1 - m_{\mathrm{BG}} H_{\mathrm{T}}(t)) E_0\,, \tag{4.3}$$

where E_0 is the amplitude of the incident light, Λ is the decay length of the near-field, and m_{BG} is a constant describing the background decay. We assume that the near-field amplitude A_{NF} exponentially decays with the tip-sample distance while the background amplitude A_{BG} decays linearly. We intentionally neglect the tip-sample coupling here, as we only want to understand the basic principles of the lock-in detection. A more detailed treatment of the tip-sample interaction is discussed in Section 4.2.4. The amplitude of the tip scattered light is proportional to the initial amplitude modulated by the tip oscillation

$$s_{\mathrm{NF}}(t) \propto A_{\mathrm{NF}}(t) = e^{-\frac{H_{\mathrm{T}}(t)}{\Lambda}} E_0\,, \tag{4.4}$$

$$s_{\mathrm{BG}}(t) \propto A_{\mathrm{BG}}(t) = (1 - m_{\mathrm{BG}} H_{\mathrm{T}}(t)) E_0\,. \tag{4.5}$$

This behavior is displayed in Fig. 4.6 (a-c). As mentioned earlier, its clear that the higher localization of the near-field leads to a stronger modulation of the scattered light. However, we neglected the phase of the fields so far. A better representation of the scattered light is a complex-valued scattering coefficient including the scattering amplitude and phase. For the two scattered fields we can write

$$\sigma_{\mathrm{NF}}(t) = s_{\mathrm{NF}}(t) e^{i\phi_{\mathrm{NF}}}\,, \tag{4.6}$$

$$\sigma_{\mathrm{BG}}(t) = s_{\mathrm{BG}}(t) e^{i\phi_{\mathrm{BG}}}\,. \tag{4.7}$$

$\sigma_{\mathrm{NF}}(t)$ is the near-field scattering coefficient and $\sigma_{\mathrm{BG}}(t)$ is the background scattering coefficient. We assume that the tip-sample distance has no influence on the scattering phase, which is a good approximation for small oscillation amplitudes. For simplicity we can ignore the phase of the scattered fields for now, but it will become important for the interferometric method. In order to get a representation of how the lock-in would decompose the scattered amplitudes, we express the scattering amplitudes as a fourier series

$$s_{\mathrm{NF}}(t) = \sum_{n=0}^{\infty} s_{\mathrm{NF},n} e^{in\Omega t}\,, \tag{4.8}$$

$$s_{\mathrm{BG}}(t) = \sum_{n=0}^{\infty} s_{\mathrm{BG},n} e^{in\Omega t}\,. \tag{4.9}$$

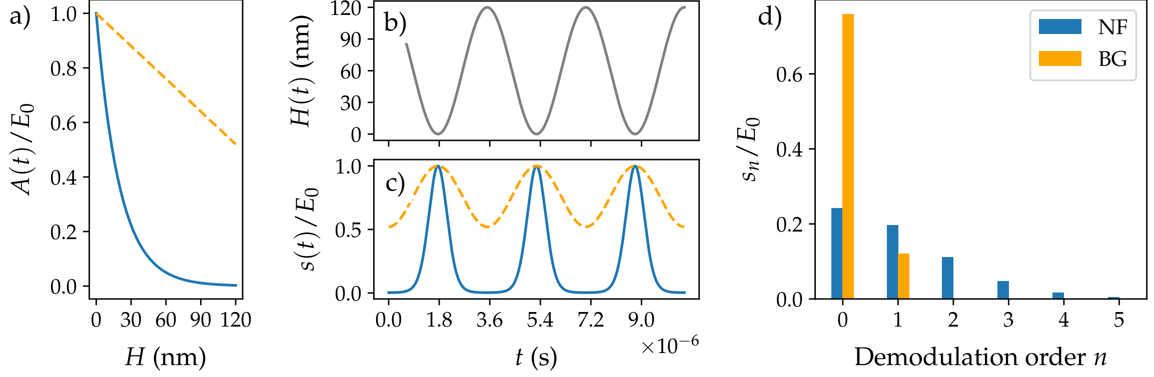


Figure 4.6: Visualization of the lock-in detection principle. Exponentially decaying near-field amplitude and linear background amplitude (a). Tip oscillation (b) and scattering amplitude of near-field and background for three tip oscillation cycles (c). Demodulated scattering amplitudes showing the suppression of the background term (d). Parameters are: $\lambda = 1\,000\,\text{nm}$, $\Omega = 2\pi\cdot\, 285\,\text{kHz}$, $\Lambda = 20\,\text{nm}$ and $m_{\text{BG}} = 1/(4 * \lambda)$. Adopted from the PhD thesis of Lars Mester [77].

The Fourier coefficients can be calculated by integrating over one period of the oscillation

$$s_{\text{NF},n} = \frac{1}{2\pi}\int_0^{2\pi/\Omega} s_{\text{NF}}(t)e^{-in\Omega t}dt\,, \tag{4.10}$$

$$s_{\text{BG},n} = \frac{1}{2\pi}\int_0^{2\pi/\Omega} s_{\text{BG}}(t)e^{-in\Omega t}dt\,. \tag{4.11}$$

The scattering coefficients at the n-th harmonic of the oscillation frequency Ω are shown in Fig. 4.6 (d). The background signal dominates the 0-th harmonic, which corresponds to the DC-component, but it quickly becomes weaker than the near-field signal due to the stronger modulation of the near-field signal. However, in reality the detector cannot differentiate between the various contributions to the scattered light, instead we have to sum all contributions. This total scattered field can be written as

$$E_{\text{T}}(t) = E_0\left[\sigma_{\text{NF}}(t) + \sigma_{\text{BG}}(t)\right]\,, \tag{4.12}$$

and expressed as a fourier series in Ω

$$E_{\text{T}}(t) = E_0\sum_{n=0}^{\infty} e^{in\Omega t}\left[\sigma_{\text{NF},n}(t) + \sigma_{\text{BG},n}(t)\right]\,. \tag{4.13}$$

Since the detector can only measure the intensity of the field, we can express the detector signal as

$$U_{\text{D}}(t) = k_{\text{D}}\left|\sum_{n=0}^{\infty} e^{in\Omega t}\left[\sigma_{\text{NF},n} + \sigma_{\text{BG},n}\right]\right|^2\,. \tag{4.14}$$

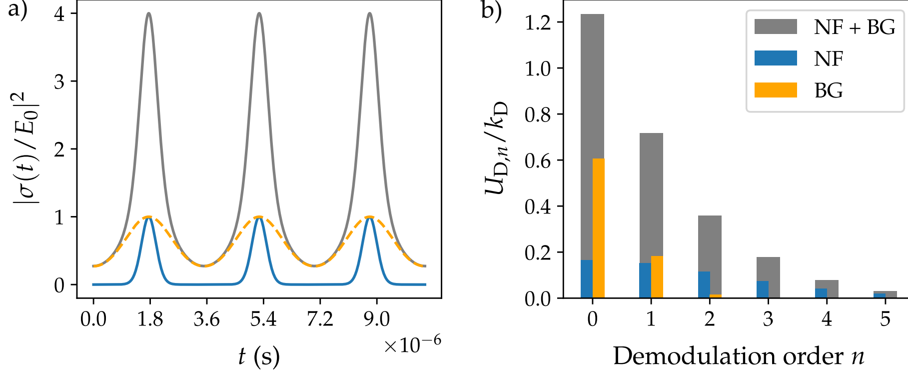


Figure 4.7: Scattering amplitudes of the near-field, background and their interference (a). Demodulated detector signal at the n-th harmonic (b) for $\phi_{\mathrm{NF}} = \phi_{\mathrm{BG}} = 0$. The interference of the near-field scattering with the background scattering leads to a significant background contribution to higher harmonics.

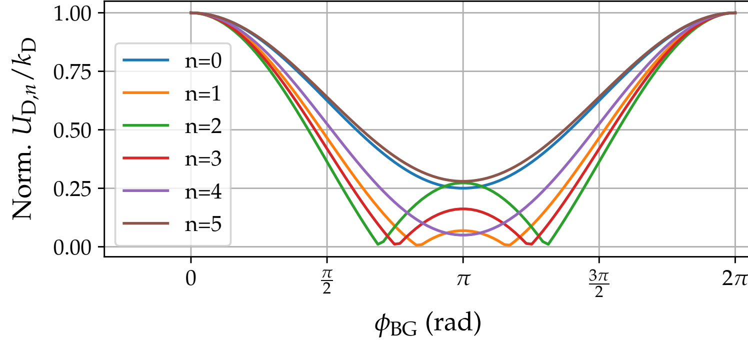


Figure 4.8: Demodulated detector signal at the n-th harmonic for different background phases ϕ_{BG}. The background phase has a significant impact on the detected signal, even for higher harmonics.

Here, k_{D} is a constant describing the detector sensitivity and other parameters such as the laser power and probing tip related parameters. The detector signal is proportional to the intensity of the sum of the near-field and background scattering contributions. This non-linearity consequently leads to a mixing of the Fourier components of the near-field and background. This effect is illustrated in Fig. 4.7. In (a) the near-field and background scattering intensities are shown along with their interference. Plotting the demodulated detector signal at the n-th harmonic, as shown in (b), one can see that the background influences higher harmonics due to interference of the near-field and background fields. In this example we assume $\phi_{\mathrm{NF}} = \phi_{\mathrm{BG},0} = 0$, but in general the background phase is much more complex and often varies over time or while measuring. The influence of the background phase is displayed in Fig. 4.8, which shows the normalized demodulated detector signal at the n-th harmonic for different background phases ϕ_{BG}. The demodulated detector signal at

the n-th harmonic can also be approximated analytically as [75]

$$U_{D,n} \approx 2k_D s_{BG,0} s_{NF,n} \cos\left(\phi_{NF,n} - \phi_{BG,0}\right) , \tag{4.15}$$

please refer to the appendix A.1.1 for a more detailed derivation. In the analytic formula the demodulated detector signal is proportional to the near-field amplitude $s_{NF,n}$, but the background term $s_{BG,0}$ remains. This is often referred to as the "multiplicative background", as it is an additional factor in the demodulated signal stemming from the interference of the tip scattered near-field with tip scattered background electric fields. Even for high harmonics, the background term can not be completely removed by the lock-in detection. Additionally, the detected signal decreases with increasing harmonic order, leading to a lower signal-to-noise ratio (SNR). The oscillation frequency of the tip is also important. The tip frequency is typically above 250 kHz and the bandwith of our detectors is limited to roughly 1 MHz. Therefore, in experiments we are usually limited to the detection of the first 4 harmonics. One may also use tips with much lower frequencies, such as force modulation tips, but then the lower frequency leads to a slower variation of the near-field signal compared to far-field variations and thus higher hamonics should be used to effectively reduce the background contributions. In order to remove the multiplicative background, an additional supression method has to be used, which will be discussed in the next section.

Pseudo-Heterodyne Detection

The pseudo-heterodyne detection is an interferometric detection scheme, which further supresses background and enables phase-resolved near-field measurements. It is similar to prior approaches such as homodyne detection and heterodyne detection [32, 72, 78]. The homodyne detection is based on taking two measurements and changing the phase of the reference beam between the two measurements by $\pi/2$. This allows to extract the amplitude and phase of the near-field signal, but it relies on the assumption that the reference beam is much stronger than the background signal. A more reliable method is the heterodyne detection, which uses a frequency modulated reference beam. The interference of the frequency modulated reference beam with the near-field signal leads to sidebands in the detector signal and a much better background suppression. However, the heterodyne detection requires a high frequency modulation of the reference beam, which is typically achieved by an acousto-optic modulator. This leads to a wavelength dependent modulation frequency, which can be problematic for measurements over a wide wavelength range. Additionally, the modulation frequency is typically in the MHz range, which requires fast detectors with a high bandwidth, which are often more expensive and have a lower SNR.

The pseudo-heterodyne detection is based on the same principles as the heterodyne detection, but it uses a phase modulation of the reference beam instead of a frequency modulation, effectively negating its shortcomings. In the following, we will first discuss the basic principles of the pseudo-heterodyne detection and then show how it can be used to suppress the background signal. Additionally, we will also elucidate how the near-field

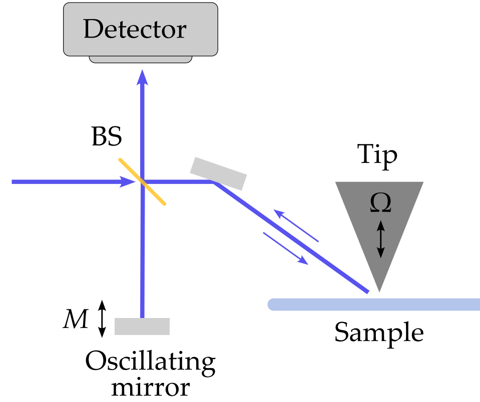


Figure 4.9: Schematic of the pseudo-heterodyne detection. The reference beam is phase modulated by oscillating the end mirror of the Michelson interferometer at a frequency M. The near-field light is scattered by the tip and interferes with the reference beam at the detector.

amplitude and phase can be extracted from the detector signal. Similar to the previous section, we will illustrate the general principles with a simple numerical example and add an analytical description of the detector signal. We will see later, that the extraction of the complex near-field scattering coefficient from the numerical example is not straightforward, but the analytical description allows to demonstrate the extraction.

Building on the basic lock-in detection, the addition of a Michelson interferometer allows to measure the phase of the near-field signal relative to the reference beam. The phase modulation is achieved by oscillating the end mirror of the interferometer at a frequency M, which is much lower than the tip oscillation frequency Ω. In Fig. 4.9 a schematic of the method is displayed. Assuming that the background signal has no temporal coherence with the reference beam, it can be suppressed by demodulating the detector signal at frequencies $n\Omega \pm mM$, where n and m are integers. Furthermore, the reference beam is typically much stronger than the tip scattered fields, leading to an amplification of the near-field signal in the generated sidebands. This is caused by the interference of the near-field signal with the strong reference beam, leading to a more pronounced modulation of the detector signal.

Similar to the lock-in detection, we can express the electric field at the detector as the sum of the near-field E_{NF} and background E_{BG} contributions with the reference beam E_{R}

$$E_{\mathrm{D}}(t) = E_{\mathrm{NF}}(t) + E_{\mathrm{BG}}(t) + E_{\mathrm{R}}(t) = E_{\mathrm{T}}(t) + E_{\mathrm{R}}(t)\,, \tag{4.16}$$

with $E_{\mathrm{R}}(t) = \rho e^{i\gamma\cos(Mt)}$, $\rho = r_{\mathrm{R}} e^{i\phi_{\mathrm{R}}}$, the amplitude r_{R} and phase offset ϕ_{R} of the reference beam, γ is the modulation depth and M is the modulation frequency. The detector signal is then again the intensity of the total electric field times an unknown factor k_{u}

$$U_{\mathrm{D}}(t) = k_{\mathrm{u}}\,|E_{\mathrm{T}}(t) + E_{\mathrm{R}}(t)|^2\,. \tag{4.17}$$

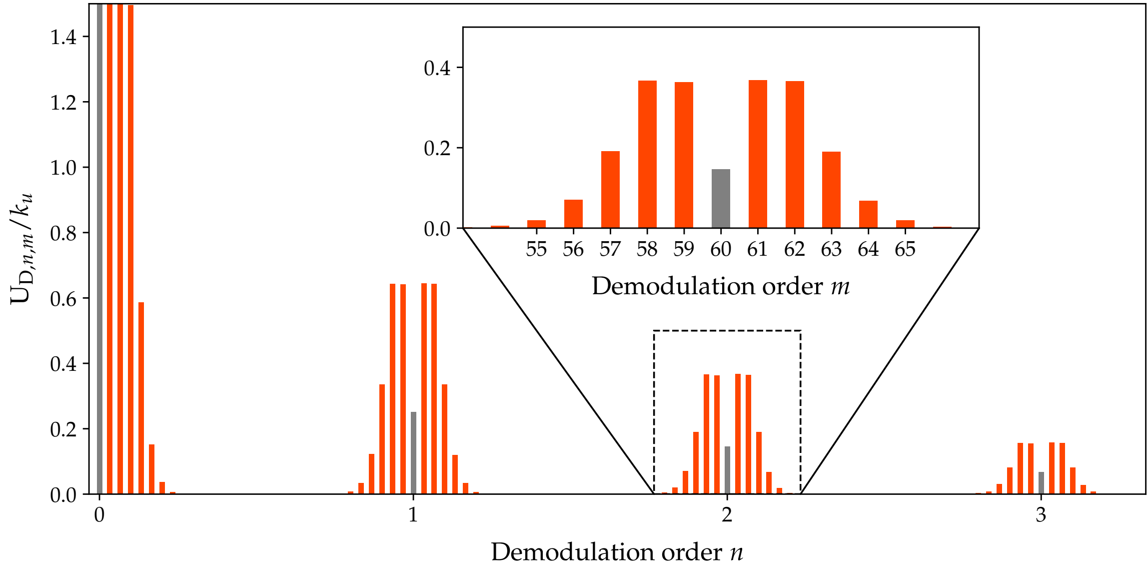


Figure 4.10: Numerical visualization of the demodulated detector signal. Shown is a wide range in frequency space, including the first 3 tip oscillation harmonics. In the inset a zoomed in version of the second tip oscillation harmonic is displayed, where the sidebands are clearly visible. This simulation corresponds to an integration over a timespan in which the background phase is varied from 0 to 2π. The parameters used for the simulation are: $\omega_0 = 100\Omega$, $M = \Omega/30$, $r_R = 5$, $\phi_{\text{NF}} = \pi/4$, $\phi_R = 0$, $\gamma = 2.63$ and $s_{\text{NF},n}$ is as defined before.

The demodulated detector signal looks qualitatively like the one shown in Fig. 4.10. In this numerical calculation, the mirror oscillation frequency M is much higher than in practice to make the sidebands visible. The inset shows a zoomed in version of the signal surrounding the second harmonic of the tip oscillation. The sidebands can have a much higher amplitude than the main harmonic, due to the higher amplitude of the reference beam. The sidbands are usually assumed to be background free. However, this is only true if we assume that the background field cannot interfere with the reference field. This is typically the case, as the background field is presumably not temporally coherent with the reference field, due to the complex mechanisms of the background generation. But in general the sidebands can still contain a significant background contribution. In order to include the background field in the higher harmonics but neglect it in the sidebands, we average the demodulated detector signal while changing the background phase ϕ_{BG} from 0 to 2π. This leads to a cancellation of the background contribution in the sidebands, while the main harmonics still contain a background contribution. This should also quantitatively correspond to the experimental situation.

From the numerical example, it is not straight forward how to extract the near-field amplitude and phase. To gain further insight into the underlying mechanisms, one can calculate the demodulated detector signal analytically. Herefore, one can express the electric fields via complex scattering coefficients $\sigma_T = \sigma_{\text{NF}} + \sigma_{\text{BG}}$, and expand both sides into a

Fourier series. The detector signal can then be expressed as

$$\sum_{m,n=-\infty}^{\infty} U_{\mathrm{D},n,m} e^{i(n\Omega+mM)t} = k_{\mathrm{u}} \left| \sum_{n=-\infty}^{\infty} \sigma_T e^{in\Omega t} + \sum_{m=-\infty}^{\infty} \rho_m e^{imMt} \right|^2 , \tag{4.18}$$

where the coefficients ρ_m are equal to

$$\rho_m = r_{\mathrm{R}} J_m(\gamma) e^{i\phi_{\mathrm{R}}+im\pi/2} , \tag{4.19}$$

where $J_m(\gamma)$ is the Bessel function of the first kind of order m. The detector signal demodulated at the n-th harmonic of the tip oscillation frequency Ω and the m-th harmonic of the end mirror modulation frequency M can then be expressed as [76]

$$U_{\mathrm{D},n,m} = 2k_{\mathrm{u}} r_{\mathrm{R}} J_m(\gamma) s_{\mathrm{T},n} \cos\left(\phi_{\mathrm{T},n} - \phi_{\mathrm{R}} - m\pi/2\right) , \tag{4.20}$$

where $s_{\mathrm{T},n}$ is the total scattering amplitude at the n-th harmonic and $\phi_{\mathrm{T},n}$ is the total scattering phase at the n-th harmonic. The resulting detector spectrum is shown in Fig. 4.11 for $n = 2$ and $m = \pm 1, 2, 3, 4, 5$. Formally, the background signal is still present in the demodulated signal. But like before, we assume that the sidebands are free from background, so $s_{\mathrm{T},n} = s_{\mathrm{NF},n}$ for $m \neq 0$. The sidebands contain the real and imaginary part of the near-field signal. To retrieve the real and imaginary part we write out the terms for the first two sidebands using Euler's formula for the cosine

$$U_{\mathrm{D},n,1} = k_{\mathrm{u}} r_{\mathrm{R}} J_1(\gamma) s_{\mathrm{T},n} \left(-ie^{i\left(\phi_{\mathrm{NF},n}-\phi_{\mathrm{R}}\right)} + ie^{i\left(-\phi_{\mathrm{NF},n}+\phi_{\mathrm{R}}\right)} \right) , \tag{4.21}$$

$$U_{\mathrm{D},n,2} = k_{\mathrm{u}} r_{\mathrm{R}} J_2(\gamma) s_{\mathrm{T},n} \left(-e^{i\left(\phi_{\mathrm{NF},n}-\phi_{\mathrm{R}}\right)} - e^{i\left(-\phi_{\mathrm{NF},n}+\phi_{\mathrm{R}}\right)} \right) . \tag{4.22}$$

In order to extract the complex near-field scattering coefficient $\sigma_{\mathrm{NF},n} = s_{\mathrm{NF},n} e^{i\phi_{\mathrm{NF},n}}$, we can combine for example the first two sidebands

$$s_{\mathrm{NF},n} e^{i\phi_{\mathrm{NF},n}} = k_{\mathrm{p}} \left(i\frac{U_{\mathrm{D},n,1}}{J_1(\gamma)} - \frac{U_{\mathrm{D},n,2}}{J_2(\gamma)} \right) e^{i\phi_{\mathrm{R}}} , \tag{4.23}$$

where $k_{\mathrm{p}} = 1/(2k_{\mathrm{u}} r_{\mathrm{R}})$ is a constant prefactor. A suitable choice of the modulation depth γ makes two subsequent sidebands directly proportional to the real and imaginary part of the near-field signal. For the first two sidebands this means $\gamma = 2.63$. Experimentally, this modulation depth corresponds to a mirror vibration amplitude of $A_M = 0.5\gamma\lambda/(2\pi)$. The constant prefactor k_{p} does not need to be determined experimentally if only relative amplitudes are of interest. The near-field amplitude and phase can then be calculated from

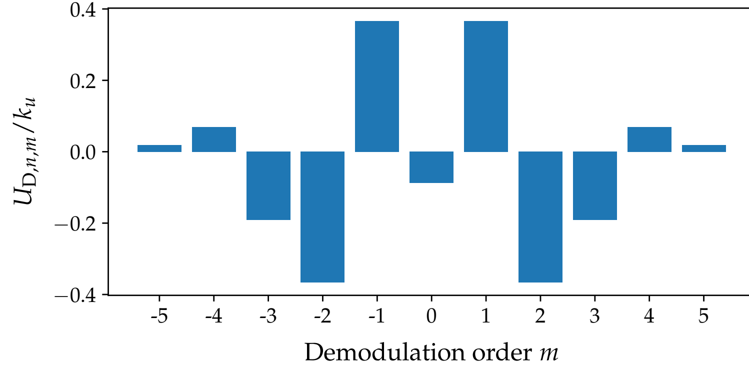


Figure 4.11: Analytic result for the second harmonic of the tip oscillation frequency Ω. The sidebands are shown for the first five harmonics of the end mirror modulation frequency M. The parameters are: $r_{\mathrm{R}} = 5$, $\phi_{\mathrm{NF},n} = \pi/4$, $\phi_{\mathrm{R}} = 0$, $\gamma = 2.63$ and $s_{\mathrm{NF},n}$ is as defined before.

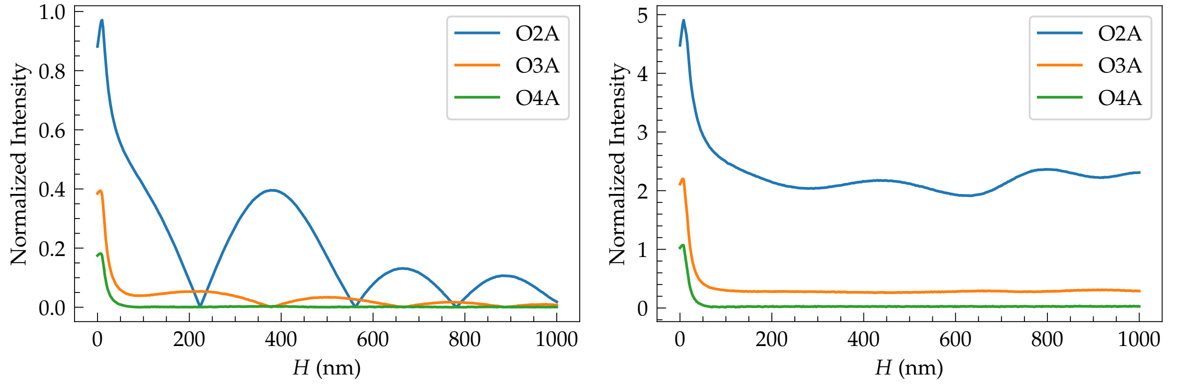


Figure 4.12: Example of approach curves on a gold substrate, demonstrating the impact of pseudo-heterodyne detection. (a) Without pseudo-heterodyne detection, (b) with pseudo-heterodyne detection.

the sidebands

$$s_{\mathrm{NF},n} = \frac{k_{\mathrm{p}}}{J_1(\gamma)} \sqrt{U_{\mathrm{D}n,1}^2 + U_{\mathrm{D}n,2}^2}\,, \tag{4.24}$$

$$\phi_{\mathrm{NF},n} = -\arctan\left(\frac{U_{\mathrm{D}n,1}}{U_{\mathrm{D}n,2}}\right) + \phi_{\mathrm{R}}\,. \tag{4.25}$$

An example of lock-in and pseudo-heterodyne detection is shown in Fig. 4.12. The left image shows an approach curve measured on a gold substrate. The second and third harmonics show a significant background contribution for large tip-distances. After adding the pseudo-heterodyne detection, the oscillations are significantly reduced, although a DC offset remains. The amplitudes are also amplified by roughly a factor of 5.

The analytic formula for the demodulated detector signal Eq. (4.20) is not strictly positive, as it would be expected for a photo-detector signal. This is also reflected in the numerical results shown in Fig. 4.10. For the correct extraction of the near-field phase the sign of the sidebands is essential. Experimentally, a lock-in amplifier is used to demodulate the detector signal, using the reference signal of the end mirror oscillation. This allows to extract the sidebands amplitude and phase shift relative to the reference signal.

4.2.4 Tip-Sample Interaction

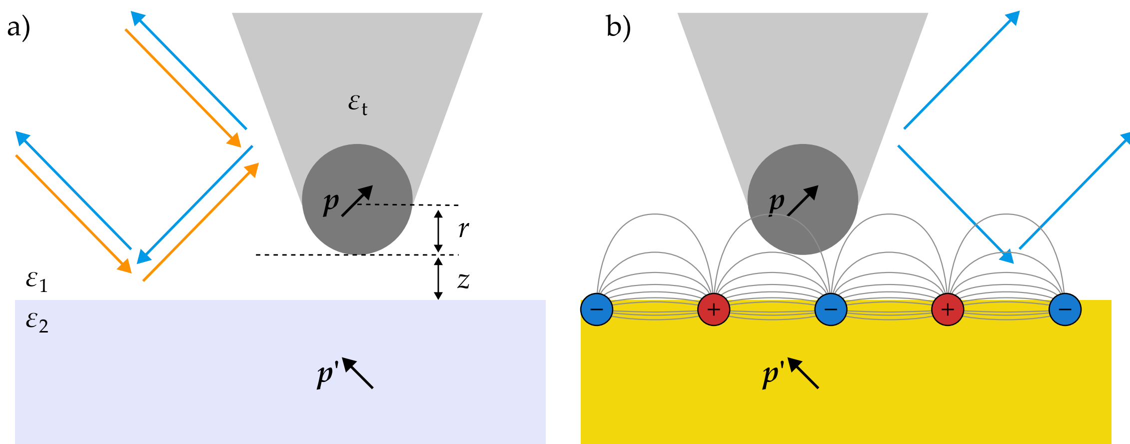


Figure 4.13: Sketch of the tip-sample interaction. In (a) an incident field (orange) excites a dipole in the tip, which then interacts with a sample and induces an image dipole in the sample. (b) depicts a situation where the tip dipole is excited via the electric field of the sample, in this case a plasmonic surface wave. In both cases the scattered light is shown in blue.

Tip-sample interactions are a crucial aspect of s-SNOM measurements, as they determine how the near-field information is converted into a far-field detectable signal. The probing tip acts as an antenna, which scatters the near-field information into the far-field. The tip-sample interaction can be understood by considering the tip as a dipole, which can be excited either by the incident light field or by the near-field of the sample, as illustrated in Fig. 4.13. The far-field excitation of the tip (a) corresponds to measurements done in the reflection mode of the setup, whereas transmission mode measurements usually only indirectly excite the tip via near-field coupling (b) of surface waves for example. In the transmission mode the sample is illuminated with normal incidence, consequently the light is polarized in the sample plane and only the horizontal dipole can be excited in the tip. This dipole usually interacts only weakly with the sample, thus this direct illumination of the tip can usually be neglected. For both cases, the induced dipole moment in the tip interacts with the sample surface, leading to the formation of an image dipole in the sample. The interaction between the tip dipole and the image dipole modifies the effective polarizability of the tip-sample

system, which in turn affects the scattered light detected in the far-field. This section will discuss the tip-sample interaction in more detail, including the analytical point dipole model.

The tip-sample interaction is influenced by several factors, such as the tip shape, size, material, and the tip-sample distance. The tip shape can affect the scattering properties of the tip, as well as the spatial resolution of the measurement. The tip size determines the strength of the near-field interaction, with smaller tip radii leading to stronger interactions and higher spatial resolution. The reason for the field enhancement at the apex of the tip is related to the increased charge density due to the cone-like shape of the tip, often referred to as the lightning-rod effect [74]. The tip material plays an important role, as different materials have different dielectric properties, which can affect the scattering behavior. For example the tip material may exhibit a material resonance depending on the frequency of the electromagnetic excitation. The tip-sample distance is another important factor, as it determines the strength of the near-field interaction and can influence the contrast and sensitivity of the measurement. The reasons for this are twofold: First, the tip-sample distance affects the coupling of the tip dipole with the image dipole in the sample, which modifies the effective polarizability of the tip-sample system. Second, we usually probe optical near-fields that decay exponentially with the tip-sample distance, leading to a strong dependence of the detected signal on the tip-sample distance. The sample topography can also influence the tip-sample interaction, as topography related effects such as edge-darkening [79] or others [80–82] can modify the near-field distribution and affect the detected signal. Many of such effects are related to interference of multiple light paths, including reflections from the sample surface or other nearby structures.

In addition to these factors, the tip-sample interaction is influenced by the polarization of the electric field, a crucial factor often disregarded due to the enhanced out-of-plane coupling. Understanding and controlling the tip-sample interaction is essential for obtaining accurate and reliable s-SNOM measurements. One of the most common and easy to use models to describe the tip-sample interaction is the point dipole model, which will be discussed in the following.

Point Dipole Model

An exact description of the tip-sample interaction is difficult, as many parameters play a role and the geometry of the tip is often not well defined. However, simplified models such as the point dipole model can provide valuable insights into the underlying mechanisms of the tip-sample interaction and help to interpret experimental results. In the following, we will discuss this point dipole model in more detail.

The following derivations are based on [83, 84]. In this model, the tip is approximated as a small sphere with a radius much smaller than the wavelength of the incident light. The sphere is characterized by its polarizability α, which describes how easily the sphere can be polarized by an incident electric field E_{inc}. The tip is located in a distance z to the sample surface, as shown in Fig. 4.14. Here, and in the following, we assume ambient conditions and set the dielectric constant ϵ_1 of the air surrounding the tip to 1. The electric field at the

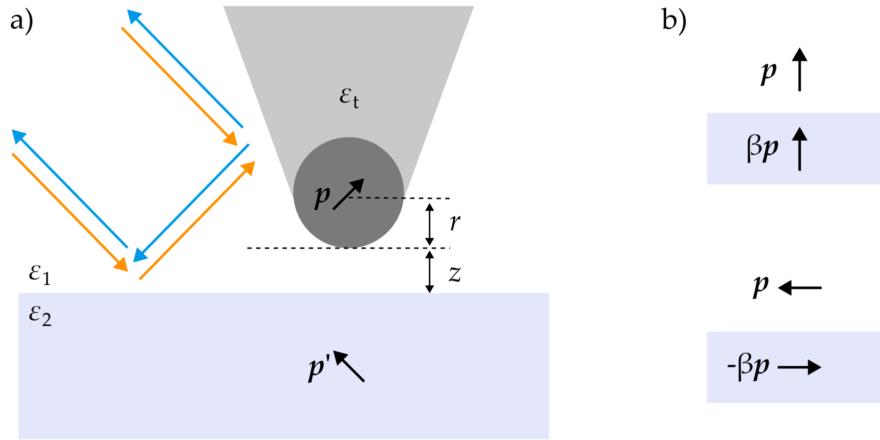


Figure 4.14: (a) Schematic of the tip-sample interaction and formation of the image dipole depending on the tip dipole and tip-sample distance z. The tip dipole can be excited with incident light (orange arrows) or the electric field generated by a sample. The tip emission is indicated by the blue arrows. A part of the incident or emitted light is reflected from the sample surface. (b) The two extreme case polarization orientations of the tip and image dipoles. Adopted from the PhD thesis of Lars Mester [77].

tip position polarizes the sphere, inducing a dipole moment $\boldsymbol{p} = \alpha \boldsymbol{E}_{\text{inc}}$. The polarizability of the sphere can be expressed by

$$\alpha = 4\pi r^3 \frac{\epsilon_t - 1}{\epsilon_t + 2}, \tag{4.26}$$

where r is the radius of the sphere and ϵ_t is the dielectric function of the tip material. The electric field at the tip position can stem from the incident field or from the near-field of the sample. The induced dipole moment of the tip generates a dipole electric field, which interacts with the sample surface. This interaction can be modeled by introducing an image dipole located at a distance $2(z + r)$, below the sample surface. The orientation and magnitude of the image dipole depend on the polarization of the initial dipole moment. For a dipole moment oriented perpendicular to the sample surface (along the z-axis), the image dipole has the same orientation and a magnitude given by

$$p'_z = \beta p_z, \quad \beta = \frac{\epsilon_2 - 1}{\epsilon_2 + 1}, \tag{4.27}$$

where ϵ_2 is the dielectric function of the sample material and β is the electrostatic reflection coefficient. For a dipole moment oriented parallel to the sample surface (along the x-axis), the image dipole has an opposite orientation and a magnitude given by

$$p'_x = -\beta p_x. \tag{4.28}$$

The electric field generated by the image dipole acts back on the tip, modifying its dipole moment. This leads to an effective tip-sample dipole moment, which can be described by an effective polarizability α_{eff}. The effective polarizability takes into account the interaction between the tip and the sample, and can be expressed as

$$\alpha_{\text{eff},\perp} = \frac{\alpha}{1 - \frac{\alpha\beta}{16\pi(r+z)^3}}, \tag{4.29}$$

for the perpendicular dipole orientation, and

$$\alpha_{\text{eff},\parallel} = \frac{\alpha}{1 - \frac{\alpha\beta}{32\pi(r+z)^3}}, \tag{4.30}$$

for the parallel dipole orientation. The scattered electric field from the tip-sample interaction can be expressed as the sum of the dipole field directly emitted from the tip dipole to the detector and the field emitted from the tip dipole which is reflected from the sample towards the detector. The reflected field is modified by the Fresnel far-field reflection coefficient r_p or r_s, depending on the polarization of the incident light. Since the same reflection occurs for a part of the incident field the far-field reflection term is squared. The total scattered electric field can then be expressed as

$$E_{\text{scat},\perp} \propto (1 + c_r r_\perp)^2 \alpha_{\text{eff},\perp} E_{\text{inc}} \tag{4.31}$$

$$E_{\text{scat},\parallel} \propto (1 + c_r r_\parallel)^2 \alpha_{\text{eff},\parallel} E_{\text{inc}}, \tag{4.32}$$

where E_{inc} is the incident electric field at the tip position and $r_\perp$ and $r_\parallel$ are the Fresnel reflection coefficients for perpendicular and parallel polarizations. The factor c_r in front of the reflection coefficient is an empirical factor often used to describe the detected strength of the reflected light relative to the incident light [85]. We include this here with a common value of $c_r = 0.9$, but in general this factor has to be determined for each setup configuration independently. The Fresnel coefficients are given by

$$r_\perp = \frac{n_1 \cos(\theta) - \sqrt{n_2^2 - n_1^2 \sin^2(\theta)}}{n_1 \cos(\theta) + \sqrt{n_2^2 - n_1^2 \sin^2(\theta)}}, \tag{4.33}$$

$$r_\parallel = \frac{n_2 \cos(\theta) - n_1 \sqrt{1 - (\frac{n_1}{n_2} \sin(\theta))^2}}{n_2 \cos(\theta) + n_1 \sqrt{1 - (\frac{n_1}{n_2} \sin(\theta))^2}}, \tag{4.34}$$

where $n_1 = \sqrt{\epsilon_1}$ and $n_2 = \sqrt{\epsilon_2}$ are the refractive indices of the ambient air and the sample, respectively, and θ is the angle of incidence. We assume that theta has a fixed value of 30°, as defined in Section 4.2. Due to the focusing an angular distribution would be more precise, but this is neglected here for simplicity. As an example, we can consider a platinum tip with a

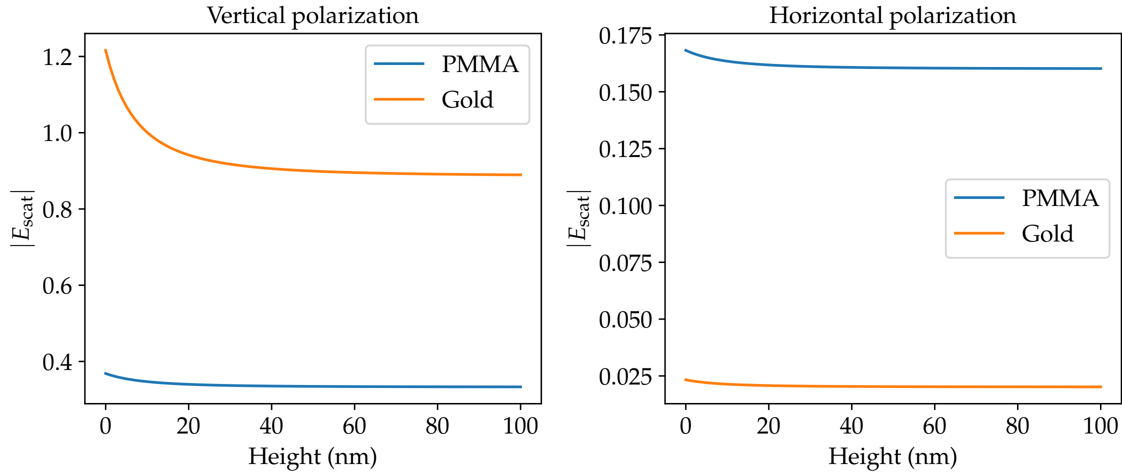


Figure 4.15: Scattered electric field amplitude as a function of tip-sample distance z for a platinum tip with a radius of $r = 30\,\mathrm{nm}$ on a gold sample and a PMMA sample. The results are shown for both dipole orientations: perpendicular (blue) and parallel (orange) to the sample surface, and all amplitudes are normalized to the overall maximum amplitude in a distance of 10 nm to the sample surface, as this corresponds to a typical mean tip position during measurements. The scattered field amplitudes are 1 (gold, perpendicular), 0.02 (gold, parallel), 0.34 (PMMA, perpendicular) and 0.15 (PMMA, parallel) at a tip-sample distance of 10 nm.

radius of $r = 30\,\mathrm{nm}$ and two different sample materials: gold and poly-methyl-methacrylate (PMMA). The dielectric functions of the materials are: $\epsilon_{\mathrm{Pt}} = -99.071 + 15.213i$ [86], $\epsilon_{\mathrm{Au}} = -38.961 + 2.653i$ [43] and $\epsilon_{\mathrm{PMMA}} = 2.196$ [87] at a wavelength of 970 nm. Calculating the scattered electric field as a function of tip-sample distance z for both sample materials and both dipole orientations, we obtain the results shown in Fig. 4.15. The curves are normalized to the overall maximum amplitude at a tip-sample distance of 10 nm, which corresponds to a typical mean tip position during a measurement. These results do not represent an accurate description of a real s-SNOM measurement, as multiple influences such as the actual tip shape, the tapping motion and deconvolution of the detected signals are not taken into account. However, they illustrate the general trends of the tip-sample interaction in s-SNOM measurements. From these results, we can estimate the relative strength of the scattered fields depending on the sample material and dipole orientation. For a gold sample, the perpendicular dipole orientation leads to a significantly stronger scattered field compared to the parallel orientation. In this case the scattered field amplitude drops to approximately 2%. This is due to the strong interaction of the tip dipole with the image dipole in the gold sample, which enhances the effective polarizability of the tip if the initial dipole is oriented perpendicular to the sample, or supresses the effective polarizability if the dipole is oriented parallel to the sample surface. For the PMMA sample, the difference between the two dipole orientations is less pronounced, as the interaction with the image dipole is weaker due to the lower dielectric function of PMMA. Compared to gold, the overall scattered field amplitude

is also significantly lower for PMMA, approximately 35%, as the Fresnel and electrostatic reflection coefficients are much smaller than for gold. For the parallel dipole orientation, the scattered field amplitude on PMMA decreases to approximately 47% compared to the perpendicular orientation. This explains why s-SNOM measurements on metallic samples typically yield a much stronger signal than on dielectric samples, and why the tip is more sensitive to the out-of-plane near-field components. Contrary to metallic samples, the in-plane near-field components of dielectric samples can significantly contribute to the overall scattered field.

Not considered in this model are layered samples. For thin layers the material of the buried layer can have a significant influence on the tip-sample interaction, but here we consider the PMMA layer to be thick enough to neglect such effects. Also, this model only considers flat samples, so any topography induced effects are neglected as well. A slightly better model was introduced with the finite dipole model (FDM) [84], which takes into account the elongated shape of the tip, leading to a more accurate description of the tip-sample interaction. However, so far no description of the tip-sample interaction with the parallel dipole orientation could be found for this model. Therefore, we restrict ourselves to the simpler point dipole model in this work. Recently, a python package called *snompy* has been introduced [85], which allows to calculate the tip-sample interaction with the point dipole model and the FDM for arbitrary sample stacks and tip materials. This package has been used to validate the calculations presented here.

In summary, the point dipole model provides a simple yet effective way to describe the tip-sample interaction in s-SNOM measurements. It allows to calculate the scattered electric field from the tip-sample system, taking into account the tip polarizability, the sample dielectric function, and the tip-sample distance. The model can be used to interpret s-SNOM measurements and extract information about the sample properties, such as its dielectric function and near-field distribution. There are more advanced models that take into account more complex tip geometries and interactions, but the point dipole model remains a widely used and valuable tool in the field of near-field optics.

4.3 Measurement Geometry Considerations

When performing s-SNOM measurements, the geometry of the measurement setup and the orientation of the sample can have a significant impact on the measured signal. This section discusses some important considerations regarding the measurement geometry, including the influence of sample orientation on the polarization components detected in transmission mode measurements.

If one wants to measure polarization components other than the usual out-of-plane (z) component, the sample orientation has to be adjusted accordingly. The reason for this is that the projection of the sample's vectorial near-field components onto the analyzer plane depends on the sample orientation. This is sketched for two example measurement configurations in Fig. 4.16. The sample orientation is especially important for the measurements

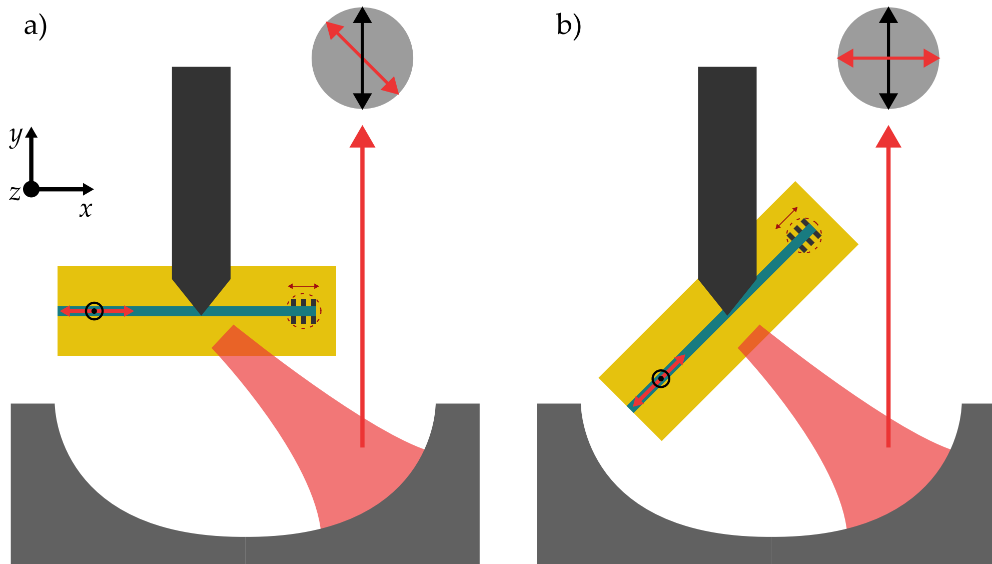


Figure 4.16: Top down sketch of two example measurement configurations to measure waveguides in the transmission mode. The illumination area is marked by the red dashed ring, the tip emission is sketched in red as well. Depending on the rotation of the waveguide, the projections of the waveguides in- and out-of-plane electric field components on the analyzing plane are rotated. In (a) a standard measurement configuration is shown, and in (b) the waveguide is rotated by 45° to separate the in- and out-of-plane components in the analyzer plane.

presented in Section 6.2, where we actively rotate the waveguide by 45° in order to separate the in- and out-of-plane components in the analyzer plane.

In a reflection type measurement, the same principle applies. Here, one usually wants to choose an initial polarization to maximize the excitation of the tip. Depending on the sample and its orientation, the incident field can also cause the excitation of surface modes on edges perpendicular to the projection of the incident polarization. If this is not desired, the sample can be rotated accordingly to minimize such effects. One could however, also invert this and choose a polarization perpendicular to the tip axis and only excite the sample.

When using a pulsed laser source for synchronized s-SNOM measurements in the transmission mode, the direction of the wave propagation on the sample can have a significant influence on the measured signal. This is due to the fact that the pulse overlap between the reference arm and the sample arm changes depending on the direction of wave propagation. The pulse overlap is initially optimized on the excitation spot of the surface wave. The worst case is sketched in Fig. 4.17. Here, the sample and parabolic mirror move to the left in order for the tip to scan the propagating wave, while keeping the illumination fixed on the excitation spot of the wave. This leads to a time delay in the waves excitation, additionally the wave also takes more time to reach the tip position. This leads to a total effective path length

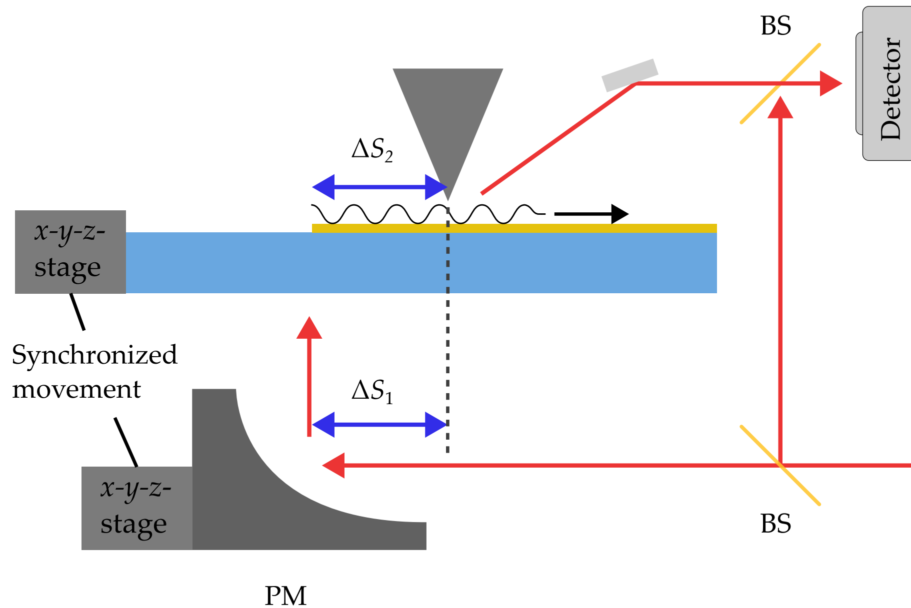


Figure 4.17: Sketch of the influence of the direction of wave propagation on the pulse overlap in transmission mode measurements. In the displayed situation, the sample and parabolic mirror move together in the same direction. If the wave on the sample propagates from left to right the effective path length difference between the sample and reference arm changes increases by $\Delta S_1 + \Delta S_2$, leading to a faster change of the pulse overlap. If the wave propagates from right to left, the effective path length difference decreases by $\Delta S_1 - \Delta S_2$.

difference change of $\Delta S_1 + \Delta S_2$. Where ΔS_1 is the optical path length difference caused by the movement of the parabolic mirror and ΔS_2 is the optical path length difference caused by the surface wave propagation. If the effective index of the surface wave is close to 1, both are very similar. However, if the wave propagates in the opposite direction, the effective path length difference changes by $\Delta S_1 - \Delta S_2$. This effectively cancels the influence of the wave propagation on the pulse overlap, leading to a more stable measurement. Therefore, it is advisable to always measure in transmission mode with the wave propagating to the left. One can also measure perpendicularly to the discussed cases, here the mirror movement does not lead to a path length difference due to the parabolic mirror geometry, only the wave propagation leads to a change in pulse overlap.

An example of the discussed path length variation is shown in Fig. 4.18. Here, a long s-SNOM scan of an isolated grating incoupler is shown. The measurement is performed in transmission mode with the synchronization of the lower parabolic mirror to the sample table movement. The amplitude image clearly shows that the amplitude of the excited surface plasmon polaritons (SPPs) decays faster on the right side of the grating compared to the left side. Additionally, in the raw phase image a much steeper phase gradient is visible on

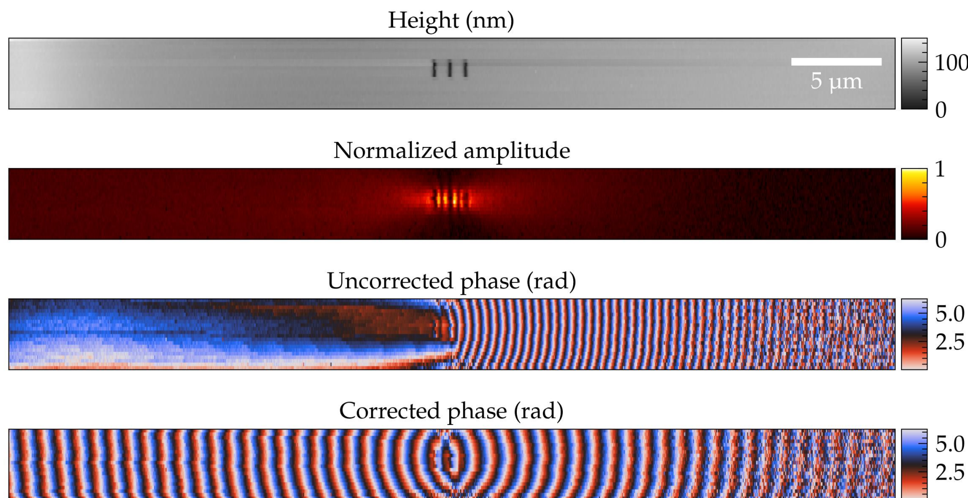


Figure 4.18: Example transmission mode s-SNOM measurement of an isolated grating incoupler milled into a 60 nm thick gold film on a glass substrate. The measurement is performed with the illuminating lower parabolic mirror being synchronized to the movement of the sample table. Shown is the leveled sample topography, the amplitude and phase data without sync-correction and with sync-correction applied of the third harmonic demodulated signal.

the right side compared to the left side. This is caused by the path length changes discussed before. By applying the sync-correction described in Section 4.4.1, the phase image becomes symmetric again, but the amplitude image remains unchanged. This shows the importance of considering the measurement geometry and applying appropriate corrections to obtain accurate and reliable s-SNOM data. If we would now try to extract the propagation losses of the travelling SPPs from the amplitude data we would get different results depending on the propagation direction. Even the direction where the path length changes only slowly would lead to an overestimation of the losses, as the slow changes in the pulse overlap lead to a slight decrease in the detected amplitude over distance.

4.4 Measurement Corrections

In this section, we discuss various measurement corrections that can be applied to s-SNOM data to improve the quality and reliability of the results. These corrections include the correction of the phase gradient introduced by the movement of the lower parabolic mirror in transmission mode measurements, pixel scaling, height leveling, slow amplitude and phase drift corrections, the calculation of real part data from amplitude and phase information, and Gaussian filtering for better visualization.

4.4.1 Sync-Correction

As already discussed in Section 4.3, the movement of the lower parabolic mirror in transmission mode measurements adds a phase gradient to the measured phase data. This phase gradient arises because the movement of the parabolic mirror changes the optical path length in the tip-sample arm, leading to a time-dependent phase shift in the detected signal. To correct for this phase gradient, we can apply a sync-correction to the measured phase data. The sync-correction involves calculating the expected phase shift based on the known movement of the parabolic mirror and subtracting this phase shift from the measured phase data. This correction can be implemented in the data analysis pipeline, ensuring that the corrected phase data accurately reflects the true near-field phase distribution without the influence of the parabolic mirror movement. The expected phase shift $\Delta\phi$ can be calculated using the formula

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta S_1\,, \tag{4.35}$$

where ΔS_1 is the change in optical path length due to the parabolic mirror movement and λ is the wavelength of the incident light. The optical path length can be directly calculated from the pixel coordinates. Only the movement of the parabolic mirror along the x-axis needs to be considered here, as the movement along the y-axis does not change the optical path length in the tip-sample arm. When the fast axis in a scan is not aligned with the x-axis, the rotation angle relative to the x-axis has to be taken into account when calculating ΔS_1.

4.4.2 Pixel Scaling

When measuring large area scans, it is often advisable to choose a non symmetric pixel size in order to reduce the overall scan time. When imaging the scans later, the individual pixels then have to be scaled accordingly to represent the correct physical dimensions. For this one should use an appropriate x and y size, such that the pixels can be scaled by an integer number. Otherwise, the pixels need to be interpolated to achieve the correct dimensions, which can lead to artifacts in the data. If an interpolation is necessary, it is advisable to apply the interpolation to complex data (amplitude and phase) instead of separately interpolating amplitude and phase data. This helps to avoid phase discontinuities that can arise when interpolating phase data alone.

4.4.3 Height Leveling

In s-SNOM or AFM measurements, the sample surface may not be perfectly flat, leading to height variations across the scanned area. Additionally, the sway and bending of the tip or cantilever can introduce further height variations, especially when the sample changes the scanning direction. To correct for these height variations, a height leveling procedure can be applied to the topography data obtained from the AFM measurements. Standard procedures such as as a three point plane leveling or polynomial fitting can be used to remove the overall tilt or curvature of the sample surface. Slow temporal drifts in height

can also be corrected by comparing lines along the slow scan axis and adjusting the height values accordingly. In order to properly apply such a slow drift correction, it is important to ensure that the scan includes sufficent reference areas without significant topographical features. Most height leveling of the data used in this thesis was performed using the open source software *Gwyddion*. More advanced height leveling techniques like a nonlinear line by line correction using one or two reference areas was employed using a custom python analysis script, detailed in Appendix A.4.2.

4.4.4 Amplitude and Phase Drift Correction

Similar to the slow height variations discussed in the previous section, slow drifts in amplitude and phase can also occur during long s-SNOM measurements. These drifts can be caused by various factors, such as temperature fluctuations, mechanical instabilities, or changes in the laser power. To correct for these drifts, a slow amplitude and phase drift correction can be applied to the measured data. This correction involves identifying reference areas in the scan that are expected to have constant amplitude and phase values, and using these areas to calculate the drift over time. The calculated drift can then be subtracted from the entire dataset to obtain corrected amplitude and phase values. The reference areas should be chosen carefully to ensure that they are representative of the overall measurement and do not contain significant features that could bias the drift calculation. In this thesis, a custom python analysis script was used to perform slow amplitude and phase drift corrections, as described in Appendix A.4.2.

4.4.5 Gaussian Filtering

Raster scanned images in general correspond to a point cloud measurement. Therefore, a pixelated image is only an approximation of the actual continuous field distribution. Additionally, in s-SNOM images the probing tip does not only scatter information of the point directly below it but of a small volume around the tip-sample interaction point. To better visualize the measured data, the larger pixels can be divided in smaller sub-pixels and a Gaussian filter can be applied to smooth the data [2]. This helps to reduce pixelation artifacts and helps the visual and numerical analysis of the measured near-field distribution. The Gaussian filter should be applied to complex data in order to avoid phase discontinuities. The width of the Gaussian filter should be chosen carefully to balance between smoothing the data and preserving important features. The scaling and smoothing of the data helps not only to improve the visual appearance, but also to supress noise and help in further analysis steps such as shifting scan lines with higher precision. The Gaussian filtering is not always necessary, but can be a useful tool for better visualization of the measured data. The modified images in this thesis are labeled accordingly in the figure captions.

[2] Other filters, like a linear interpolation, can be used as well.

4.4.6 Calculating the Real Part Data

In some cases, it can be beneficial to represent the measured near-field data in terms of its real part, rather than using separate amplitude and phase information. The real part of the complex-valued near-field data can provide a more intuitive understanding of the spatial distribution of the near-field, as it directly corresponds to the electric field distribution at the sample surface. Especially for traveling surface waves, the real part of the near-field can give a more direct visualization of the wave propagation. To calculate the real part from amplitude and phase information, we can use the following formula

$$E_{\text{real}} = A \cdot \cos(\phi)\,, \tag{4.36}$$

where E_{real} is the real part of the near-field, A is the amplitude, and ϕ is the phase.

CHAPTER 5

Chiral Couplers

Light posesses a variety of interesting properties, of which some have been known for centuries, while others have only been discovered in the last decades. Among these newly discovered properties are the spin- and orbital angular momentum of light, which give rise to spin-orbit interactions (SOIs) [88]. These effects are typically very weak in free space, but can be significantly enhanced in the presence of sub-wavelength inhomogeneities. Due to the advances in nanofabrication techniques over the last decades, it has become possible to engineer structures at the nanoscale that can manipulate light at subwavelength scales, resulting in stronger light-matter interactions and enhanced SOI effects. The SOI of light involves the coupling between different forms of angular momenta, These are the intrinsic spin angular momentum (SAM) of photons, associated with their polarization state, the intrinsic orbital angular momentum (IOAM), corresponding to optical vortices and the extrinsic orbital angular momentum (EOAM), related to the beam trajectory [88]. These couplings are analogous to the spin-orbit interactions observed in relativistic quantum particles [89, 90] and electrons in solids [91, 92]. Several examples of SOIs have been observed in optics, such as the polarization-dependent transverse shift experienced by a circularly polarized beam upon reflection or refraction at an interface. This effect is also known as the Imbert-Fedorov effect [93–95] and is related to the Goos-Hänchen effect [96]. A second example is the generation of optical vortices when tightly focusing circularly polarized beams [97, 98]. Additionally, SOI effects have been observed in plasmonic waveguide systems as well [33–36]

Many of these SOIs become significant only at subwavelength scales and have been largely neglected in traditional optics. However, they become particularly relevant in the context of optical near-fields, where the spatial variations of the electromagnetic field are pronounced. A promising platform to study and utilize these effects are surface plasmon polaritons (SPPs). SPPs exhibit strong field confinement at subwavelength scales, leading to large field gradients and significant SOIs. Notably, SPPs possess an inherent transverse spin, which is oriented perpendicular to the direction of propagation. The direction of this transverse spin is locked to the propagation direction, allowing for chiral coupling effects.

For example, a rotating dipole placed above a metal surface can preferentially excite SPPs in one direction along the surface, depending on the handedness of the dipole rotation [99]. A similar behavior was also observed for the evanescent fields around an optical fiber [100].

In this chapter, we present our work on chiral couplers that utilize the photonic spin-Hall effect to achieve directional excitation of plasmonic strip-waveguide modes based on the polarization state of the incident light. We discuss the design, fabrication, characterization, and simulation of these couplers, as well as measurements demonstrating their performance. Finally, we outline potential future directions for this research, based on the use of orbital angular momentum states of light.

5.1 Photonic Spin-Hall Effect

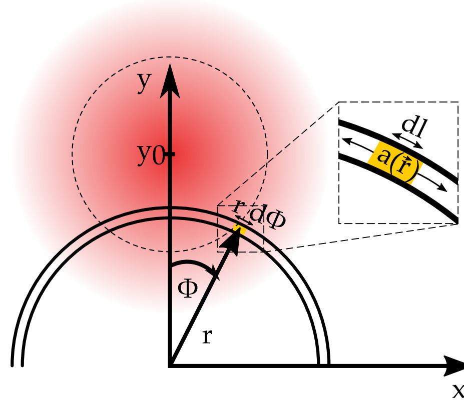


Figure 5.1: Sketch of the SOI interaction and excitation of two counter propagating modes. The half ring corresponds to the curved waveguide bend, the insert shows a closeup of an infinitessimal piece of the waveguide. The Gaussian beam is displayed in red, the displacement y_0 of the laser spot along the y-axis is important for small beam sizes. r is the radius of curvature of the coupler and Φ the azimuthal angle around the z-axis.

The chiral couplers presented in this chapter exploit the photonic spin-Hall effect to achieve unidirectional coupling of light into waveguide modes based on the polarization state of the incident light. The underlying principle of this effect is based on spin angular momentum transfer from photons, along the z-axis with circular polarization, to the extrinsic orbital angular momentum of waveguide mode propagating on a curved trajectory in the x-y-plane [33]. The derivation by Y. Lefier et al. is based on a plasmonic slot waveguide but should also apply to strip waveguides, as both are closely related. The presented argument is, that the impinging photons carry SAM of $\sigma\hbar$ per photon for right- and left-circular polarization, respectively, where $\sigma = \pm 1$ defines the polarization helicity. A curved single mode waveguide can support two counter-propagating modes, each carrying EOAM

of $\pm\hbar\beta r$ per photon along the z-axis, where β is the propagation constant of the mode and r is the radius of curvature of the waveguide. These two modes have an azimuthal phase dependence of $\exp(\pm i\beta r\phi)$, where ϕ is the azimuthal angle around the z-axis. These modes can be viewed as eigenmodes of the angular momentum operator $L_z = -i\hbar\partial/\partial\phi$, and carry momentum of $\pm\hbar\beta r$ per photon. The incoupled light is the superposition of these two modes, with amplitudes a_r and a_l for the right- and left-propagating modes, respectively. The probabilities of coupling into these modes are given by $|a_r|^2$ and $|a_l|^2$. Therfore, the extrinsic angular momentum of the guided mode along the z-axis is given by $(|a_r|^2 - |a_l|^2)\hbar\beta r$. Assuming elastic coupling and conservation of angular momentum along the z-axis, the two probability amplitudes can be derived as [33]

$$|a_r|^2 = \frac{1+\sigma/\beta r}{2} \quad \text{and} \quad |a_l|^2 = \frac{1-\sigma/\beta r}{2}. \tag{5.1}$$

For linear polarization ($\sigma = 0$), or a large radius of curvature ($r \to \infty$), the coupling is symmetric, with equal probabilities of coupling into the right- and left-propagating modes. However, when the condition $r = \sigma/\beta$ is satisfied, the coupling becomes unidirectional, with all the incident photons coupling into either the right- or left-propagating mode, depending on the polarization helicity σ. This geometrical condition also corresponds to a resonance condition for a closed circular waveguide of radius r, where the accumulated phase for one roundtrip of the guided mode equals 2π.

It was also shown, that the directionality of the coupling depends on the position and size of the excitation beam [33], more precisely on the spatial overlap of the excitation beam with the waveguide mode. An uneven illumination of the curved waveguide can lead to a local variation of the geometric resonance condition along the waveguide, resulting in a changed directionality. This effect is not relevant for our measurements, as our laser spot size ($\approx 3\,\mu$m) is much larger than the investigated coupler radii ($< 300\,$nm). Thus, we can assume a uniform illumination of the coupler structure and the general resonance condition given above.

The analytical model presented here provides an intuitive understanding of the resonance condition for unidirectional excitation of chiral couplers. However, in practice, a number of factors can influence the actual coupling efficiency and directionality. First of all, the coupler structure is not a closed ring, but a half-ring connected to two straight waveguides. Therefore, the simple analytical model does not fully apply, because the straight waveguides look like a continuation of the ring with changing radius, when viewed from the apex of the coupler. Since it is not possible to only excite the ring part where the radius is constant, this will affect the directionality of the coupling. The second reason is that the overall coupling efficiency depends on the overlap of the excitation beam with the waveguide mode. For surface waves, such as surface plasmon polaritons (SPPs), the wavevector is always larger than that of free-space photons, making direct excitation impossible. Subwavelength features such as gratings or nanoantennas are typically used to provide the necessary momentum boost for coupling. In this case the curvature of the waveguide itself provides the required momentum

matching for coupling. This implies that the coupling efficiency will depend on the radius of curvature as well and decreases for larger radii. This effect is in principle independent of the coupling direction, meaning it does not alter the directionality directly, but i may strengthen the impact of the previous argument. The exact dependence of the coupling efficiency on the radius of curvature, waveguide width, spot size and position is challenging to predict analytically, and is best investigated using simulations and measurements.

5.2 Design

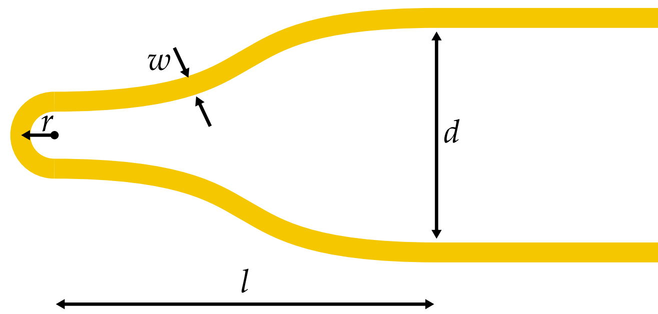


Figure 5.2: Sketch of the chiral coupler design. The parameters are the waveguide width w, the coupler radius r, the length of the bending section l and the distance of the two waveguides after the bend d.

The desing of the chiral coupler is based on a half-ring waveguide structure connected to two straight strip waveguides, as shown in Figure 5.2. In order to reduce the coupling of the two waveguides, an additional s-shaped bend is introduced after the half-ring section, increasing the distance d between the two waveguides. As long as this bend is shallow enough no measureable influence could be observed. The main design parameters are the waveguide width w and the coupler radius r. The length of the bending section l is chosen to be 4 µm and the distance increase is set to 2 µm. The total center to center distance d of the waveguides is then the distance increase plus the coupler diameter. The mode index of such a strip waveguide is around $n_{eff} = 1.45$ for a width of 120 nm at a wavelength of 1 550 nm [101]. Therefore, the resonant radius for unidirectional coupling is expected to be around $r = \frac{\lambda}{2\pi n_{eff}} \approx 170\,\text{nm}$.

5.3 Simulation

The chiral couplers are simulated using finite element method (FEM) simulations in *Comsol Multiphysics* with the *Waveoptics* module. A 60 nm thick gold layer is modeled on a glass substrate with a refractive index of 1.5. The dielectric data for gold is taken from Johnson

and Christy [43]. The simulation geometry is surrounded by scattering boundary conditions to avoid reflections. The chiral coupler structure is excited from the bottom with a circularly polarized Gaussian beam with a waist of 3 µm positioned at the apex of the half-ring coupler at a wavelength of 1 550 nm. Different coupler radii and waveguide widths are simulated. The complex electric field is exported in a cutplane 10 nm above the gold surface for the out-of-plane component E_z. From the complex data, the amplitude, phase and real-part data can be calculated. The amplitude data is used to extract the intensity in the two waveguides after the coupler, allowing to determine the directionality of the coupler. The resulting simulation results for a variation of the input polarization state are shown in Fig. 5.3 (c-f). The circular arrows, indicating the helicity of the polarization, are defined from the point of view of the emitter. For the two circular polarization states, a clear directionality can be observed, with most of the signal going into one waveguide depending on the handedness of the polarization. For linear polarization, the coupling is symmetric, with equal intensities in both waveguides. The main difference here being the relative phase difference between the two waveguides. For the linear polarization perpendicular to the waveguide, the two waveguides are out of phase, while they are in phase for the linear polarization parallel to the waveguide.

Fig. 5.3 (a) shows the amplitude distribution in a cut-plane across a single isolated waveguide excited from the bottom with a Gaussian beam linearly polarized perpendicular to the waveguide. This excitation leads to a dipole-like field distribution in the waveguide. In (b) the mode distribution of the single mode waveguide is displayed. The field is mainly concentrated at the edges of the metal strip waveguide, with a higher concentration at the lower gold/glass interface. The orientation of the field lines indicates a primarily out-of-plane field at the top gold/air interface. In combination with the higher field concentration at the edges, the waveguides will show the strongest signal on the edges in the near-field measurements. The reason for this is the tips sensitivity to the out-of-plane component discussed in Section 4.2.4. In the amplitude images of the chiral coupler (c-f) an oscillation can be seen, which decays along the waveguide. This stems from an interference of the waveguide mode with the dipolar excitation of the waveguide from direct laser illumination. To clarify this effect, a simple numerical model has been implemented in python, which simulates the interference of the waveguide mode with a perpendicular dipolar excitation along the waveguide.

In this numerical model, the waveguide mode is represented as a propagating wave with an exponential decay along the waveguide. The dipolar excitation is modeled as a perpendicular dipolar amplitude and phase distribution along the waveguide, which decays with the Gaussian amplitdue distribution of the excitation beam along the waveguide. To better represent the measurements, the pixels between the two sides of the waveguide are calculated from a Gaussian superposition of the two sides. The total field along the waveguide is then given by the superposition of the mode and dipolar field. The results of this model are shown in Fig. 5.4, where the amplitude, phase, real part and imaginary part of the total field along the waveguide are displayed. The oscillations in the amplitude are the result of the interference of the two fields, leading to constructive and destructive

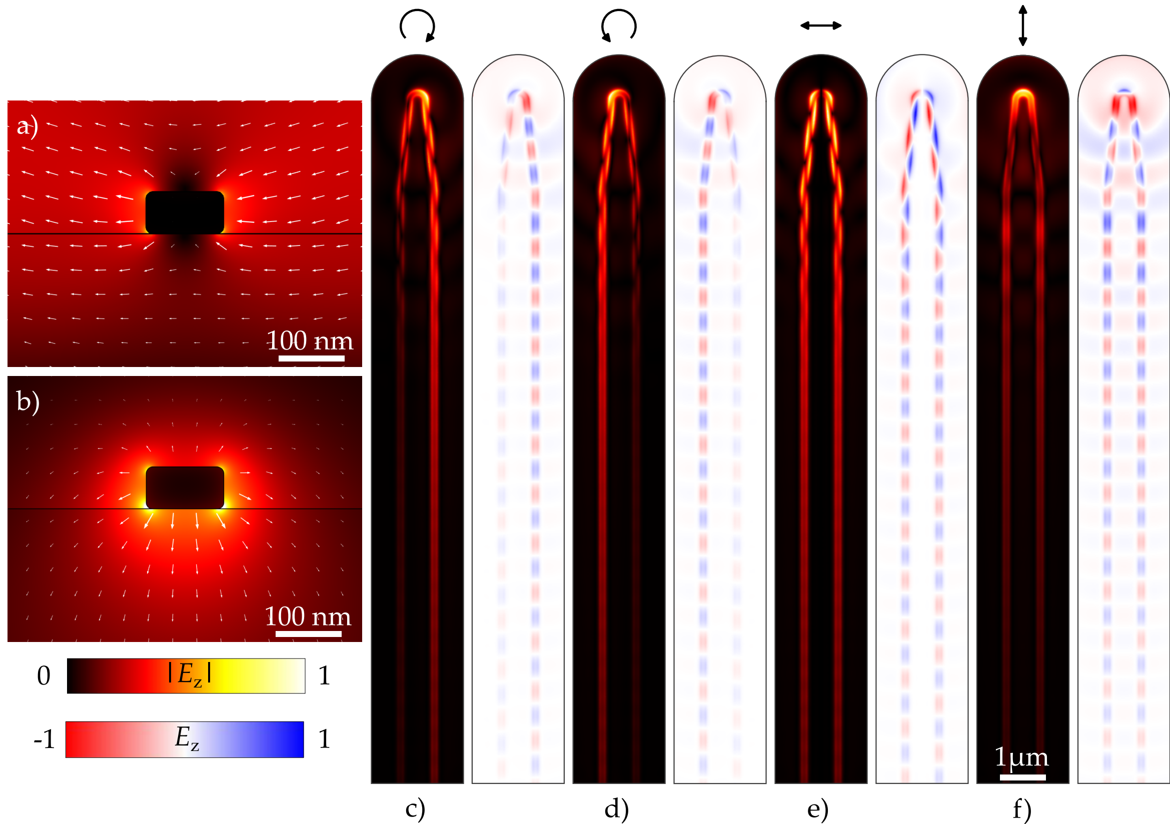


Figure 5.3: Simulation results of the chiral coupler for different excitation conditions. (a) Dipolar excitation of a single waveguide with linear polarization perpendicular to the waveguide. (b) Mode profile of the strip waveguide. The waveguide has a width of 120 nm and a height of 60 nm. (c-f) Chiral coupler with a radius 170 nm, excited with different polarization states : (c) right-circular, (d) left-circular, (e) linear perpendicular, (f) linear parallel. Shown is the amplitude of the out-of-plane electric field component E_z and its realpart 10 nm above the gold surface.

interference along the waveguide. The phase shows initially a phase jump of π at the first minimum of the amplitude, because at this point the dipolar excitation is equally as strong as the waveguide mode. This phase shift decreases along the waveguide, as the intensity distribution of the Gaussian beam decays faster than the waveguide mode. Interestingly, the real-part data shows no direct sign of the interference, as the dipolar distribution, decaying with the Gaussian laser distribution, only leads to an amplitude factor to the exponential decay of the waveguide mode. The reason for this is that in the model we have chosen a phase offset of $\pm\pi/2$ for the upper and lower side of the dipolar distribution. Therefore, the real part is symmetric for both sides. The imaginary part however is not symmetric, as it can be seen in (d).

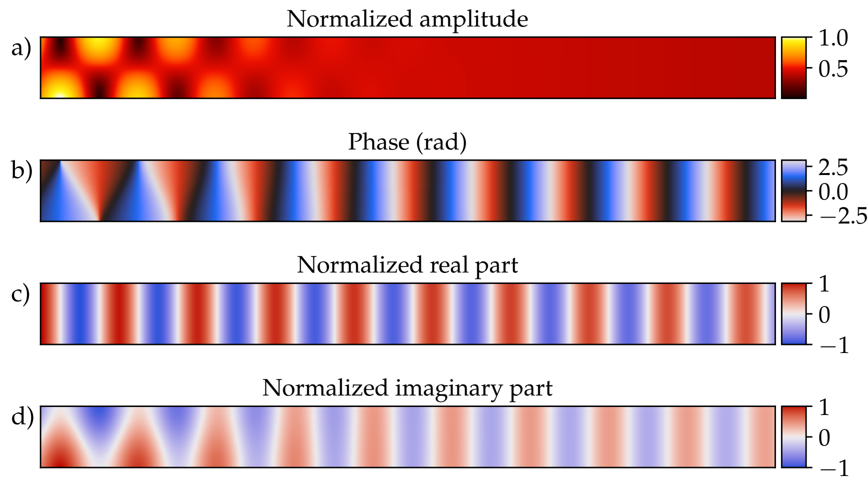


Figure 5.4: Numerical model of the interference of a waveguide mode with a dipolar excitation along the waveguide. Shown are the amplitude (a), phase (b) real-part (c) and imaginary part (d) of the total field along the waveguide. The pixel size is not symmetric and arbitraryly chosen for better visibility.

5.4 Fabrication

The chiral couplers and waveguides are fabricated using positive tone electron beam lithography (EBL) with poly-methyl-methacrylate (PMMA) as described in Section 3.1.3. For the default waveguide width of around 120 nm a design width of 60 nm is used. The structures are written as continous lines with a spacing of 5 nm between the lines and a line dose factor of 1. After the development of the resist, 5 nm of chromium is evaporated as an adhesion layer, followed by 60 nm of gold.

For the parameter variation of the width, the design width is varied from 20 to 200 nm in steps of 20 nm. For the coupler radius variation, the radius is varied from 80 to 300 nm in steps of 20 nm. For the bending variation, the distance increase is varied from 0 to 2 000 nm in steps of 400 nm. An example scanning electron microscope (SEM) image of fabricated samples with different coupler radii is shown in Fig. 5.5.

5.5 Measurement

The chiral couplers are characterized using the scattering-type scanning near-field optical microscope (s-SNOM) setup described in Section 4.2. The samples are excited from the bottom with a circularly polarized Gaussian beam with a waist radius of around 3 µm at a wavelength of 1 600 nm. Our laser system frequently exhibits more noise around a wavelength of 1 550 nm, thus the slightly changed illumination wavelength. The polarization state of the input is controlled with a linear-polarizer, $\lambda/2$ waveplate and a dedicated circular

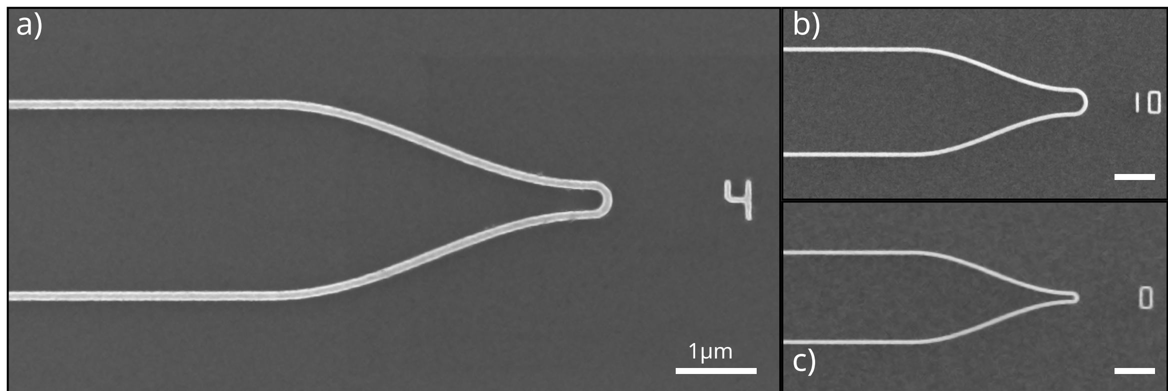


Figure 5.5: SEM images of the samples for different coupler radii (a: 160 nm, b: 80 nm, c: 280 nm) and a waveguide width of 120 nm. The scalebar is 1 µm.

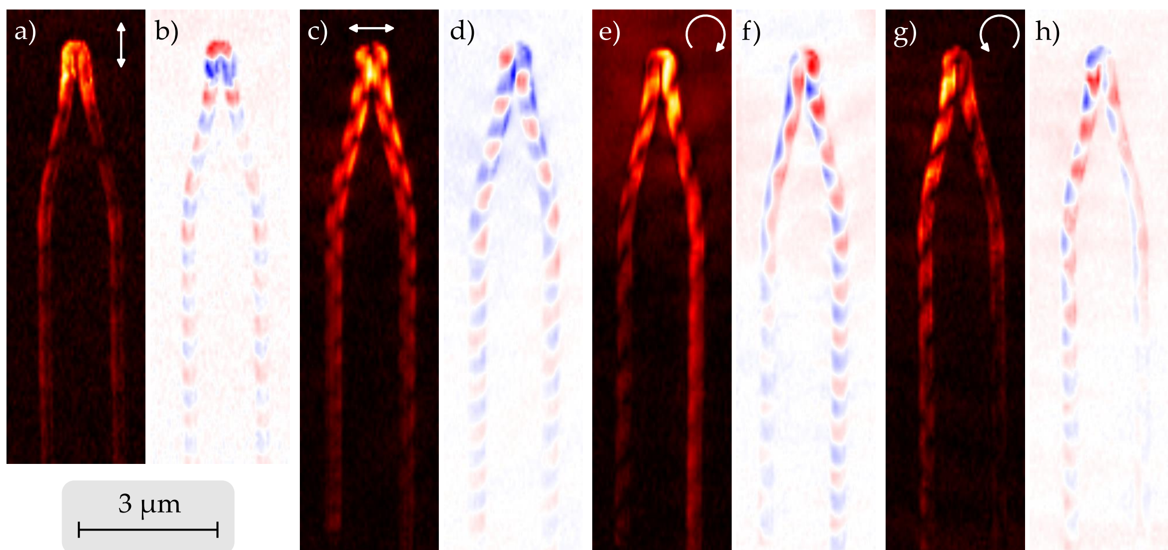


Figure 5.6: Near-field measurements of the chiral coupler for different excitation conditions. The coupler radius is 170 nm and the waveguide width is 120 nm. Shown is the amplitude (a, c, e, g) and reconstructed real-part (b, d, f, h) of the fourth demodulation order.

polarizer (*Thorlabs*, CP1L1550 or CP1R1550). The near-field is recorded in transmission mode, using the synchronized movement of the lower parabolic mirror to keep the excitation spot fixed on the coupler apex during the scan. The out-of-plane electric field component E_z is recorded using standard platinum coated tips and filtered with an analyzer to reduce stray light.

First, we perform measurements of the chiral coupler for different excitation polarization states, similar to the simulations shown in Fig. 5.3. For this we use a $\lambda/2$-waveplate for the linear polarization states and two dedicated circular polarizers for the circular polarization states. The recorded amplitude and reconstructed real-part of the fourth demodulation order are shown in Fig. 5.6. The two linear polarization states (a-d) show a symmetric coupling into both waveguides, with a phase difference of 0 for the parallel polarization (c, d) and π for the perpendicular polarization (a, b). This is in good agreement with the simulation results in Fig. 5.3 (c-f). The beating on the waveguide is absent for the linear polarization parallel to the waveguide, as the dipolar excitation from direct illumination is minimized in this configuration. For the linear polarization perpendicular to the waveguide, the dipolar excitation is maximized, leading to a strong interference effect. For the circular polarization states (e-h), a clear directionality can be observed. The measurements are not directly comparable as the experimental conditions differ slightly, but qualitatively the results agree well with the simulations. All follwoing measurements are performed with the right-handed circular polarizer, as indicated by the clockwise arrow in (e, f).

Next, we investigate the influence of the s-bend on the waveguide mode. The bending is used to separate the two waveguides after the coupler, reducing the coupling between the two waveguides. A too large bending can introduce additional losses, and may act as a secondary coupling feature, similar to the half ring coupler. Furthermore, the measurement time increases linearly with the distance increase, as the total measurement area increases. A sequence of measurements for various distances after the bending is shown in Fig. 5.7. From qualitative observation, no significant influence of the bending can be observed, as the directionality and intensity in both waveguides remains similar for all distances. Therefore, we will use an intermediate bending, which increases the distance to 2 µm plus the coupler diameter for the following measurements. This allows us to ignore the coupling between the two waveguides after the coupler, while keeping bending losses and the measurement time reasonable.

After the bending, the width of the waveguides is investigated. Some exemplary measurements are depicted in Fig. 5.8 for a coupler radius of 170 nm and a waveguide separation of 2 340 nm. A clear directionality can be observed for all widths, with most of the signal going into the right waveguide. For widths above 200 nm, the coupler can not be considered as a ring any more and the waveguide starts to support higher order modes. This sets an upper practical limit for the waveguide width.

Finally, the coupler radius is varied. Some representative measurements are shown in Fig. 5.9 for a waveguide width of 120 nm and a waveguide separation of 2 µm plus the coupler diameter. From the simple geometrical model we expect a radius dependent resonance, with the optimal radius around 170 nm. A clear directionality can be observed

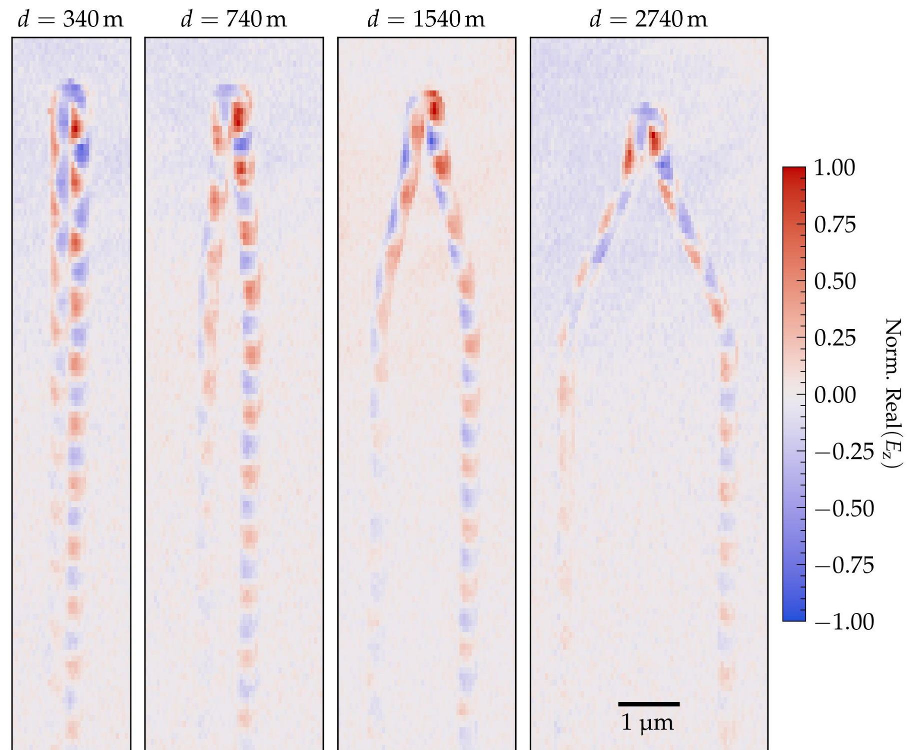


Figure 5.7: Measurements for different waveguide separations. The coupler radius is 170 nm and the waveguide width is 120 nm. Shown is the reconstructed real-part of the fourth demodulation order.

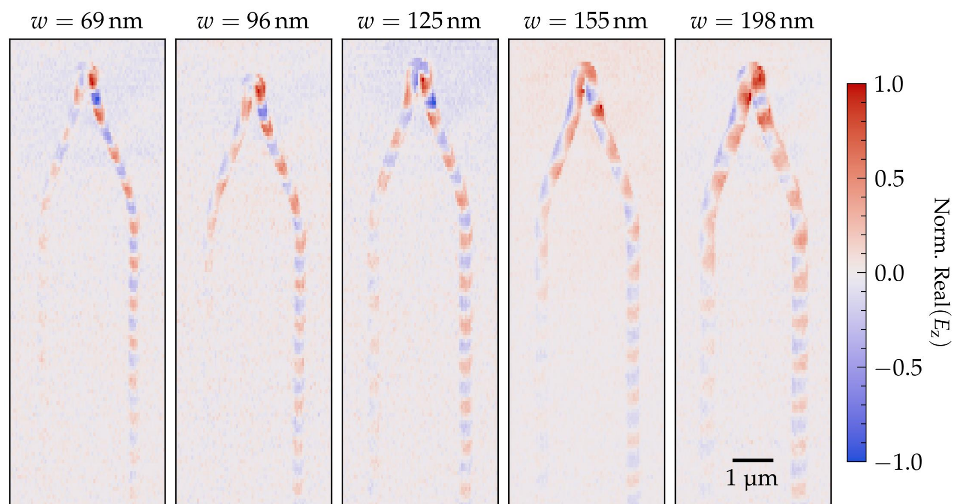


Figure 5.8: Measurements for different waveguide widths. The coupler radius is 170 nm and the waveguide separation is 2 340 nm. Shown is the reconstructed real-part of the fourth demodulation order.

for all radii, with most of the signal going into the upper waveguide, but no clear resonant behavior can be observed.

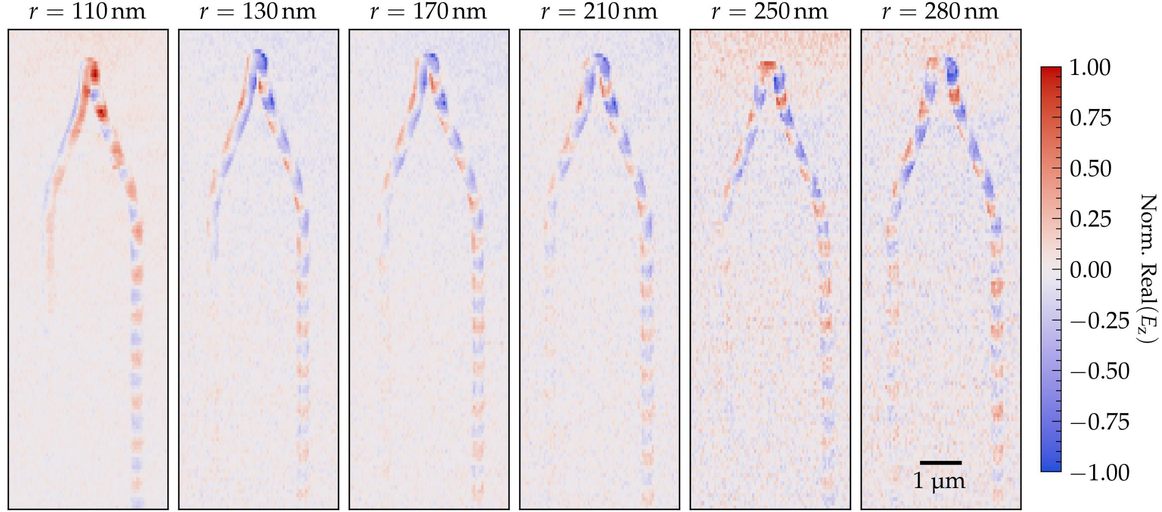


Figure 5.9: Measurements for different coupler radii. The waveguide width is 120 nm and the waveguide separation is 2 μm plus the coupler radius. Shown is the reconstructed real-part of the fourth demodulation order.

We observe a clear directionality in all measurements, with most of the signal going into the right waveguide for right-handed circular polarization. The measurements show a low signal-to-noise ratio (SNR) due to the weak signal from the waveguides and the strong background caused by the transparent substrate. Due to the low polarizability of the substrate it should not be visible compared to the gold waveguides, but it leads to large amounts of stray light. Stray light reflected from the probing tip can cause background interferences. This makes it challenging to extract quantitative values for the directionality. Our approach to quantitatively analyze the directionality will be discussed in the next section.

5.6 Directionality

The main property of the chiral couplers is the directionality of the coupling into the two waveguides after the coupler. We define the directionality η as

$$\eta = \frac{|a_\mathrm{r}|^2}{|a_\mathrm{r}|^2 + |a_\mathrm{l}|^2}, \tag{5.2}$$

where a_r and a_l are the amplitudes in the right and left waveguide, respectively. The directionality can take values between 0 and 1, where 1 corresponds to all the intensity

going into the right waveguide, 0 to all the intensity going into the left waveguide and 0.5 to equal intensities in both waveguides. We chose this definition, as all measurements with parameter variations have been performed with the same right handed circular polarization state.

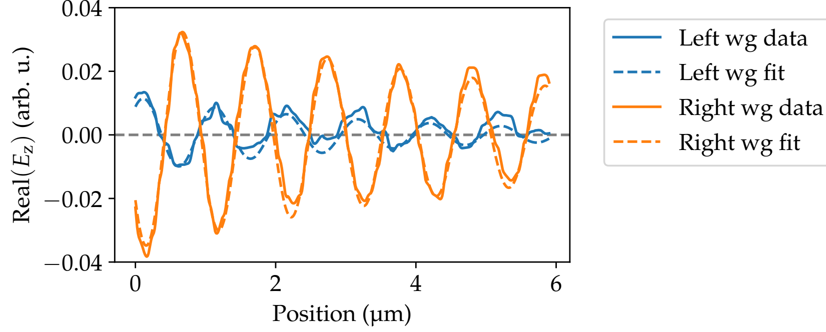


Figure 5.10: Example cutlines of the averaged real part along the two waveguides of the chiral coupler with radius 170 nm displayed in Fig. 5.9. The data is fitted according to Eq. (5.3).

To this end, the amplitude data from the two scan directions of each measurements [1] are averaged to reduce noise. Then, the real-part and imaginary part data is reconstructed from the data of the fourth demodulation order, using the averaged amplitude and original phase data. We use the real part and imaginary part data to reduce the impact of the beating pattern as discussed in Section 5.3. The position of the waveguides is then determined from Gaussian fits to each horizontal line across the waveguide height data. Using the waveguide positions, the real-part and imaginary-part data is extracted along the waveguides, and averaged over the width of the waveguide. This results in two line profiles along each waveguide, for the real-part and imaginary-part respectively. The individual line profiles are then fitted with an exponentially damped oscillation like

$$f(x) = Ae^{\frac{-x}{(2L_{prop})}} \cos(k_{spp}x + \phi), \tag{5.3}$$

where A is the amplitude, L_{prop} the propagation length, k_{spp} the wavevector of the SPP and ϕ a phase offset. An example of cutlines extracted and fitted like described here is shown in Fig. 5.10. The only relevant parameter for the directionality extraction is the amplitude A, as this corresponds to the amplitude of the waveguide mode at the beginning of the waveguide after the coupler. From the fitted amplitudes A_{right} and A_{left} for the right and left waveguide, the directionality can then be calculated using Eq. (5.2).

A similar procedure is used for the simulation data, where the amplitude ($|E_z|$) data is exported in a cut-plane 10 nm above the gold surface.

The resulting directionality values for a variation of the coupler radius and waveguide

[1] Each scan line is scanned forwards and backwards as described in Section 4.1.

width are shown in Fig. 5.11. A trend can be observed in both the measurements and simulations. The directionality decreases for larger coupler radii and waveguide widths. However, no clear resonance condition can be observed, contrary to the prediction of the simple analytical model. As discussed before, this is due to the influence of the waveguides and the coupling efficiency dependence on the radius of curvature. The simulations and measurements show a reasonable agreement in the overall trend of the directionality with varying parameters. Although, the directionality extracted from the measurements is consistently overestimated, which may be caused by the strong background and weak signal in the left waveguide. The near-field measurements are quite sensitive to tip conditions and environmental factors. Especially the largely transparent substrate can lead to substantial background signals, which can influence the directionality.

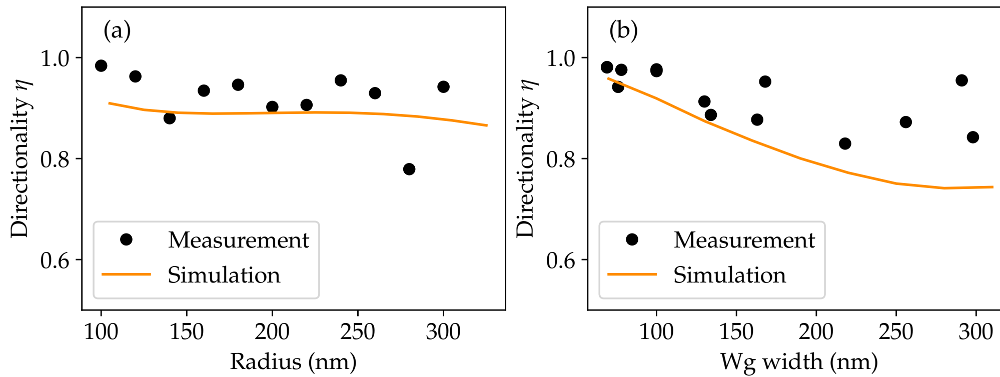


Figure 5.11: Directionality from the near-field data of demodulation order 4 and *Comsol* simulations for a variation of the coupler radius (a) and waveguide width (b).

5.7 Conclusion

In this chapter, we have presented our work on plasmonic chiral couplers that utilize the photonic spin-Hall effect. We discussed the analytically expected resonance condition, the design, simulation and fabrication of these couplers. An unexpected beating pattern observable on the waveguides was explained via the interference of the waveguide mode with a dipolar field distribution caused by the direct illumination with the laser light. The directionalities extracted from the measurements show a clear dependence of the directionality on the coupler radius and waveguide width, in reasonable agreement with simulations. However, no clear resonance condition could be observed, likely due to the influence of the connected waveguides and coupling efficiency variations. Overall, this work demonstrates the potential of chiral couplers for controlling light propagation at the nanoscale using SOI.

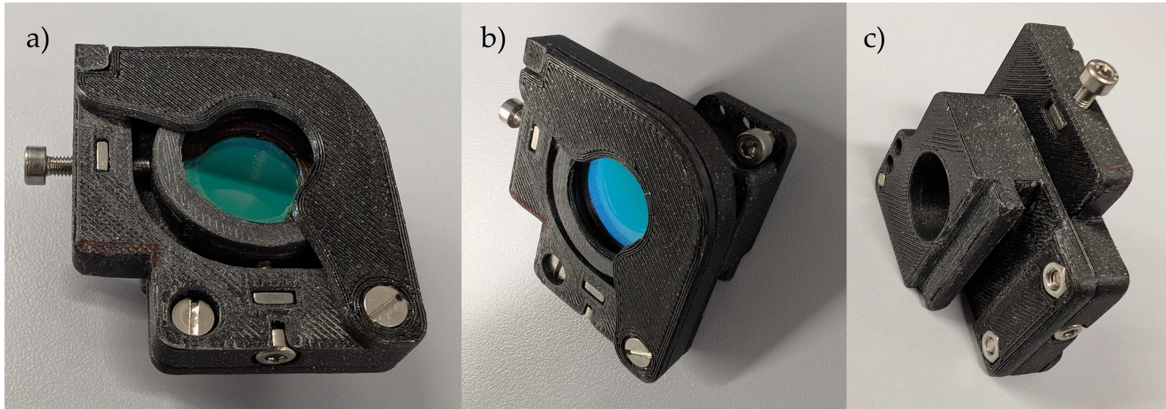


Figure 5.12: 3D-printed adapter to mount the vortex retarder plate in the beam path of the transmission imaging module in front of the lower parabolic mirror. Shown are three different angles of the adapter (a) front view, (b) front/side view of the mounting and (c) backside.

5.8 Outlook

An interesting extension of the chiral coupler concept would be the use of orbital angular momentum (OAM) states of light for excitation. OAM states, also known as optical vortices, possess a helical (corkscrew) phase front characterized by an azimuthal phase dependence of $\exp(il\phi)$, where l is the topological charge and ϕ is the azimuthal angle around the beam axis [102]. These states carry IOAM of $l\hbar$ per photon along the propagation axis. The coupling of OAM states to waveguide modes could also lead to directional excitation effects, similar to the spin-Hall effect discussed earlier. So far, no obervations of such effects have been reported in literature, making this a promising area for future research.

To explore this concept, we have performed preliminary experiments using a vortex phase plate (VPP) to generate a donut-shaped beam with OAM. The VPP imparts a helical phase structure onto the incident Gaussian beam, resulting in a ring-shaped intensity profile with a phase singularity at the center. We use a commercial VPP (*Thorlabs*, WPV10L-1550) designed for a wavelength of 1 550 nm with a topological charge of $l = 1$. The VPP is placed in the beam path after a linear polarizer, allowing us to control the polarization state of the generated donut beam. By rotating the VPP relative to the linear polarizer, we can generate different polarization states around the ring, such as radial or azimuthal polarization. Due to space constraints in the setup, the VPP is placed in a custom 3D-printed mount screwed into a slot in the transmission interferometer. The 3D-printed mount, shown in Fig. 5.12, allows for a positional offset of the plate via two set screws, enabling a rough alignment of the VPP within the beam. If the VPP is correctly aligned to the center of the beam, a clean donut beam is generated. However, if the plate is misaligned, the resulting beam profile can be distorted, leading to asymmetries in the intensity distribution and phase structure. Careful alignment is therefore crucial for generating high-quality OAM beams.

We used this donut beam to excite plasmonic disc arrays and ring structures, similar to the chiral couplers discussed earlier. The measurements were performed in transmission mode using the s-SNOM setup. The results of these measurements are shown in Fig. 5.13 and Fig. 5.14. In Fig. 5.13, the field distribution of a disc array excited with the donut beam is shown. The ring shape of the excitation beams amplitude is clearly visible in the amplitude pattern, along with the phase singularity at the center. In Fig. 5.14, a ring of discs is excited with the donut beam. The individual discs show the positional polarization state of the donut beam. The measured polarization states vary from expectations. A sketch of the measured and expected polarization state around the ring is shown in Fig. 5.14 (lower panel). The polarization state can be controlled by the rotation of the vortex plate relative to the linear polarizer in front of it. The actual local polarization state can be estimated by looking at the amplitude and phase distribution of the individual gold discs arranged in the ring pattern. However, the observed polarization states deviate from the expected polarization state. The deviation could be due to imperfections in the vortex plate, misalignment in the setup or distortions introduced by the focusing optics. Further investigations are needed to clarify these effects and optimize the generation of OAM states for coupling experiments. So far no attempts to excite chiral couplers with OAM states have been successful, due to time constraints and difficulties in generating a clean donut beam with the current setup. Future work could focus on improving the OAM beam generation, for example by using a spatial light modulator for more flexible beam shaping, and performing systematic studies of the coupling efficiency and directionality of chiral couplers excited with OAM states.

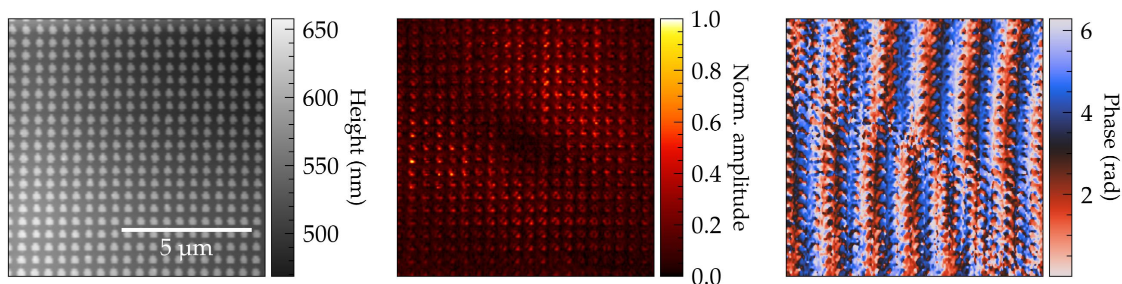
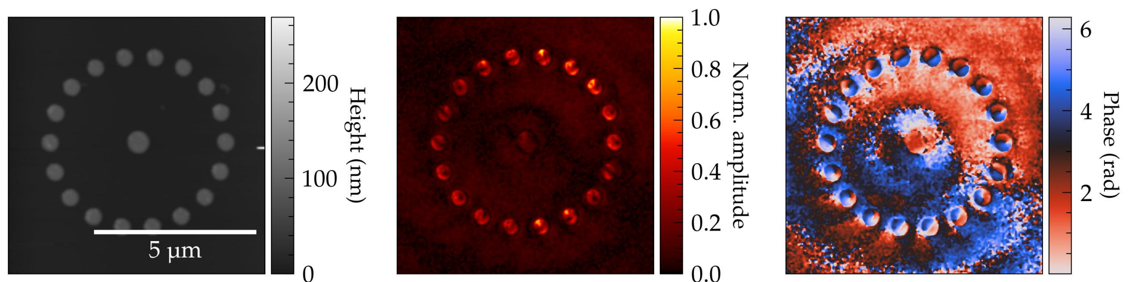


Figure 5.13: Measurement of a disc array (left) excited with a donut beam in transmission mode with the beam fixed to the middle of the image. The donut shape of the excitation beam is visible in the amplitude data (middle) of the disc array and the phase data (right) reveals the expected singularity in the center.



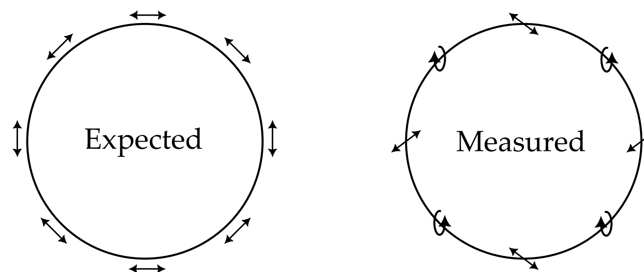


Figure 5.14: Measurement of a ring of discs excited with a donut beam in transmission mode with the beam fixed to the middle of the image. The individual discs show the positional polarization state of the donut beam. Shown is the fourth demodulation order and all images have been Gauss filtered for better visibility. The sketch below shows the expected (left) and measured (right) polarization states around the ring of discs.

CHAPTER 6

Single DLSPP Waveguides

Plasmonic waveguides confine light to subwavelength dimensions by coupling it to collective electron oscillations at a metal-dielectric interface, known as surface plasmon polaritons surface plasmon polaritons (SPPs). This confinement leads to much smaller mode volumes than for dielectric waveguides, enabling strong light-matter interactions which are beneficial for integrated photonic circuits with a small footprint and efficient coupling to quantum emitters. However, this comes at the cost of higher propagation losses due to ohmic heating in the metal. Dielectric-loaded surface plasmon polariton waveguides (DLSPPWs) offer a good compromise between confinement and losses, compared to other plasmonic waveguides like slot waveguides [103, 104] or strip metal waveguides [101]. This makes them attractive for various applications in nanophotonics, including photonic computing and the modeling of condensed-matter systems [105].

A DLSPPW can be realized by placing a dielectric ridge on top of a metal film [106]. As described in Section 2.2 and demonstrated in [38], the mode of a DLSPPW can be excited via grating coupling. Several optical elements such as reflectors, splitters and interferometers have been demonstrated based on DLSPPWs [107, 108].

DLSPPWs can be imaged with leakage radiation microscopy (LRM) due to leakage of the waveguide mode into the substrate below the metal film, if the metal film is thin enough. This is a powerful technique which allows to take large scale images in real and Fourier space. Furthermore, they are also interesting for near-field microscopy due to the accessibility of the electric fields. Near-field microscopy allows to complement LRM with the ability to measure the local phases of the electric field, further boosting the usefulness of DLSPPWs. These waveguides have been studied before [37], even in near-field studies [38, 39], but a central property has not been discussed. SPPs are transverse magnetic (TM) waves, meaning that they have not only an out-of-plane electric field component but also an in-plane component. Usually, the in-plane component can not be detected as well as the out-of-plane component due to the tip-sample coupling described in Section 4.2.4. Therefore, the in-plane component is often neglected in the scanning near-field optical microscopy (SNOM) community. For certain structures, such as DLSPPWs, the in-plane component can be quite significant.

In the following, measurements of single DLSPPWs with different geometries and illumination wavelengths are presented and complemented with *Comsol* mode simulations. In gerneral, the scattering-type scanning near-field optical microscope (s-SNOM) is mostly sensitive to the out-of-plane component of the electric field. In order to measure the in-plane component separately a reflection grating is implemented. The gratings are simulated and measured in a suitable configuration to separate the two components. Since the fabrication of DLSPPWs can be quite challenging, the fabrication process is also explained in detail.

6.1 Single Waveguides

We will start with a mode analysis simulation to investigate the mode profiles. The focus here will be on the ratio of the in- and out-of-plane components of the electric field. The exact ratio of the in- and out-of-plane components depends on the waveguide geometry and illumination wavelength. The simulations are done for two different wavelengths. Historically, these waveguides have been studied in our group at a wavelength of 980 nm, because it is a good compromise between losses and camera sensitivity in the LRM setup. This wavelength is above the limit of our laser system, so all measurements and simulations are done for 970 nm instead. Using a similar wavelength allows us to compare near-field measurements to previously performed LRM measurements. For near-field measurements it is often easier to use a longer wavelength, as this usually leads to better signal-to-noise ratio (SNR) and less background. Additionally, a longer wavelength changes the mode properties of the waveguides, which will be discussed in the following sections. Therefore, we also perform measurements and simulations at a wavelength of 1 500 nm, which is also a common wavelength close to the telecommunication band.

6.1.1 Simulations

The mode profiles of DLSPPWs are simulated by the finite element method (FEM) with *COMSOL Multiphysics*. Such a mode analysis simulation calculates the possible eigenmodes for a given simulation geometry and excitation frequency. For the materials poly-methyl-methacrylate (PMMA) and gold we use the built in material library, which uses the data from Johnson and Christy [43] for gold and Sultanova [87] for PMMA. For the glass substrate we assume a constant refractive index of 1.5.

The mode simulations yield several properties of the waveguide modes. The first is the complex-valued effective mode index n_{eff}. This gives us information about the phase velocity and effective wavelength of the mode and also about the propagation length L_{p}, which can be calculated from the imaginary part of the effective mode index as $L_{\mathrm{prop}} = \frac{\lambda}{4\pi \mathrm{Im}(n_{\mathrm{eff}})}$. The propagation length is defined as the distance over which the intensity of the mode drops to $1/e$ of its original value. The real part of the effective mode index and propagation length are closely related to the mode confinement. Finally, the relative strength of the in- and out-of-plane components of the electric field is investigated.

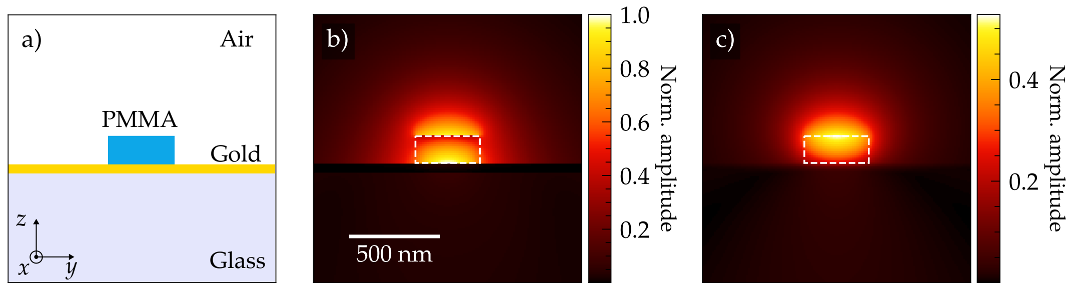


Figure 6.1: Simulation of a DLSPPW with a width of 350 nm and a height of 150 nm at a wavelength of 970 nm. In (a) the simulation geometry is shown with the materials glass (substrate), PMMA (waveguide), air (above metal film and waveguide) and gold (metal film). The colormap images display the out-of-plane (z) (b) and in-plane (x) (c) electric field amplitude normalized to the maximum field amplitude. The waveguide is highlighted by the white dashed box in the amplitude images.

An example of such a mode simulation is shown in Fig. 6.1. The three images display a zoomed in view of the materials (a) out-of-plane (b) and in-plane (c) electric field amplitude. The phase is not displayed, it is constant over a cut-plane accross a waveguide. The phase of the out-of-plane and in-plane field is shifted by $\pi/2$. This will be important later on as it can lead to interference if both components are detected at the same time. According to the simulation, the out-of-plane component is stronger overall but less localized to the waveguide, especially above the metal film next to the waveguide. The in-plane component is more localized to the top of the waveguide and has little to no contribution to the sides. Therefore, in the following measurements we expect to see that the signal collected from the sides of the waveguide will be primarily caused by the out-of-plane component, while on top of the waveguide both components will contribute. A variation of the waveguide geometry changes the mode profiles and allows to tailor the waveguide properties to specific applications. More details about the geometry variation can be found in the appendix A.2.1.

In Fig. 6.2 the effective mode index and propagation length are shown for a variation of the waveguide width and height at two different wavelengths. The effective mode index increases with increasing waveguide dimensions while the propagation length decreases in the investigated parameter range. Both effects are related to the confinement of the mode. An estimation of the confinement of the modes for the different geometries can be found in the appendix A.2.1, but in general the confinement increases with the dimensions of the waveguide and is generally higher for the shorter wavelength of 970 nm. Naively, one might expect that a larger waveguide leads to a less confined mode, but here the opposite is the case. The reasons for this behavior is the deep subwavelength mode volume of these waveguides. We operate these waveguides already in a weakly guiding regime, if the geometry of the waveguide is too small, the difference in dielectric constant between the waveguide and the surrounding medium is not sufficient to confine the mode well. A similar behavior can be

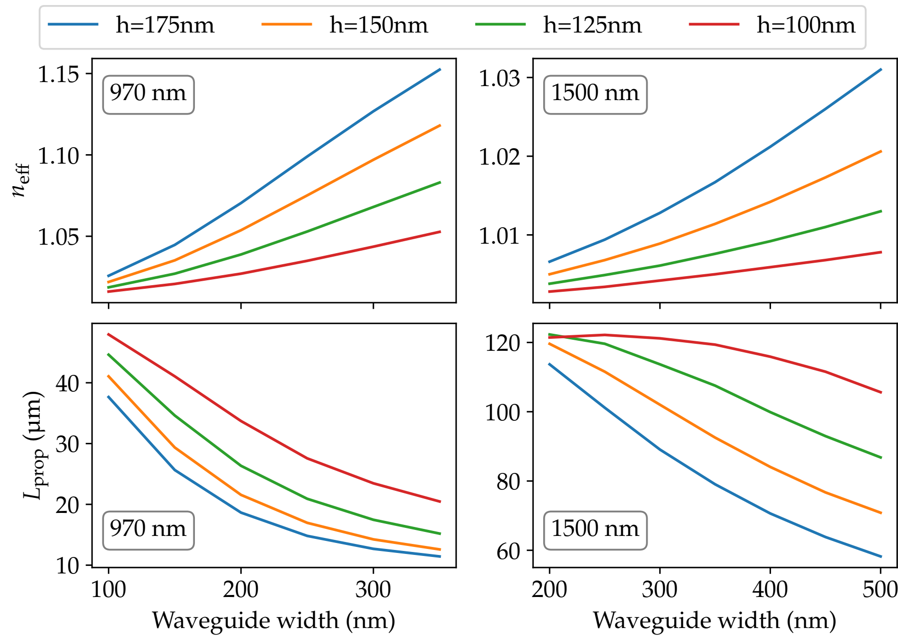


Figure 6.2: Displayed are the effective mode index (upper row) and propagation length (lower row) for a variation of the waveguide width and height at two different wavelengths. The effective mode index increases with increasing waveguide dimensions while the propagation length decreases in the investigated parameter range.

observed in dielectric waveguides when pulling them to subwavelength dimensions [109]. If the waveguide geometry increases, so does the confinement, leading to a larger portion of the mode being concentrated to the high index dielectric waveguide, resulting in a larger effective mode index. On the other hand, a higher confinement also leads to stronger losses, especially in the metal film due to ohmic losses, thus a shorter propagation length. If longer propagation lengths are required a metal with lower losses like silver could be used instead of gold.

The choice of a high confinement or propagation length is always a compromise. A waveguide with a high confinement is desirable, because it increases the local field strength and allows to place structures like other waveguides closer. On the other hand, the propagation length needs to be sufficient for our applications. For experiments we usually chose a waveguide width of 300 to 350 nm and a height close to 150 nm as this offers a good compromise between confinement and losses for a wide range of wavelengths. For the longer wavelength the waveguide height would need to be increased to maintain a similar confinement. However, we limit the waveguide height to 150 nm, which allows us also to fabricate multiple waveguides for various excitation wavelengths on one sample with the positive tone lithography process discussed in Section 6.1.2. To compensate for the limited height we can increase the waveguide width to a certain extend.

The interplay of the in- and out-of-plane components is important to investigate. Especially for the near-field measurements a good understanding of the ratio of the in- to out-of-plane components is necessary. For this, we use the same mode profile simulations and export the amplitudes of the electric field for the two components, as shown in Fig. 6.1 (b, c). In order to mimic the measurement situation, where the probing tip is oscillating at a small distance above the sample surface, the ratio is calculated by averaging the absolute values of the electric field components over a small area above the waveguide. If we calculate the ratio for multiple waveguide width variations for different waveguide heights we get Fig. 6.3. One can see that the out-of-plane component is always dominant and often by a factor of two to four times stronger. The ratio varies with the waveguide dimensions and gets smaller for wider waveguides. Higher waveguides lead to a smaller ratio, but for the highest simulated waveguide the ratio becomes larger again. This indicates an optimal height around 150 nm to measure the in-plane component. For the larger wavelength the behavior is similar, but the ratio is generally larger, indicating a weaker in-plane component. Due to the localization of the in-plane component to the top of the waveguide, the ratio is even larger if it would be calculated above the metal film besides the waveguide. Therefore, it will be very difficult to measure the in-plane component there.

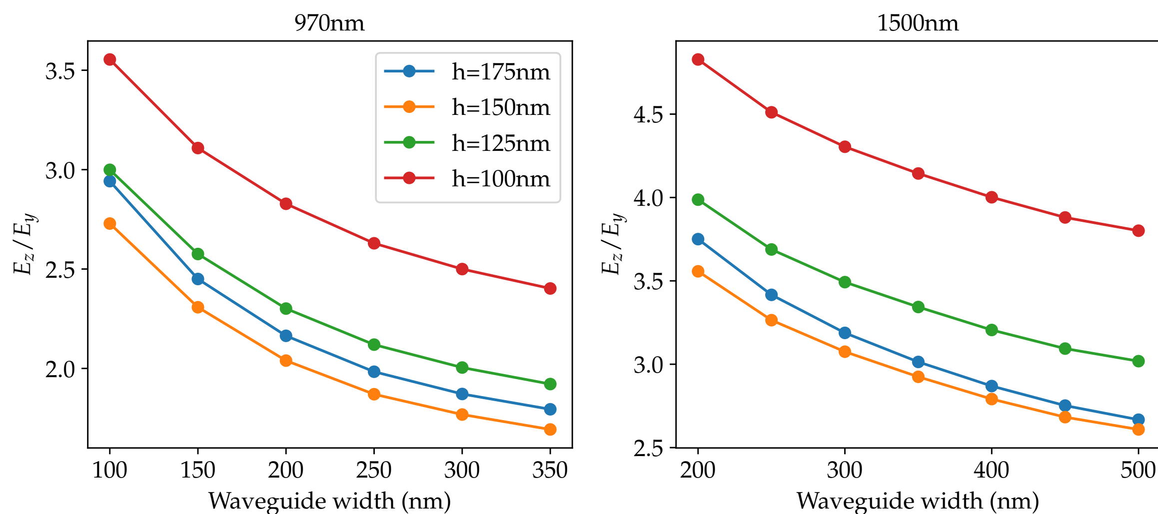


Figure 6.3: Ratio of the absolute value of the out-of-plane (z) component to the in-plane (y)-component of the electric field for different waveguide widths and heights for two wavelengths. The ratio is averaged over a small area above the waveguide.

6.1.2 Fabrication

In previous works, DLSPPWs were fabricated using negative tone electron beam lithography (EBL) in our group. Naturally, this was the starting point for this project. However, as the project progressed, problems occured, which are linked to the use of PMMA as a negative tone resist. PMMA is in principle a positive tone resist, but for a high enough dose the

resist crosslinks and can be used as a negative tone resist as well. The problem with this process is, that the crosslinking is dependent on the exact composition of the resist, which may vary depending on the batch and manufacturer. This can lead to unwanted sample defects, which are discussed in the appendix A.2.4. In order to circumvent these problems, the fabrication method was changed to positive tone EBL with PMMA. This process is more robust and leads to better results overall. The height of developed structures using this process is fixed by the resist thickness after spin coating and can not be controlled by the dose like in negative tone lithography. This is an advantage here, because we do not want height variations in our samples.

Negative tone lithography

Negative tone lithography with PMMA is a grayscale lithography technique. The general process is described in Section 3.1.4. The resist thickness after spin coating defines the maximum height of the waveguides, but due to the proximity effect and the crosslinking of the resist, the actual height can be much lower. This can be used to create waveguides with different heights on the same sample, which can be useful for certain applications. The waveguides are written as a rectangle consisting of individual lines to have a better control over the final shape.

An example dose test is displayed in Fig. 6.4. This height profile is measured with the atomic force microscopy (AFM) and shows that the height of the waveguides increases with increasing dose, as expected. The width of the waveguides also increases with increasing dose. Two example dosetests with different dose variations are shown in Fig. 6.5. The height and width of the waveguides are plotted against the dose factor. Although both dosetests are fabricated in the same fabrication run, the height and width data are slightly off. The width data for the large variation are smoothened due to a low resolution of the AFM scan. The height increases almost linearly after passing the crosslinking threshold, and then starts to saturate. The width increases more or less linearly in the investigated range. In order to control the final width and height of the waveguides, both the design width and the dose must be adjusted. The dose tests shown here are meant as an example and should be repeated frequently.

Typical parameters for a waveguide with a height of 140 nm and a width of 250 nm are a design width of 100 nm width a line distance of 5 nm and a dose factor of 3. One could also opt for a larger design width and a smaller dose factor like shown in the dose test example, but in practice little difference was observed and the proximity effect is very strong anyway. However, as mentioned before, this process has some serious limitations. The biggest problems are the reproducibility of the results and the shape and surface quality of the waveguides, as described in the appendix A.2.4.

In conclusion, negative tone lithography with PMMA did not lead to consitent results for the fabrication of DLSPPWs. The fabrication problems seem not to have occured in previous works. But here, they impose serious limitations and can not be ignored. Especially thin and narrow waveguides can not be fabricated.

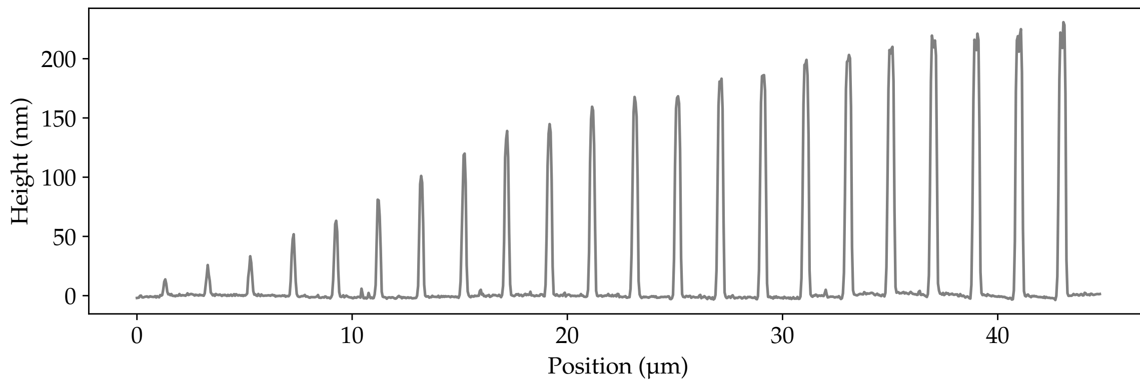


Figure 6.4: Height profile measured with the AFM of a dose test for PMMA negative tone lithography. The waveguides are written with a design width of 350 nm and a line distance of 5 nm. The waveguides are spaced by 2 µm to avoid proximity effects and the dose factor is increased from left to right in steps of 0.025 up to 1.

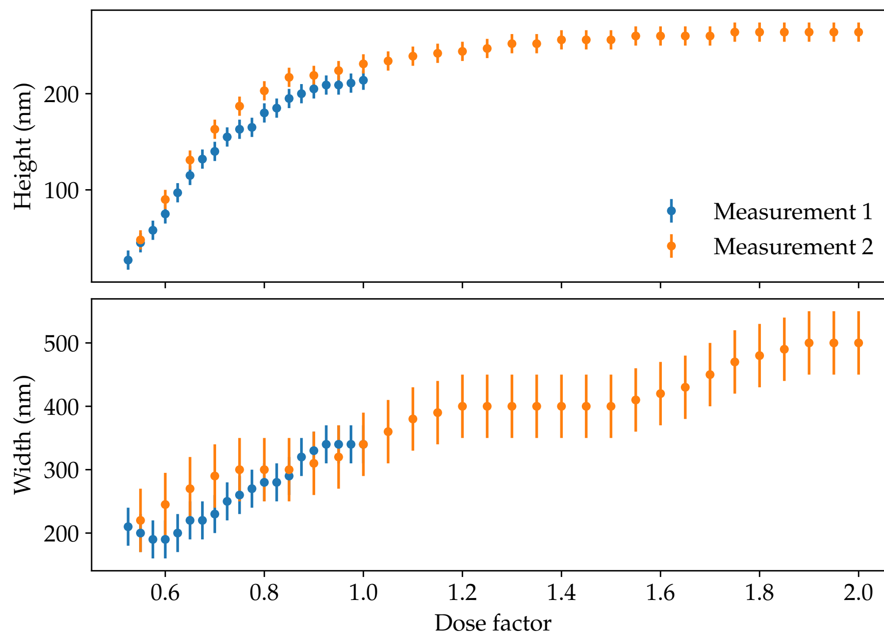


Figure 6.5: Evaluation example of a dose test for PMMA negative tone lithography. The waveguides are written with a design width of 350 nm and a line distance of 5 nm. The dose factor affects the final width and height of the waveguides. The width data for the large variation had to be smoothened due to a low resolution of the AFM scan.

Positive tone lithography

Positive tone lithography with PMMA was implemented as an alternative to negative tone lithography. As explained in Section 3.1.3, this method is usually used to create masks for the evaporation of various materials. This could be done here as well, for example a dielectric material such as silicon dioxide could be used, but in our case we use it to create the waveguides themselves. This makes the design phase more involved as one has to expose around the structures which should remain after the developement process. The benefits to this approach will be discussed later. The resist thickness after spin coating defines the height of the waveguides, as the exposed areas are removed during the developement process. A table with some example thicknesses achievable with different spin speeds for the 2 % PMMA (950k) solved in anisol is shown in Table 6.1. For very thin waveguides however, the

Final thickness (nm)	Recipe
98 nm	PMMA (950k, 2 %), 1500 rpm, 90 s
116 nm	2× PMMA (950k, 2 %), 4000 rpm, 90 s
135 nm	2× PMMA (950k, 2 %), 3250 rpm, 90 s
160 nm	2× PMMA (950k, 2 %), 2500 rpm, 90 s

Table 6.1: Examples of resist thicknesses achievable with different spin speeds for the 2 % PMMA (950k) solved in anisol.

height decreases a bit due to the proximity effect. This can be limited and without much optimization it is possible to create waveguides with a height of 140 nm at a width of less than 200 nm, which is not possible with PMMA negative tone. The parameters for this are a design width of 250 nm with a line distance of 20 nm at a dose factor of 0.3.

An example design of a waveguide with a tapered incoupling area is shown in Fig. 6.6. The tapered incoupling area is written on top of the incoupling grating and the long gold area shown before in Fig. 3.10. Contrary to the negative tone process, the dose is not the main parameter to control the width of the waveguides, but the design width. The dose is chosen just high enough to ensure a complete exposure of the written areas. This minimizes the proximity effect and leads to sharper edges. This process is also much more controlled than the crosslinking in the negative tone process making it easier to replicate the same structures. AFM scans reveal a much smoother surface and sharper edges as shown in Fig. 6.8 for examle.

The tapered incoupler is a modification to allow for a bigger excitation area, which in turn should lead to a better coupling efficiency. Such a design is also frequently used in the literature [38, 101, 110, 111]. To realize such a shape with negative tone lithography while keeping the height constant would be challenging. Exposing areas for positive tone lithography is much faster, due to the lower dose threshold for exposure. However, larger areas around the waveguides need to be exposed, leading to a longer patterning time. In order to keep the writing times fast enough, the area around the waveguide is limited to

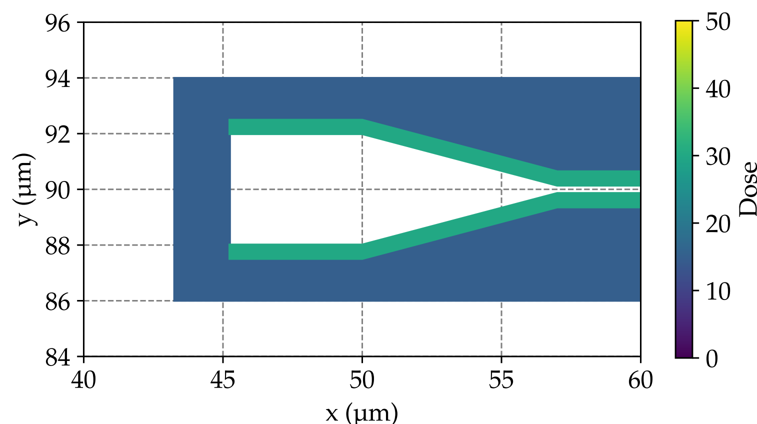


Figure 6.6: Design of the second step of the two step EBL process for the fabrication of DLSPPWs. In This step the waveguides are written on top of the incoupling gratings and the long gold areas (not shown here). The edges closest to the waveguides are written first and as lines to have a better control over the final shape. The areas surrounding the waveguides are written as one connected area.

a total width of 8 µm for wavelengths below 1 000 nm and 10 µm for wavelengths above. All structrues, which are to be measured with the s-SNOM need to be exposed as well in order to remove the PMMA layer on top. A simple example is a bare metal reference area, which is required to adjust the upper parabolic mirror, or small resonant antennas, which are useful to help aligning the interferometer. Since we can only use a mild solvent (methyl isobutyl ketone (MIBK)) to remove the exposed resist, PMMA residues in the exposed areas can remain on the substrate. Over time or during AFM or s-SNOM scans this can lead to adhesion problems and the degradation of samples. The residues can be removed by a short plasma cleaning step. Furthermore, the exposure of PMMA can lead to an undercut during the developement process, which can be problematic for very thin waveguides. This effect seemed not to be a problem in our case, but it should be kept in mind. Additionally, one key advantage of this process is that the samples can be rewritten or additional structures can be added without the need of a new coating. This makes the process very flexible and allows to adapt the samples to new requirements. This was a major drawback of the negative tone process, as new samples had to be fabricated each time, including the creation of the incoupling gratings in the first step. This is because the crosslinked resist becomes suprisingly resilient to solvents and can not be removed easily.

Final Fabrication Process

The final fabrication process for the DLSPPWs with positive tone lithography is as follows:

1. Spin coating of PMMA (950k A2, *Allresist*) at 3200 rpm for 90 s on a indium tin oxide (ITO) covered glass slide (1 cm^2). Heat at 170 °C for 120 s. Repeat the spin coating and baking step to achieve a resist thickness of approximately 140 nm. But the resist for the

second layer should only be added when the rotation speed ramps up, to avoid the dissolution of the first layer. The resist is taken out of the fridge shortly before use to avoid heating and degradation.

2. EBL of the waveguides with 20 kV and a 10 μm aperture. The writefield needs to be aligned to the first step. The casing of the waveguide has a width of 500 nm, is written as lines with a line distance of 20 nm and a dose factor of 0.3. The surrounding area is written as a connected polygon area with a dose factor of 0.155. For a waveguide with a width of roughly 350 nm a design width of 400 nm is used.

3. Development of the resist in MIBK for 55 s followed by a rinse in isopropanol and drying with nitrogen.

6.1.3 Measurement

All measurements shown in this chapter have been performed with the s-SNOM setup described in Section 4.2.1. The waveguides are excited through the glass substrate with a focused laser beam at normal incidence. The polarization of the excitation beam is always rotated to be parallel to the waveguide axis to maximize the coupling efficiency. Standard tapping mode tips (Arrow NcPt, *NanoWorld*) are used, which have an electrically conductive platinum-iridium coating of 23 nm thickness, a typcial tip radius of 25 nm, a tapping frequency around 285 kHz and are operated at a tapping amplitude of around 65 nm. All measurements are performed in ambient conditions interfereometrically in the syncronized mode, meaning the illuminating lower parabolic mirror moves during the measurement to keep the excitation spot fixed on the incoupler. All measurements are corrected for the phase gradient introduced by the movement of the illuminating parabolic mirror as described in Section 4.4.1.

An example measurement of a single DLSPPW is shown in Fig. 6.7. Shown is the near-field amplitude and phase together with the height map. The height shows the location of the grating coupler milled with the focused ion beam (FIB) into the gold film, as well as the horizontal waveguide fabricated with negative tone lithography. The amplitude and phase images reveal multiple features. First of all, the grating coupler excites a mode in the DLSPPW propagating to the right. This can be seen from the high amplitude confined to the waveguide and the linear phase gradient on the waveguide. The coupler also excites free SPPs which can propagate in any direction, but are excited favorably horizontal as well due to the orientation of the grating and polarization of the illumination. This can be seen from the circular wavefronts in the phase image and the cone shaped amplitude distribution, to the left and right of the coupler.

The free SPPs can overlap and interfere with the waveguide mode. They have a lower effective mode index and thus a longer wavelength. The waveguide mode and free SPPs are initially excited in phase at the grating coupler but due to the different effective wavelengths they will shift out of phase after a certain propagation distance. This can be seen at the right end of the measurement (Fig. 6.7), especially in the last 5 μm. Here, the SPPs phase is

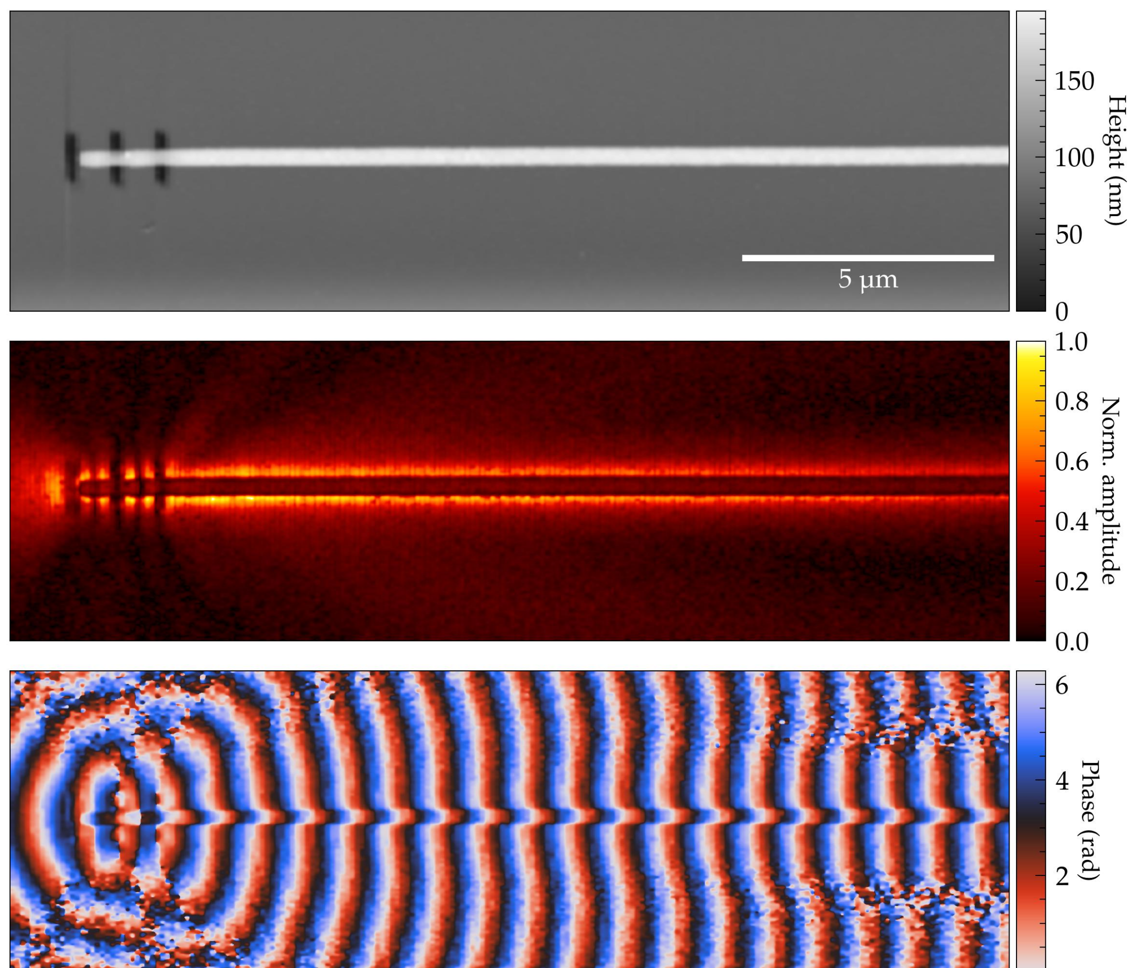


Figure 6.7: Transmission mode s-SNOM measurement of a single DLSPPW at a wavelength of 970 nm. Shown is the third demodulation order (O3). The waveguide has a width of 300 nm and a height of 115 nm. This sample is fabricated with negative tone lithography. The height data is leveled and all channels are filtered with a gaussian filter for a cleaner image.

shifted by π relative to the waveguide mode leading to destructive interference. This can be seen in the amplitude image as a significant drop of the amplitude and by a phase jump of π in the phase image. A further indication is the appearance of noise in the phase data, as the amplitude is very low here. The interference occures parallel to the waveguide in a certain distance from the waveguide. This distance shows that the field surrounding the waveguide until this distance is dominated by the waveguide mode and not free SPPs. The fields further away from the waveguide are dominated by the free SPPs. But this distance can also change depending on the losses of the waveguide mode and free SPPs, and the confinement of the waveguide mode. The knowledge, that the field close to the waveguide corresponds purely to the waveguide mode will be important later on, as it allows to not only use the field amlitude and phase on top of the waveguide but also on the metal film next to it.

The measurement also shows two unexpected features. The mode simulations, which do not include the measurement tip, suggest, that the electric field is strongest on top of the waveguide. In stark contrast, the waveguide is much darker than the surrounding metal film in the measurement. This can be explained by the polarizability of the materials involved. As discussed in Section 4.2.4 the signal scattered by the tip depends largely on the polarizability of the material it is probing. The polarizability of PMMA is much lower compared to gold, thus probing the near-field on PMMA leads to a lower signal. Therefore, the field on top of the waveguide appears weaker compared to the field on the adjacent gold. This effect increases with the height of the PMMA layer on top of the gold film [1]. The respective polarizabilities can be estimated analytically using a tip sample coupling model such as the point dipole model (PDM) described in Section 4.2.4.

The second unexpected feature is a phase shift between the waveguide mode on the PMMA and adjacent gold. This phase shift does not show up in simulations and is purely a measurement artifact. It can be quite significant but it's origin is much harder to determine. A similar shift can be seen in many published articles on s-SNOM measurements on similar waveguides [39, 101, 112, 113], but it is usually not mentioned or discussed. We will discuss this phase shift in more detail in Section 6.1.4.

As a second example at a larger wavelength, a different measurement at 1 600 nm wavelength is shown in Fig. 6.8. The overall behavior is very similar to the previous measurement. The tapered incoupler is an improvement over the thin FIB milled coupler and improves the coupling efficiency. Due to the positive tone fabrication method the incoupler and waveguide are very well defined with very little height variation and smooth surfaces.

From such measurements, we can extract the effective mode index [2] for a variaty of waveguide widths, like shown in Fig. 6.9. The values for the wavelength of 970 nm and 1 600 nm are in reasonable agreement with the values from the mode simulations. The values for 970 nm with a waveguide width above 500 nm start to show a multimode behavior. One

[1] Due to the finite thickness of the PMMA layer, the tip sample coupling is influenced by the metal film below the PMMA.

[2] The mode index can be calculated from the slope of the phase gradient of the propagating mode.

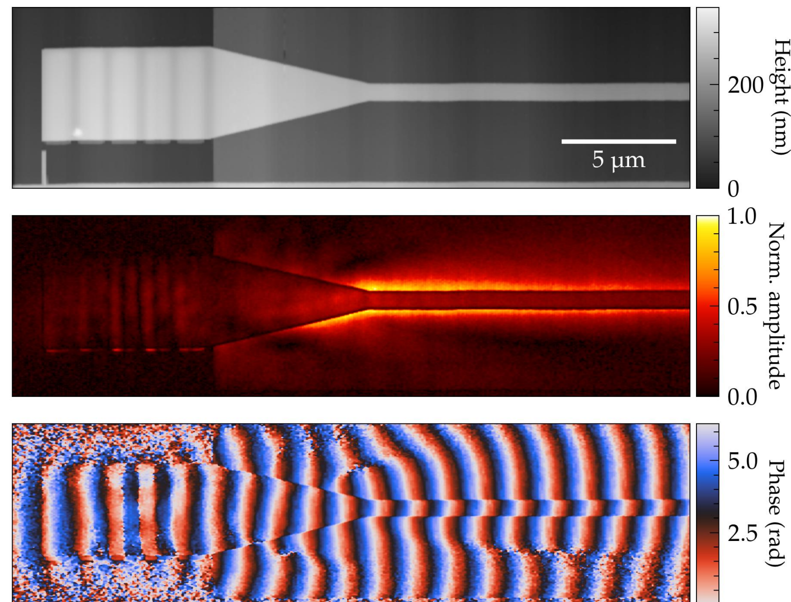


Figure 6.8: s-SNOM measurement of a single DLSPPW with a tapered incoupler at 1 600 nm. Shown is the third demodulation order (O3). The waveguide has a width of 720 nm and a height of 155 nm. This sample is fabricated with positive tone lithography. All channels are filtered with a gaussian filter for a cleaner image.

should note that these measurement are not optmimal for the extraction of the effective mode index as we intentionally scan the probing tip perpendicular to the waveguide.

The reasons for this are manifold. First of all, for AFM or s-SNOM imaging it is advisable to scan the fast axis perpendicular to edges to reduce imaging artifacts, as scanning the tip parallel to rough edges can lead to slipping and bending. Secondly, during a scan the amplitude and phase can slowly vary along the slow scan axis due to instabilities in the AFM feedback, probing tip degradation, laser system or thermal drifts. Especially for the measurements of coupled waveguides, scanning perpendicular is advantageous, as here the amplitudes and phases of all waveguides are more comparable, because the slow variations occur along the waveguides and not across them. Furthermore, slow variations along the waveguides are much easier to correct than those across the waveguides. In order for all measurements to be comparable we therefore always scan the probing tip perpendicular to the waveguides. Since this scanning orientation can lead to larger variations along the waveguide axis, the extration of the effective mode index and propagation length is less accurate. But these parameters were not the main focus of these measurements. The effective mode index and propagation length are better measured with LRM as it has been done in our group previously [114, 115].

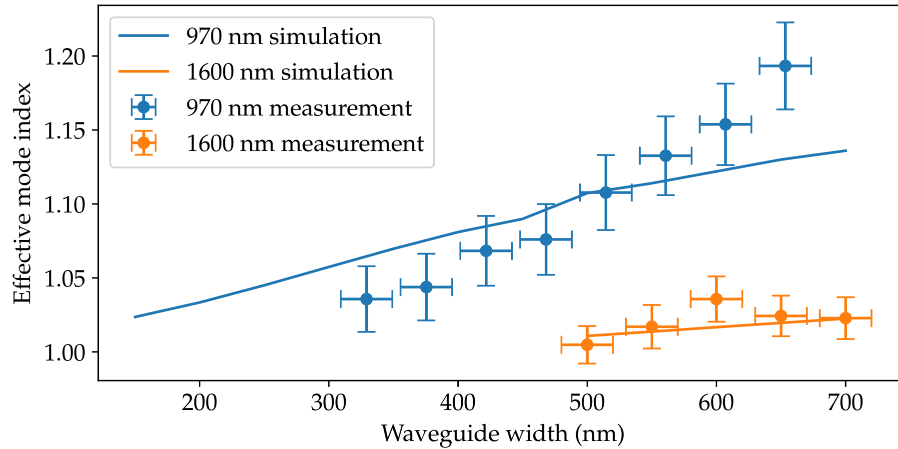


Figure 6.9: Effective mode index for a variation of the waveguide width.

6.1.4 Discussion of the Phase Shift

The origin of the measured phase shift between the signal on and next to the waveguide is complicated and can be attributed to a combination of several factors:

1. When the tip is scanned over the waveguide the sample table moves down (closer to the light source). This leads to a phase shift because the coupler is excited earlier. This effect can be corrected by substracting a phase shift based on the relative height per pixel and illumination wavelength. But because of the next point we do not apply this correction.

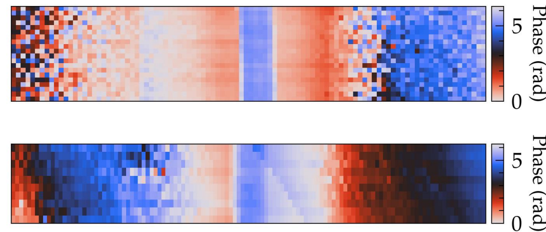


Figure 6.10: Phase map of two DLSPPWs (O2). The first measurement shows a symmetric phase shift for both sides of the waveguide, while the second measurement shows an asymmetric phase shift. The data is shown without any modification.

2. The tip-sample interaction leads to the scattering of light, which is partly directed towards the detector and partly first reflected from the sample surface. The reflected beams path depends on the sample material and topography but the direct path does not. The position of the tip (on top of the waveguide or next to it) leads to a variation of the reflection beam path. This effect and the first might cancel each other but the

exact influence is difficult to determine, so we will not try to correct either of them. Furthermore, it is not clear which contribution of the tip scattered light is dominant. An example of this effect can be seen in Fig. 6.10. Here, two measurements of similar waveguides are shown, but due to different experimental conditions like tip shape and interferometer settings, the phase shift can be very asymmetric. The asymmetry can be attributed to differences in the light emission and reflection from the sample plus path length changes and connected interference effects.

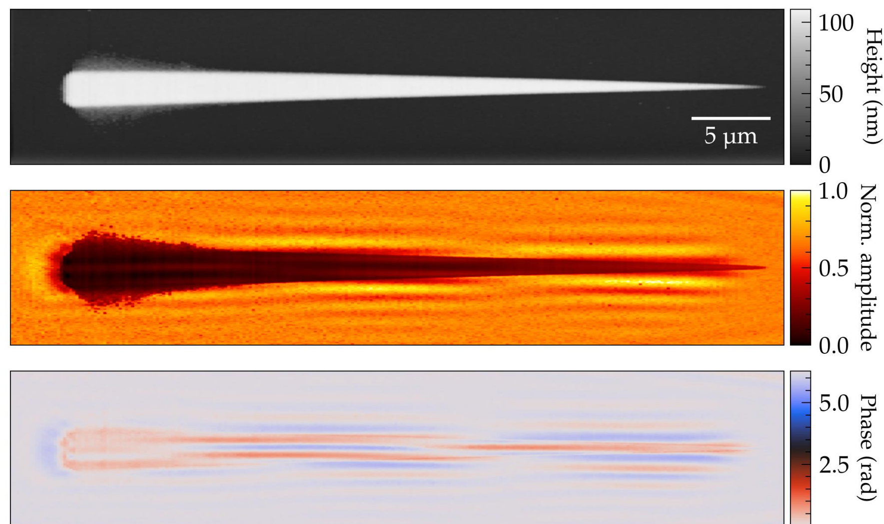


Figure 6.11: Reflection mode s-SNOM measurement of a PMMA wedge on a gold film. Shown is the third demodulation order (O3). The polarization of the incident light is parallel to the tip axis and the edges of the wedge to avoid an excitation of SPPs from the PMMA edges. The standing wave pattern originates from SPPs excited by the tip itself. The data has been pre processed to remove an amplitude and phase drift over time, the height data is leveled and all channels are filtered with a gaussian filter for a cleaner image.

3. The tip-sample coupling can also lead to a phase shift in the emitted light. The tip is not simply emitting light relative to the electric field it is exposed to. The tip is coupled to the sample and can therefore also introduce additional effects in the sample. A simple example are samples with a resonant behavior [116], but other effects can also play a role. To illustrate this a PMMA wedge on a gold film is measured in the reflection mode. The polarization of the incident light is rotated to be parallel to the tip axis and the sample is rotated such that the PMMA edges are also parallel to the light polarization. This suppresses the excitation of SPPs at the PMMA edges. The measurement displayed in Fig. 6.11 clearly shows a standing wave pattern which originates from SPPs that are not directly excited by the incident light but by the tip itself. These SPPs propagate radially from the tip and can be reflected from the PMMA edges. The interference of the incident and reflected SPPs leads to the standing wave

pattern. The standing wave pattern outside of the PMMA wedge shows a periodic modulation of the field intensity along the edge of the wedge, which is caused by the interference of the SPPs reflected from the two sides of the wedge. Inside the PMMA wedge a similar standing wave pattern can be seen, but with a different periodicity due to the changed SPP wavelength inside the PMMA. The most interesting part is however the phase shift in the middle of the wedge relative to the gold film. This phase shift varies with the width of the PMMA wedge. Similar processes can also occure during a transmission mode measurement where the tip is not directly excited by the incident light but the fields of the sample, e.g. a waveguide mode.

4. As mentioned before, the tip is not only sensitive to the out-of-plane component of the electric field but also to the in-plane component. If the projection of both components have an overlap at the detector, they can interfere. Since the components have a phase shift of $\pi/2$ according to the mode simulations, this can lead to a phase shift in the detected signal. The interference of both components can become relevant especially on top of the waveguide where the in-plane component is strongest. The resulting phase shift mainly depends on the ratio of the two components, and the tip's polarization sensitivity.

5. The demodulation order can also be important. Higher demodulation orders are more sensitive to fast changes in the near-field, as we have seen in Section 4.2.3 and thus to smaller volumes. This can lead to a different ratio of the in- and out-of-plane components, as they are differently localized. Because the in-plane component is more confined to the top of the waveguide a stronger phase shift should be visible in higher demodulation orders. This is of course only true if the lower orders are sensitive to the near-field at the bottom of the waveguide. This sensitivity decreases fast with an increasing waveguide height, as several studies on subsurface near-field imaging imply [117–119].

6. The illumination wavelength may have an influence on the phase shift as well. The wavelength changes the ratio of the in- and out-of-plane components as shown in Fig. 6.3. However, the wavelength also changes the geometrically induced phase shift, caused by the movement of the sample table. The wavelength also affects material properties and is therefore not a good parameter to investigate the origin of the phase shift. Nevertheless, we made several measurements at different wavelengths for similar waveguides, as displayed in Fig. 6.12. From the large variations and uncertainties it is difficult to draw a conclusion. But it seems that the wavelength has no significant influence on the phase shift. It looks like measurements, which were done in one session, show similar phase shifts, which indicates that the tip and interferometer alignment are more important.

7. Similar to the wavelength, the height of the waveguides is also a parameter which could affect the phase shift. But this would again change multiple parameters at

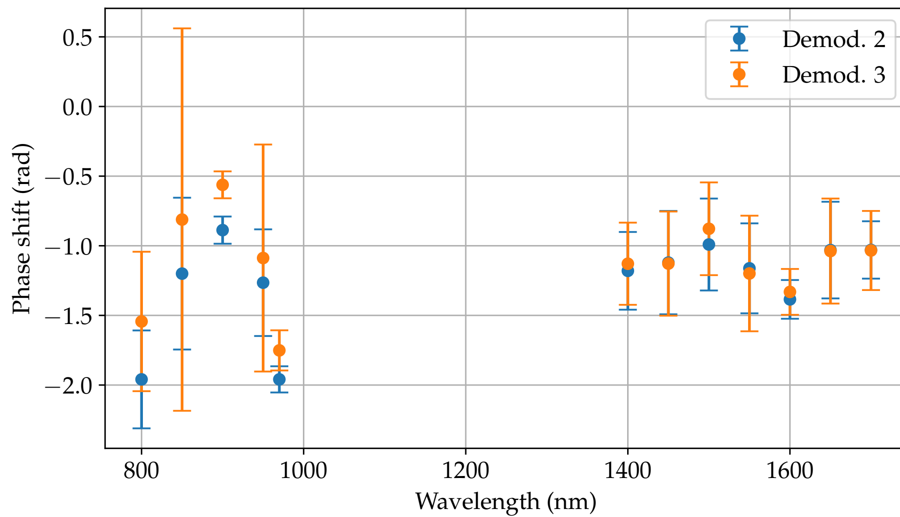


Figure 6.12: Phase shift extracted from various measurements of single waveguides for different excitation wavelengths. Plotted are the phase shifts extracted for the second and third demodulation order. The individual data points are averaged over an arbitrary amount of similar measurements and the errors are error propagated standard deviations.

once, as the height changes the ratio of the in- and out-of-plane components, the polarizability induced amplitude ratio and the geometrically induced phase shift.

8. A material resonances close to the illumination wavelength, either in the sample or tip, can also cause a phase shift [116, 120]. However, for this wavelength range neither gold nor PMMA should have material resonances other than the very broad SPP resonance.

9. If the mechanical properties of a sample changes, the mechanical phase can change as well. This also may affect the phase of the scattered light. Judging by the very minor mechanical phase variation between gold and PMMA, the observed optical phase shift can not be attributed to mechanical shifts.

In conclusion, the phase shift is most likely a combination of multiple effects. The geometrically induced phase shift, due to the sample table moving closer to the light source and the path length to the detector, is certainly one of them. The interference of the in- and out-of-plane components could contribute, but in standard measurements, without an analyzer or if the analyzer is aligned to the out-of-plane component, this effect should be minuscule. This is a consequence of the small ratio of the in-plane component to the out-of-plane component and the polarization sensitivity of the tip. The interference of SPPs excited by the tip itself could also have an influence, but is difficult to quantify. In transmission measurements, standing waves parallel to the waveguide would be an indication for this, but were not observed. Measurements, which are done either in one session or without changing the tip or interferometer alignment, show similar phase shifts. This indicates that the tip and interferometer alignment play a major role. The fact, that the phase shift is typically not

symmetric for both sides of the waveguide underlines this assumption, because the phase shift asymmetry can change depending on the interferometer alignment.

6.1.5 Indication of the In-Plane Component

The previously discussed mode simulations showed, that the waveguide mode has both an in- and out-of-plane component. The out-of-plane component is generally stronger, but the influence of the in-plane component on the detected signal is unknown. So far we did not demonstrate the presence of the in-plane component experimentally. In this section a polarization dependent measurement is performed to gain insight into the influence of the in-plane component on the measurements. We assume that the emission from the tip-sample interaction corresponds to the orientation of the electric field vector at the positon of the tip. The polarization of the emitted light can be determined by rotating the analyzer. The rotation of the analyzer is performed with a home-built rotation mount, that was created with help from Jonas Strubich, a student assistant at the time. If the tip is only sensitive to the out-of-plane component one would expect a maximum signal for an analyzer orientation of 0° and a minimum at 90°, where the angles are defined relative to the z-axis like in Fig. 4.5. The signal should be 180° periodic. If the tip is also sensitive to the in-plane component one would expect a different behavior, especially on top of the waveguide where the in-plane component is strongest. We perform two measurements, where the tip is once besides the waveguide on gold and once on top of the waveguide. At each position we rotate the analyzer and record the demodulated detector signal. In general the background suppression of the interferometer is essential, but this complicates the measurement. Ideally, we would need to rotate the $\lambda/2$-plate in the interferometer to match the analyzer rotation, but this is not possible in our case due to the limited space. Alternatively, one could repeat each single measurements twice, once with the interferometer optimized for the out-of-plane component and once for the in-plane component. The increased complexity is not necessary for a first indication of the in-plane component, because there is a third option. Rotating the interferometer polarization to 45° allows to measure both components simultaneously without adjusting the interferometer polarization. The downside is, that the signal is reduced as the interferometer is not optimized for either component, and the shape of the polarization dependence becomes asymmetric. The asymmetry comes from the fact that any component at the angle of the interferometer will be amplified, while the orthogonal component will be suppressed. For example a signal with a 45° polarization will be amplified, while 135° will be suppressed. The asymmetry is, however, not important for our purpose, as we are mainly interested in the relative strength of the two components.

We perform this measurement at two different wavelengths (970 nm and 1 600 nm). It is expected that the shorter wavelength has a stronger in-plane component, which should lead to a more pronounced effect. The results of these measurements are shown in Fig. 6.13. Before each measurement the analyzer rotation is optimized for the out-of-plane component. This rotation is used as the zero position. It should be noted that the projection of the tip axis in the analzer plane leads to a small rotation offset of roughly 12.7°, see A.1.2.

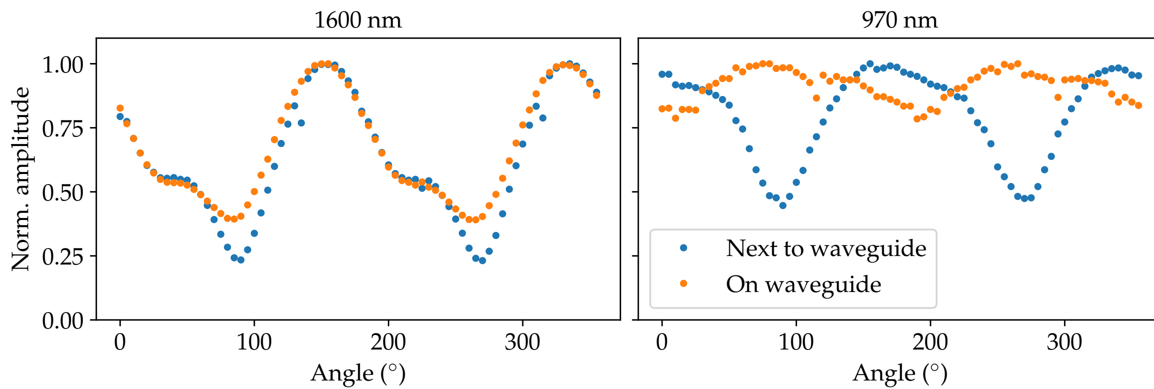


Figure 6.13: Second order demodulated detector signal for an angle variation of the analyzer. Measured at two points, on the waveguide (blue) besides the waveguide (orange) in a distance of 20 µm from the in-coupler. The data points are averaged over 5 cycles with an integration time of 400 ms per point and normalized to emphasize the differences. The measurment is done interferometrically whith the interferometer polarization being rotated to 45° to allow a simultaneous measurement of the in- and out-of-plane polarization projection. Measured were two different waveguides at two different wavelengths.

For the measurement at 1 600 nm we can see only a minor difference between the two positions. The slightly higher signal in the minima on the waveguide indicates the presence of the in-plane component, but the effect is rather small. The expected asymmetry can clearly be seen especially for the 1 600 nm measurement around an angle of 45° and 225°. The measurement at 970 nm on the other hand shows a very clear difference between the two positions. Besides the waveguide we see a clear minimum close to 90° and 270°, which is the expected behavior for a pure out-of-plane component. On the waveguide the minima and maxima are almost exactly antiproportional, indicating a strong influence of the in-plane component. However, the data quality is not sufficient to extract the exact ratio of the two components, but the measurement clearly shows, that the in-plane component can play a significant role. Overall, the measurements confirm our expectations, but the sample was not rotated perfectly, such that the projection of the in-plane component of the waveguide mode is closer to the 45° position of the analyzer than to the 90° position. This amplifies the influence of the in-plane component in this measurement, explaining how the in-plane component on the waveguide can be stronger than the out-of-plane component. Due to time constraints the measurement could not be repeated successfully with a better alignment. However, the measurement still gives a good indication of the influence of the in-plane component.

It would be interesting to measure the in- and out-of-plane components separately in a two dimensional scan. However, measurements of single waveguides are not sufficient, as both components are always present at the same time. There would be only a minor difference in amplitude on the waveguide. In order to separately image both components, we can add a reflection grating to the waveguide, which will create a standing wave pattern

on the waveguide. We will see later that the standing wave pattern is shifted for the two components, allowing us to spatially differentiate them. This approach will be discussed in the next section.

6.2 Waveguide with Reflection Grating

As noted before, adding a reflection grating allows a separate analysis of the two field components, which would not be possible for a standard waveguide. The reflection grating works just like a bragg reflector for the DLSPPW mode [107]. A bragg reflector for a waveguide mode is based on the periodic modulation of the effective mode index. This can be achieved by periodically changing the waveguide dimensions, modifying the underlying metal film or by adding a material in the waveguide with a different refractive index. Here, we use a combination of the first and last method by periodically removing parts of the PMMA waveguide and adding air gaps instead. To improve the modulation, the length of these gratings lines is chosen to be slightly larger than the waveguide width.

Each grating line reflects a certain amount of the incoming mode, leading to interference between all individual reflections. Depending on the grating period these reflections can add up constructively, leading to a significant reflection of the incoming mode. The reflection is maximized when the grating period Λ matches the condition

$$m \cdot \lambda_{\text{eff}} = 2 \cdot \Lambda, \tag{6.1}$$

with m being an integer describing the grating order and λ_{eff} the effective wavelength of the mode. For first order reflection $m = 1$ this means that the grating period has to be half of the effective wavelength of the mode. If the grating period and filling factor are chosen correctly, a significant part of the incoming mode will be reflected.

In order to understand the reflection of the mode and the resulting standing wave patterns of the two components, one can think of the grating as a perfect mirror. Based on the boundary conditions of the electric field upon reflection from a perfect mirror, the tangential components of the electric field have to be zero at the mirror surface, while the normal component can be non-zero and typically decays exponentially in the mirror material. Here, this means that the out-of-plane component has to be zero at the grating surface, while the in-plane component can be non-zero. The resulting reflection of the waveguide mode is therefore mainly determined by the out-of-plane component, which accumulates a phase shift of π upon reflection. We ignore here the influence of the in-plane component on the mode reflection completely. But since both components are strictly linked, the reflected mode also has an in-plane component. Since the two components of the mode have a fixed phase relation of $\pi/2$ and the direction of the mode propagation is flipped, the in-plane component will be reflected with a phase shift of π - π resulting in a net phase shift of 0 relative to the incoming mode. Therefore, the in-plane component shows an amplitude maximum at the grating surface, and the standing wave pattern of the in-plane component is shifted by $0.25 \cdot \lambda_{\text{eff}}$ relative to the out-of-plane component. The amplitude minimum of the

out-of-plane component at the grating surface and the corresponding amplitude maximum of the in-plane component are visible in the simulation results shown in Fig. 6.14. The respective extrema are slightly shifted into the first grating line, this could be due to the finite reflectivity of the grating and the non ideal boundary conditions. The proposed bragg reflector is not a perfect mirror after all.

6.2.1 Simulation

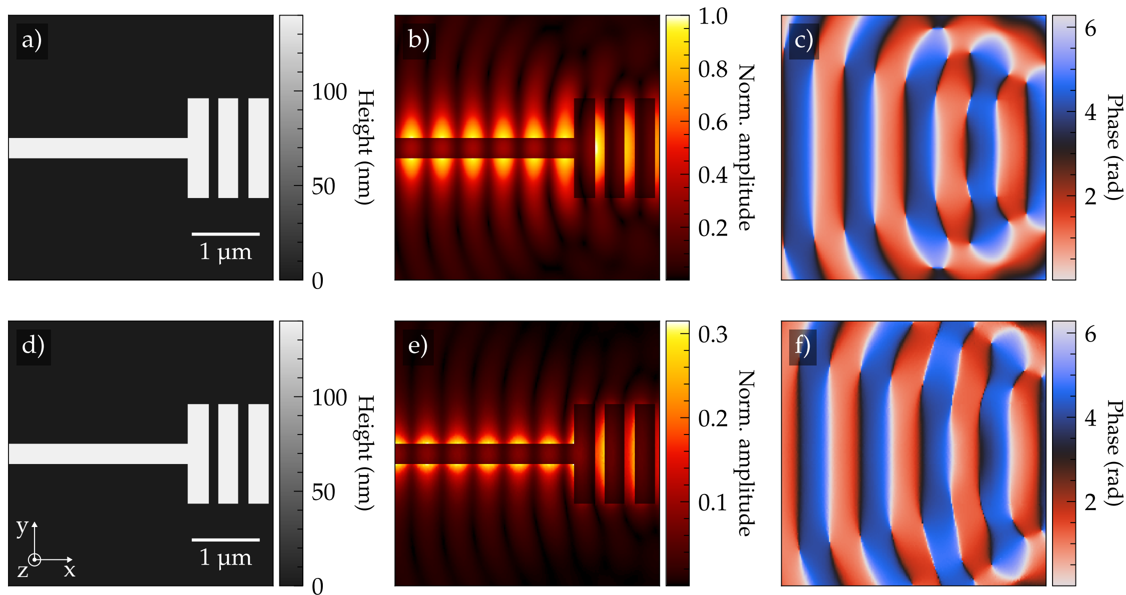


Figure 6.14: Simulation of the reflection grating for the out of plane (b, c) and in plane (e, f) component of the electric field corrected by the relative polarizabilities. (a, d) are identical and represent the simulation geometry.

A simulation [3] of a DLSPPW with a reflection grating is shown in Fig. 6.14. The height data shown in (a) and (d) are identical and created based on the simulation geometry. (b, c) and (e, f) show the amplitude and phase for the out-of-plane component and in-plane component of the electric field respectively. In the measurements the tip probes the near-field at a certain height above the sample surface. Therefore, we extract cut-planes of the electric field 10 nm above the sample surface to create comparable images. The two cutplanes are chosen 10 nm above the metal film and 10 nm above the waveguide respectively. The height map allows to combine both cut-planes to obtain a map of the electric field uniformly 10 nm above the sample surface [4]. The amplitude images are further corrected by the relative

[3] A three dimensional *Comsol* simulation, where the mode profile of a wavgeuide is used as the excitation.

[4] A complete three dimensional simulation with the scattering from the tip at different positions would of course be more accurate, but is computationally very expensive and usually unnecessary.

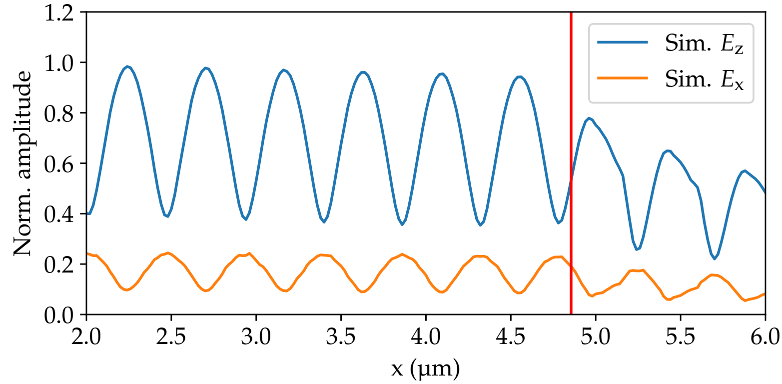


Figure 6.15: Amplitude profiles on the waveguide for the out-of-plane (blue) and in-plane (orange) component of the electric field. The profiles are extracted from the center of the waveguide along the waveguide axis. The horizontal line indicates the center of the first grating line. The amplitude is normalized to the overall maximum.

polarizabilities of gold and PMMA as discussed in Section 4.2.4. One should note though, that this estimation of the relative polarizabilities is only calculated for the dipolar excitation along the tip axis. Meaning, we do not calculate two seperate polarizabilities for the in- and out-of-plane electric field components, only the one for the out-of-plane component. This neglects the tip's polarization sensitivity. A more accurate approach using the respective in- and out-of-plane polarizabilities is exemplified in the appendix A.2.2 but did not lead to meaningful results. The amplitude images show a clear standing wave pattern on the waveguide, which is expected for a reflection grating.

A better view of the standing wave pattern can be seen in Fig. 6.15, where the amplitude profiles on the waveguide are plotted for both components. Here, E_z refers to the out-of- and E_x to the in-plane component. The period of the amplitude maxima corresponds to half of the effective wavelength of the mode. The position of the amplitude maxima is shifted between the two components by approximately $0.25 \cdot \lambda_{\text{eff}}$. This shift of the amplitude maxima is what we want to exploit in the measurements. The phase image does not show a perfectly stationary wave pattern, visible from the remaining phase gradient in x-direction. This is because of the finite reflectivity of the grating and means that a substantial part of the mode is transmitted through the grating. The waveguide behind the grating can therefore still be used to analyze the propagating mode and serves as a reference.

6.2.2 Fabrication

The gratings are created in the same step as the waveguides and thus share the same fabrication method discussed earlier in Section 6.1.2. We have fabricated gratings both with negative and positive tone lithography, the following description is based on the positive tone approach. Inspiration for the design was taken from [39, 121, 122].

The grating is created by writing thin rectangles vertical to the waveguide with a spacing

of the grating period. The rectangles are written as lines from inside out. Similar to the waveguides themselves the border region close to the edges surrounding the reflection grating are written as lines as well. The rest of the area surrounding the waveguide and grating is written as one connected area. This ensures a good edge quality and minimizes the effects of scanning electron microscope (SEM) stage drifts during the exposure. This design is displayed in Fig. 6.16. This image is colorcoded to show the chronological order of

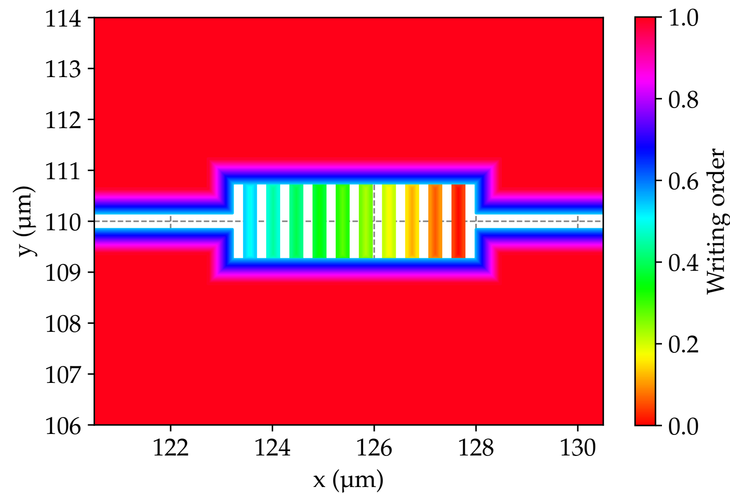


Figure 6.16: Design of the second step of the two step EBL process for the fabrication of DLSPPWs with a reflection grating. In This step the waveguides are written on top of the incoupling gratings and the long gold areas. The grating is written as lines for better control over the final shape. The areas surrounding the waveguides and grating are written as one connected area. Displayed here is the chronological order of the writing process.

the writing process. Examplary AFM scans of two fabricated reflection gratings are shown

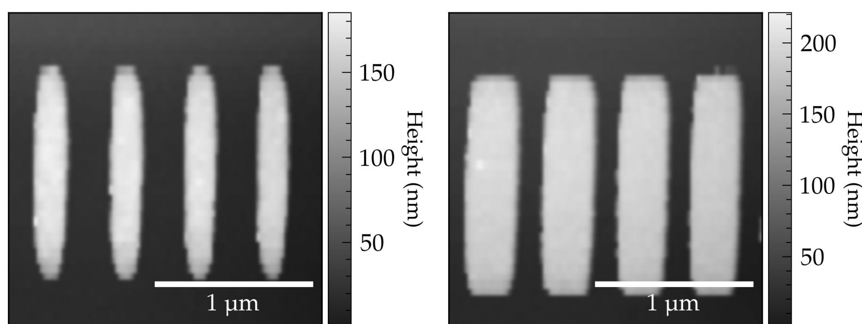


Figure 6.17: Example AFM scan of two reflection gratings with different widths. The PMMA film used here has a height of approximately 175 nm.

in Fig. 6.17.

Final Fabrication Process

The final fabrication process for the DLSPPWs for 970 nm with a reflection grating with positive tone lithography is very similar to single waveguides without a grating. The only difference is the design used in the second step. The casing of the waveguide and reflection grating has a width of 500 nm, is written as lines with a line distance of 20 nm and a dose factor of 0.3. The surrounding area is written as a connected polygon area with a dose factor of 0.155. For a waveguide with a width of roughly 350 nm a design width of 400 nm is used. The reflection grating is written as rectangles of individual lines vertical to the waveguide with a line distance of 20 nm and a dose factor of 0.35. The period of the grating is chosen to be half the wavelength of the dielectric-loaded surface plasmon polariton (DLSPP) mode (440 nm). The design width of the rectangles is half of the grating period (220 nm). The length of the rectangles of the reflection grating were initially 1.5 µm but later increased to 4 µm to increase the reflectivity. The length has a minor but non negligible influence on the necessary dose factor.

6.2.3 Measurement

In order to confirm the simulations we conduct s-SNOM measurements of a DLSPPW with a reflection grating. The grating has only 6 periods to allow a part of the mode to propagate further. The reason being that this allows to directly compare the phase of a standing and propagating wave. An example measurement is shown in Fig. 6.18. The measurement shows the third harmonic demodulated amplitude and phase at a wavelength of 970 nm for an analyzer position optimized for the out-of-plane component. The height image helps to visualize the incoupling grating and the waveguide with the reflection grating. The amplitude image shows a standing wave pattern on the waveguide, indicating that the reflection grating is working. Behind the reflection grating no interference is visible and only the pure propagating mode can be seen. The phase image shows a linear phase gradient on the waveguide before and after the reflection grating, which is expected for a propagating mode. Close to the reflection grating the phase gradient decreases slightly, indicating the presence of the standing wave pattern. The phase on the waveguide is a superposition of the linear phase gradient of a propagating mode and the partially reflected wave. Closer to the reflection grating the influence of the reflected mode is stronger because the reflected mode decays and is weaker than the incident mode. Since the phase gradient of the incoming and reflected mode have an opposite sign, they cancel each other out. But the phase gradient never completely vanishes, indicating that the reflection is not perfect. It is interesting to note that the phase shift on the waveguide relative to next to the waveguide never diminishes completely, even in close proximity to the reflection grating. This indicates that the phase shift can not be explained solely by the interference of the in- and out-of-plane components, as the fringe positions clearly indicate a dominant out-of-plane component. Since in this case the field on and next to the waveguide is dominated by the out-of-plane component, no interference between the two components is expected. In order to investigate the influence

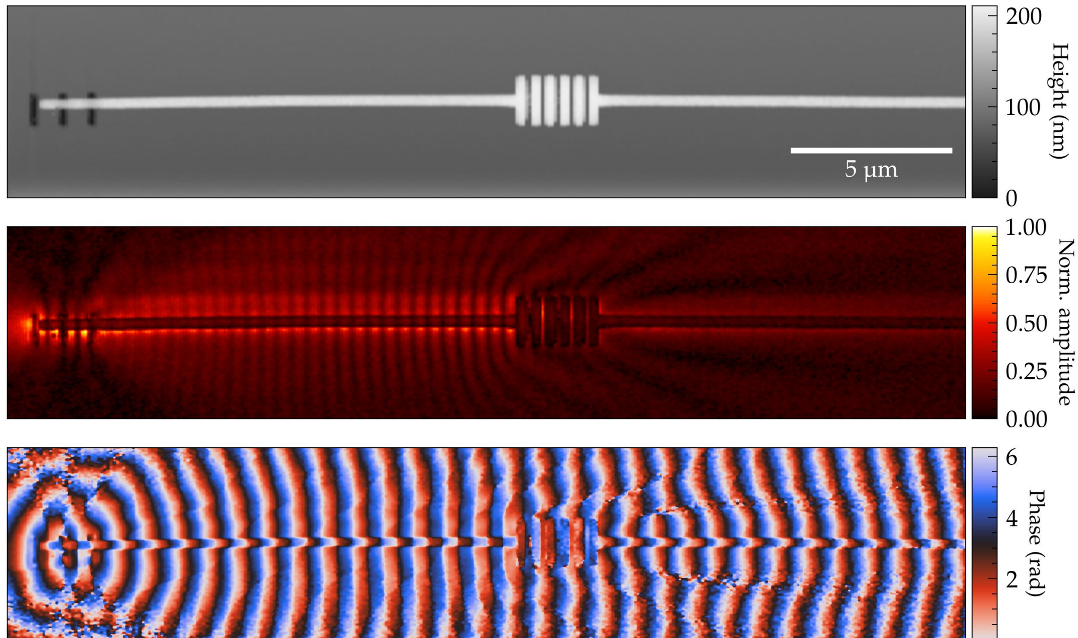


Figure 6.18: s-SNOM measurement of a DLSPPW with a reflection grating. The reflection grating has only 6 periods to allow a part of the mode to propagate further. Shown is the third harmonic demodulated amplitude and phase at a wavelength of 970 nm. The height data is leveled and all channels are filtered with a gaussian filter for a cleaner image.

of the in- and out-of-plane components further, we rotate the analyzer in front of the detector to be either parallel to the out-of-plane component or the in-plane component.

The two measurements are shown in Fig. 6.19. The measurement in the upper row is taken with the analyzer optimized for the out-of-plane component, while the measurement in the lower row is taken with the analyzer optimized for the in-plane component. Both waveguides are rotated by 45° to seperate the projected polarizations in the analyzing plane as described in Section 4.3. The amplitude images show a clear difference in the position of the amplitude maxima on the waveguide, which is expected from the simulations. In contrast, the maxima on the sides of the waveguide are not shifted in the two measurements. The reason for this is the different confinement of the two components discussed in Section 6.1.1. The in-plane component is mainly localized on the top of the waveguide, while the out-of-plane component extends further laterally. Thus, next to the waveguide the out-of-plane electric field is dominant in both measurements. Additionally, the tip is more sensitive to the out-of-plane component as described in Section 4.2.4. The amplitude next to the waveguide is suppressed by the analyzer in the in-plane measurement. However, it is still present in the measurements, which indicates that the analyzer position might not have been perfect. Therfore, the same standing wave pattern next to the waveguide can be observed in both measurements.

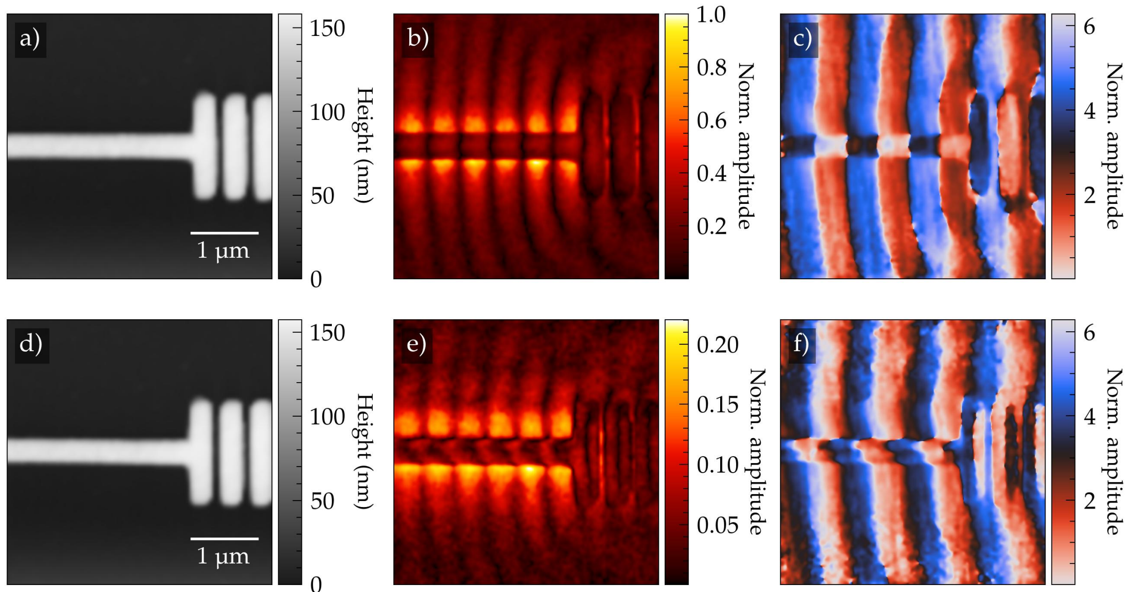


Figure 6.19: s-SNOM measurement of a DLSPPW with a reflection grating (O3). The two measurements are taken with the analyzer optimized for the out of plane component (upper) or the in plane component (lower). The height data is leveled and all channels are filtered with a gaussian filter for a cleaner image.

The phase images are overall relatively similar, but an asymmetry on the waveguide can be seen for the measurement of the in-plane component. The asymmetry seems to be a measurement artifact, as the amplitude image is affected as well.

The simulation of the reflection grating in Fig. 6.14 does not match the measurements perfectly. The reason for this is that in the simulation the in- and out-of-plane components are perfectly seperated, while in the measurements both components are always present. To better describe the measurements we can combine the two components with a certain ratio to emulate a non perfect analyzer position. We can infer the ratio from the measurements by comparing the maximum amplitude next to the waveguide in both measurements. Modifying the simulation of the in-plane component by this ratio leads to Fig. 6.20. The modified simulations (a) and (b) show a good qualitative agreement with the measurements (c) and (d) respectively, and thus confirm our understanding of the dynamics.

As a quantitative measure, we compare the shift between the two standing wave patterns on the waveguide for the two measurements shown in Fig. 6.20. As a reference point, the center position of the first grating line is used, which is extracted from a height profile parallel to the waveguide. Two shifted amplitude profiles taken on the waveguides are shown in Fig. 6.21. Although the shift between the two profiles is visible, the center of the grating line is difficult to determine, so instead a profile prallel the one on the waveguide but on the gold film is used as the reference. This profile corresponds to the out-of-plane

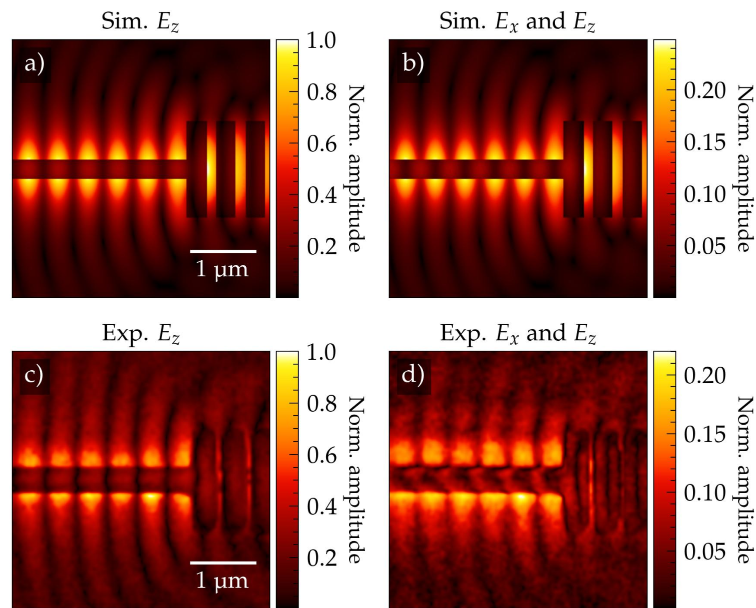


Figure 6.20: Direct comparison of the measurements and simulations of the reflection grating. The first row (a, b) shows the modified simulation data, and the second row (c, d) the measurements. The simulation data in (a) is the combination of the out of plane component above the gold film and waveguide respectively, corrected by the relative polarizabilities. The simulation data in (b) is a combination of the out-of-plane component above the gold film combined with the in-plane-component above the waveguide corrected by the relative polarizabilities. The out-of-plane component above the gold film is further multiplied with a factor to emulate the reduction of this component due to the analyzer. (c, d) are the respective measurements of the out-of-plane and in-plane component.

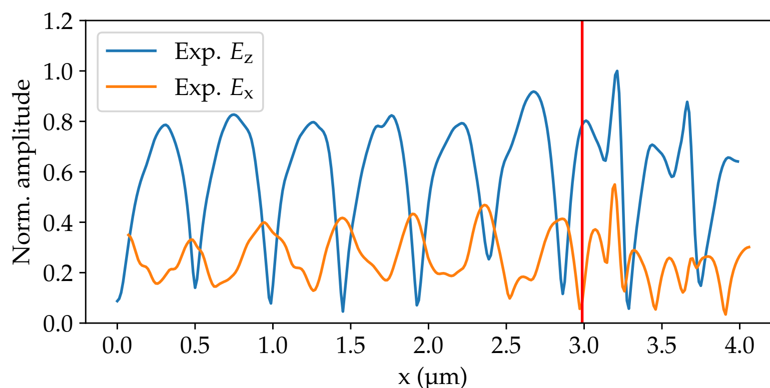


Figure 6.21: Amplitude profiles on the waveguide for the out-of-plane (blue) and in-plane (orange) component of the electric field. The profiles are extracted from the center of the waveguide along the waveguide axis and are shifted to align the center of the first grating line. The horizontal line indicates the center of the first grating line. The amplitude is normalized to the overall maximum.

component and the peak positions in this profile are independent of the analyzer rotation. The average amplitude maxima positions from the profiles on the waveguide are determined relative to the respective reference profile. From these positions the average peak to peak distance and the average shift between the two profiles on the waveguide can be calculated. The average peak to peak distance corresponds to half the effective wavelength of the mode. The average shift between the two profiles on the waveguide should be $0.25 \cdot \lambda_{\text{eff}}$ if we are indeed measuring the in- and out-of-plane components seperately.

The average distance of the amplitude maxima is approximately 463 nm, which corresponds to an effective wavelength of $\lambda_{\text{eff}} = 926\,\text{nm}$ for a free space wavelength of 970 nm. For the measurement of the out-of-plane component we get a shift of $0.032 \cdot \lambda_{\text{eff}}$ between the maxima on and next to the waveguide. It shows that the method is robust enough for a quantitative assessment. For the measurement of the in-plane component the average shift is approximately $0.26 \cdot \lambda_{\text{eff}}$. This is less than half the measurement pixelsize (50 nm) above the expectation and a strong indication that we are indeed measuring the in-plane and out-of-plane components of the electric field seperately. The small deviation can be explained by the uncertainty in the determination of the maxima positions, as the amplitude profiles are quite noisy and slightly asymmetric, especially for the second measurement.

In future work it would be interesting to determine the relative strength of the two components. It is challenging due to the different scattering efficiencies of the two components and other influences like the rotation of the analyzer, the bad signal quality and the strong polarization dependence of the half wave plates used in the setup.

6.3 Conclusion

In this chapter DLSPPWs have been investigated in detail, especially with a focus on their use in s-SNOM. The mode profile, localization and in- and out-of-plane components have been studied. The out-of-plane component was found to be dominant, but the in-plane component can have a significant impact. This is often overlooked in literature, but especially in polarization sensitive measurements it can play an important role. The unexpectedly low amplitude on the waveguides in s-SNOM measurements could be explained via the change in the effective polarizabilities of the coupled tip-sample system for gold and PMMA. The phase shift on the waveguide is not fully understood, but many contributions have been identified and studied. This shows that the tip-sample interaction is more complex than often assumed, especially since the samples here are not flat. A reflection grating was designed to measure the in-plane component seperate from the out-of-plane component. Together with simulations a good agreement was found, indicating that the in-plane component can be measured seperately. The fabrication of incoupling gratings, single waveguides and waveguides with a reflection grating was discussed in detail, as this was a big part of the work. During the fabrication multiple problems occured, which are also discussed and led to the developement of a new robust method based on positive tone EBL with PMMA.

CHAPTER 7

Coupled DLSPP Waveguides

Building on the understanding of single plasmonic waveguides, coupling between two or more waveguides arises when their modal fields spatially overlap. This interaction leads to the hybridization of the individual modes, forming new collective modes and giving rise to characteristic phenomena such as the periodic exchange of optical energy between adjacent waveguides. Combined with the properties of single waveguides, such coupling effects can be exploited to realize a variety of functional photonic components, including waveguide crossings [123], splitters [107, 124, 125] and wavelength selective elements like waveguide ring resonators [126]. Even active components such as electro-optic and thermo-optic modulators have been demonstrated [127, 128].

Coupled plasmonic waveguides share many similarities with their purely dielectric counterparts. However, the much stronger field confinement in plasmonic systems results in significantly enhanced coupling strengths. At first, the higher confinement might seem disadvantageous due to the shorter range of interaction, but it leads to significantly increased electric field intensities in the evanescent field and allows to reduce the distance between waveguides significantly. This enables devices with a substantially reduced footprint, albeit at the expense of increased propagation losses due to the intrinsic absorption in metals. To counter higher losses in plasmonic waveguides, interfaces between dielectric and plasmonic waveguides can be employed to combine the low-loss propagation in dielectric waveguides with the strong field confinement in plasmonic waveguides [129].

In this chapter, we investigate coupled dielectric-loaded surface plasmon polariton waveguides (DLSPPWs) using near-field microscopy. We begin with the simplest case of two coupled DLSPPWs and develop robust methods for analyzing scattering-type scanning near-field optical microscope (s-SNOM) data to quantitatively extract key parameters, such as the coupling length and phase information. Subsequently, we extend this approach to study arrays of coupled waveguides that implement the Su-Schrieffer-Heeger (SSH) model, a prototypical one-dimensional topological system supporting localized edge states. Using s-SNOM measurements, we directly probe these edge states and analyze their modal distribution.

7.1 Two Coupled Waveguides

If two waveguides are in close proximity, their evanescent fields can overlap, leading to a coupling between the waveguides. The coupling strength primarily depends on the distance between the waveguides and their mode profiles. In this section, we investigate the coupling between two DLSPPWs using s-SNOM measurements. We develop a robust method to analyze the s-SNOM data to extract the coupling length and phase effects.

7.1.1 Theory

The coupling between two waveguides can be described by the coupled mode theory (CMT) [130], in the limit of strongly coupled waveguides the CMT needs to be modified [131]. The amplitudes A_1 and A_2 of the modes in the two waveguides evolve along the propagation direction x according to the following set of differential equations

$$\frac{dA_1(x)}{dx} = i\beta_1 A_1(x) + iC_{12}A_2(x)\,, \tag{7.1}$$

$$\frac{dA_2(x)}{dx} = i\beta_2 A_2(x) + iC_{21}A_1(x)\,, \tag{7.2}$$

where β_1 and β_2 are the complex-valued propagation constants of the individual waveguides and C_{12} and C_{21} are the coupling coefficients between the two waveguides. The coupling coefficients C_{12} and C_{21} depend on the overlap of the evanescent fields of the modes in the two waveguides and thus on the distance between the waveguides and their mode profiles. They are in principle dependent on the in- and out-of-plane components of the fields [130], but due to the stronger confinement of the in-plane component in the transversal direction its influence is neglected here. The equations above describe the evolution of the fields in the two waveguides. If one assumes identical waveguides ($\beta_1 = \beta_2 = \beta$ and $C_{12} = C_{21} = C$), the solution of the coupled mode equations leads to a periodic exchange of energy between the two waveguides with a coupling length L_C given by [130]

$$L_C = \frac{\pi}{2C}\,. \tag{7.3}$$

The coupling length L_C is the distance along the propagation direction over which the energy is transferred from one waveguide to the other.

In order to get the coupling coefficients and coupling length one can perform mode simulations of the coupled waveguide system, as displayed in Fig. 7.1. When two waveguides are brought close to each other, the modes of the individual waveguides hybridize and form new symmetric (a) and antisymmetric (b) eigenmodes. The propagation constants of these eigenmodes differ from those of the individual waveguides. Exciting only one waveguide initially corresponds to an equal superposition of the two modes. The small difference in the propagation length of the two modes leads to a beating on each waveguide with a beat

length L_B given by [132]

$$L_B = \frac{2\pi}{\beta_s - \beta_a}, \tag{7.4}$$

where β_s and β_a are the propagation constants of the symmetric and antisymmetric eigenmodes, respectively. The envelope of the beating is maximal when both modes interfere constructively. But since one of the modes is asymmetric, the beating of the second waveguide is shifted by half a beating period. This is because the asymmetric mode has opposite phase in the two waveguides and flips its sign once per beating period. Thus, when the modes interfere constructively in one waveguide, they interfere destructively in the other waveguide. This leads to the periodic exchange of energy between the two waveguides.

From the *Comsol* mode simulation, performed at a wavelength of $\lambda = 970\,\text{nm}$, of two waveguides with identical dimensions (width $= 300\,\text{nm}$, height $= 150\,\text{nm}$) and a center to center distance of 600 nm, we get the values $\beta_s = 2\pi/\lambda \cdot (1.12 + 0.005i)$ and $\beta_a = 2\pi/\lambda \cdot (1.04 + 0.0048i)$, leading to a beat length of $L_B = 12.5\,\mu\text{m}$. The coupling length L_C is then related to the beat length by

$$L_C = \frac{L_B}{2}, \tag{7.5}$$

resulting in a coupling length of $L_C = 6.25\,\mu\text{m}$ and a coupling coefficient of $C = 0.251\,\mu\text{m}^{-1}$ from the *Comsol* simulations.

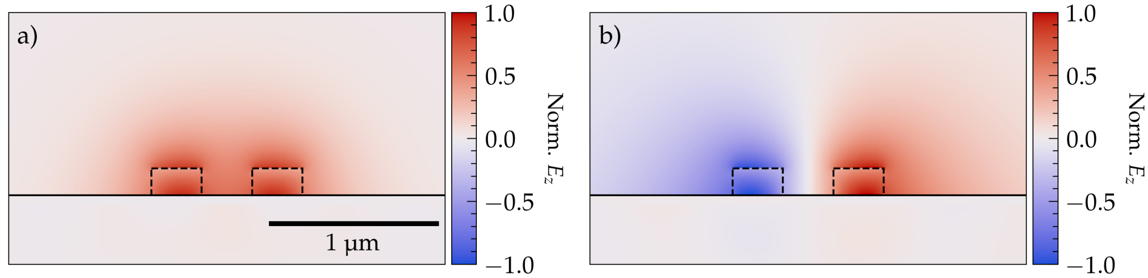


Figure 7.1: Simulated mode profiles of two coupled DLSPPWs with a center to center distance of 600 nm. The color scale represents the out-of-plane electric field component E_z. The mode profiles show the symmetric (a) and antisymmetric (b) eigenmodes formed by the coupling between the two waveguides. The dashed lines indicate the position of the PMMA waveguides and the horizontal line corresponds to the gold surface.

7.1.2 Fabrication

The fabrication process is similar to the single waveguides discussed in Section 6.1.2, with some minor changes due to the proximity effect in electron beam lithography (EBL) for closely spaced waveguides. The incoupling gratings are identical to the ones used for the single waveguide samples and are discussed in Section 3.4. All measurements shown here are performed on samples fabricated with a combination of the focused ion beam FIB

and negative tone EBL process. An alternative approach using positive tone EBL was implemented as well and is discussed in the appendix A.2.5.

Both waveguides are written as individual lines from inside out with a linespacing of 5 nm. They are written parallel to each other and not one after the other [1]. This reduces proximity effects, which are usually more pronounced on already exposed structures. The secondary waveguide without the incoupler is offset by 5 μm to mitigate its excitation by the illuminating laser. The parameters used for the samples measured in Section 7.1.3 are summarized in Table 7.1. The other steps like the the focused ion beam (FIB) milling, spin coating of the resist, the development, and the resist removal are identical to the single waveguide fabrication process described in Section 6.1.2.

#	Distance (nm)	Height (nm)	Width (nm)	Design width (nm)	Dose factor
1	500	100	300	350	0.62
2	500	120	300	350	0.67
3	500	140	300	350	0.71
4	500	150	300	350	0.83
5	600	100	300	350	0.64
6	600	120	300	350	0.69
7	600	140	300	350	0.73
8	600	150	300	350	0.86
9	1000	100	300	350	0.68
10	1000	120	300	350	0.73
11	1000	140	300	350	0.77
12	1000	150	300	350	0.94

Table 7.1: Parameters for the fabrication of two coupled waveguides with negative tone EBL with PMMA.

7.1.3 Measurements

To investigate the coupling between two DLSPPWs, the samples (4, 8, 12) according to Table 7.1 are fabricated. All measurements are performed at 970 nm. In Fig. 7.2, an example s-SNOM measurement of two coupled DLSPPWs with a center to center distance of 600 nm is shown. The height data shows the incoupling grating on the left end and the two waveguides extending to the right. The amplitude image reveals the periodic exchange of amplitude between the two waveguides along the propagation direction due to the coupling between them. The phase image shows a linear phase gradient in each waveguide indicating the propagation of a single mode in each waveguide.

To further investigate the coupling between the two waveguides, we performed s-SNOM measurements on samples with different distances between the waveguides. In Fig. 7.3,

[1] First line in waveguide one, then first line in waveguide two, then second line in waveguide one etc.

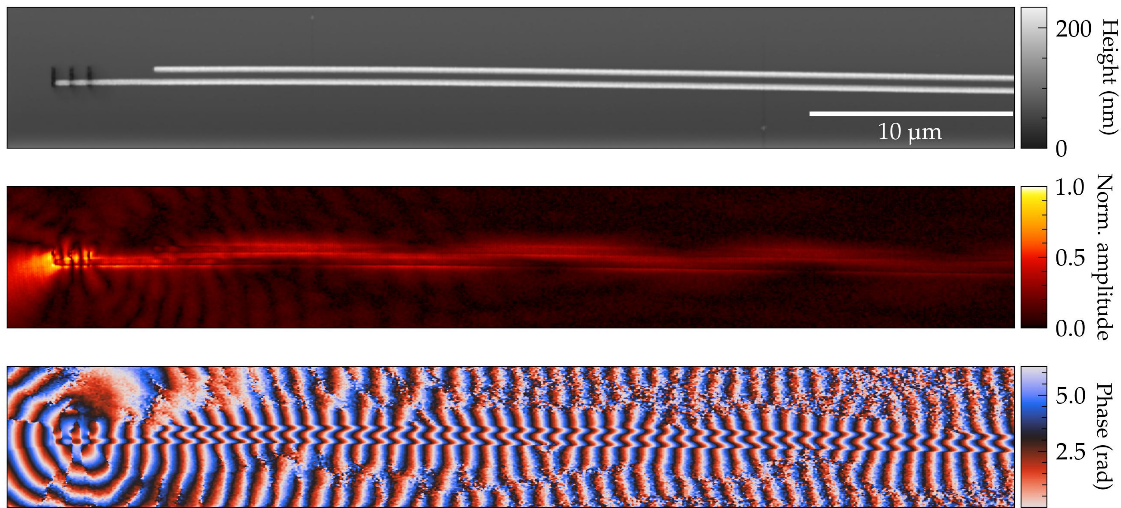


Figure 7.2: s-SNOM measurements of two coupled DLSPPWs with a center to center distance of 600 nm, a width of 300 nm and a height of 150 nm. Shown is the height (a) amplitude (b) and phase data (c). All images have been gauss filtered for a better clarity.

s-SNOM measurements of three different samples with center to center distances of 500 nm, 600 nm, and 1 000 nm are shown. As the distance between the waveguides increases, the coupling strength decreases, leading to a longer coupling length. This is evident from the amplitude images, where the periodic exchange of amplitude between the two waveguides occurs over a longer distance for larger separations.

To quantitatively analyze the coupling between the two waveguides, we extracted the amplitudes of the modes in each waveguide from the s-SNOM measurements. The amplitude of the two waveguides can be extracted in different ways. The simplest method is to take a line-profile along the center of each waveguide. However, this method can be sensitive to noise and may not accurately capture low amplitude values. Especially on the waveguides the amplitude signal is low, due to the low polarizability of the waveguide material, as discussed in Section 4.2.4. The different polarizabilities lead to an amplitude distribution per waveguide that shows similarities to a double Gaussian. Therefore, we also employed a more robust method by fitting a double Gaussian distribution per waveguide to the amplitude profile across the waveguides at each position along the propagation direction. An example of this technique is shown in Fig. 7.4 (b, d). In the example the waveguide separation is sufficent to distinguish both waveguides, but especially for the smaller distances the

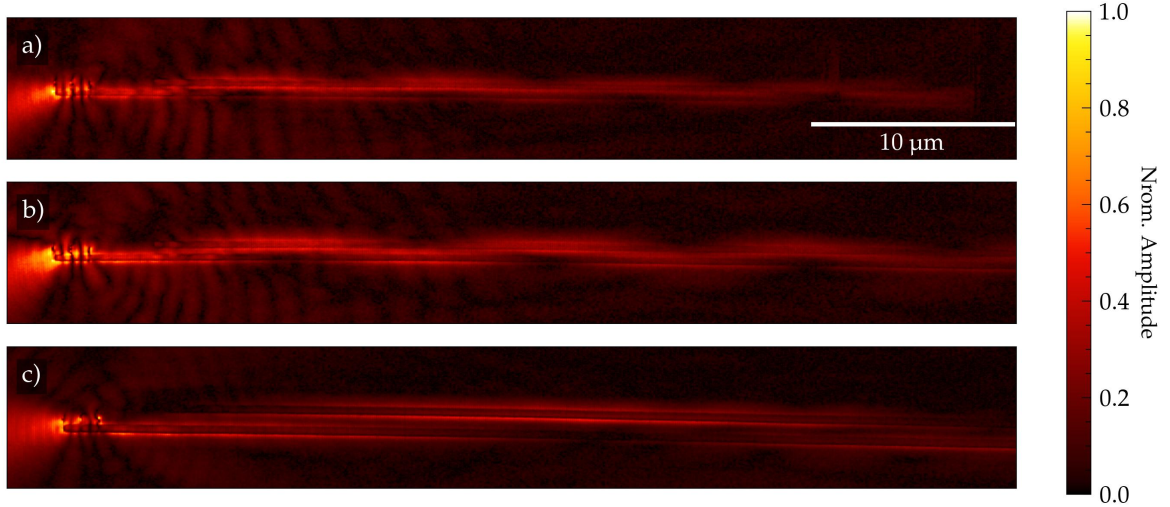


Figure 7.3: s-SNOM measurements of two coupled DLSPPWs with varying distance between the waveguides. The center to center distances are 500 nm (a), 600 nm (b), and 1 000 nm (c). The amplitude of the two waveguides is periodically exchanged along the propagation direction due to the coupling between the waveguides. The coupling length increases with increasing distance between the waveguides. Shown is the second demodulation order and the amplitudes have been gauss filtered for a better clarity.

amplitude distributions overlap. The fitting function is given as

$$f(x) = y_0 + f_1(x) + f_2(x),$$
$$f_{1,2}(x) = \frac{a_{1,2}}{\sigma * \sqrt{2\pi}} * \left[\exp\left(-\frac{(x - x_{1,2} - x_{\text{sep}}/2)^2}{2\sigma^2}\right) + \exp\left(-\frac{(x - x_{1,2} + x_{\text{sep}}/2)^2}{2\sigma^2}\right)\right],$$

where a_1 and a_2 are the amplitudes of the two double Gaussians, x_1 and x_2 are the center positions of the two waveguides, x_{sep} is the separation between the two Gaussians per waveguide, σ is the standard deviation of the Gaussians, and y_0 is an offset. First the position of the waveguides is extracted by fitting two Gaussians like $f_1(x)$ to the height data, as exemplified in Fig. 7.4 (a, c). Then, the fit of $f(x)$ is performed at each position along the propagation direction to get initial estimates for σ and x_{sep}. Alternatively, the fit parameters could be optimized using the measurement of a single waveguide with identical dimensions. But experimentally, we find that the shape of the amplitude distribution can vary significantly between different measurements, likely due to changes in the tip and sample conditions. For example the tip degrades over time, and due to the proximity effect during fabrication the waveguide dimensions can vary slightly. Therefore, we use the described fit parameter estimation. Finally, the fit is repeated with fixed and averaged parameters and only the amplitudes $a_{1,2}$ and the offset y_0 are fitted. In Fig. 7.5, the extracted amplitudes of the two coupled waveguides with a center to center distance of 600 nm are shown.

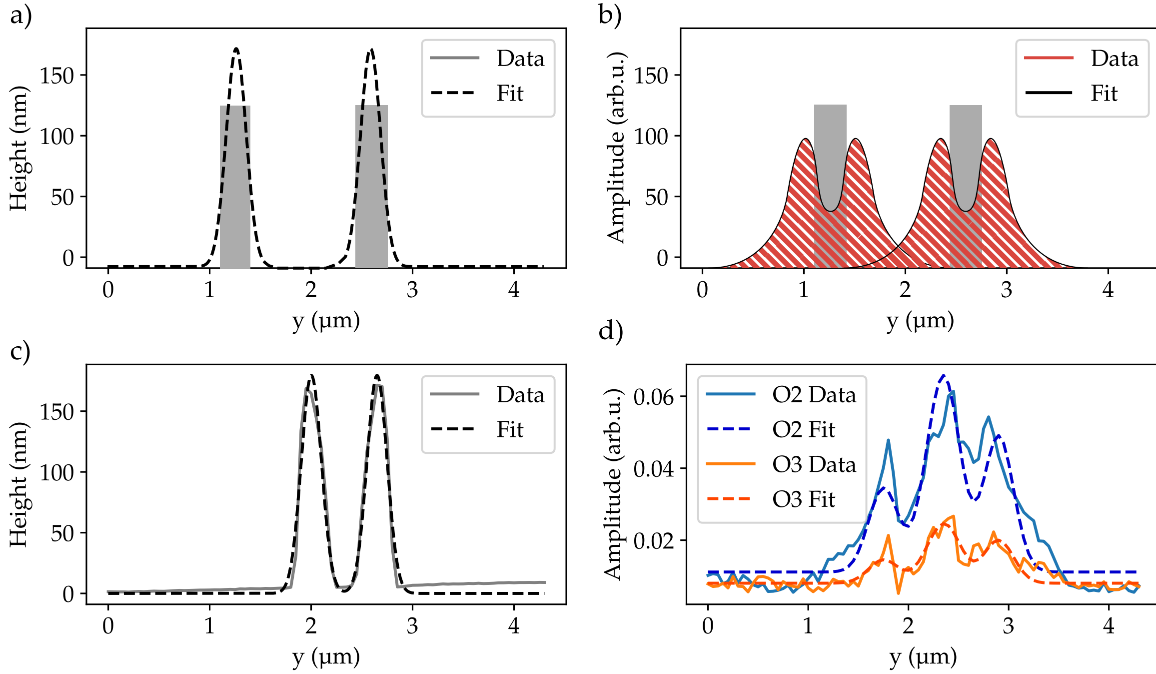


Figure 7.4: Sketch and example data for the fit procedure of the height and amplitude data of the two coupled waveguide measurements. Sketches of the height fit using two individual Gaussian distributions (a), and the double Gaussian-like amplitude distribution of the two waveguides (b). Example cutline data at an arbitrary position along the propagation direction for the height (c) and amplitude (d) data of two coupled DLSPP waveguides with a center to center distance of 600 nm. The height data is fitted with a double gaussian to extract the position of the two waveguides. The amplitude data is fitted with a sum of two double gaussians to extract the amplitudes of the two waveguides.

In order to reduce noise, the amplitude data is averaged over 1 µm intervals. Both methods yield similar results, confirming the periodic exchange of amplitude between the two waveguides. The double Gaussian fit method provides a more robust estimate of the amplitudes, especially in regions where the amplitude is low and noise can significantly affect the line-profile extraction. A slight amplitude offset between the two waveguides in Fig. 7.5 can be noticed, which is likely due to the collection efficiency of the s-SNOM tip varying with position. The second waveguide partially blocks the light that is scattered from the tip and reflected from the sample, leading to a reduced collection efficiency for the lower waveguide.

To gain further insights into the coupling dynamics between the two waveguides, we reconstructed the real part of the electric field from the s-SNOM measurements. This is achieved by combining the fitted amplitude data with the phase data extracted from the middle of each waveguide. The resulting real part of the electric field for the two coupled DLSPPWs with a center to center distance of 600 nm is shown in Fig. 7.6. The periodic beating pattern resulting from the interference of the symmetric and antisymmetric modes

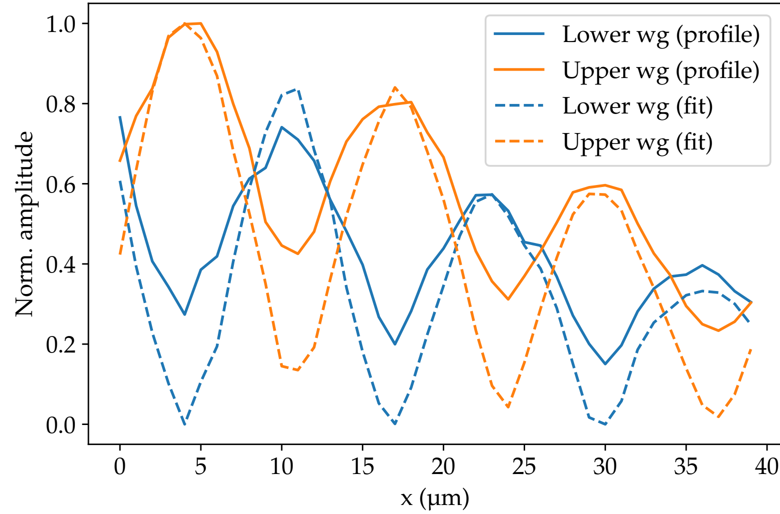


Figure 7.5: Extracted amplitudes of the two coupled waveguides with a center to center distance of 600 nm from the s-SNOM measurement shown in Figure 7.2. The amplitudes are extracted in two ways. From a cutline along the center of each waveguide and from a repeated fits of two double Gaussians along the waveguides.

is clearly visible, confirming the coupling dynamics between the two waveguides. In order to better understand the coupling dynamics and accurately extract the coupling length, we analyse the measurement for the case of 600 nm waveguide separation in more detail.

In Fig. 7.7, we compare the real part of the electric field of the upper waveguide with numerical simulations. The numerically calculated data is based on the beating of two modes with the mode indices from the *Comsol* simulations. The numerical data is shifted to account for the arbitrary starting position. Both data sets show the beating pattern resulting from the interference of the symmetric and antisymmetric modes, confirming the coupling dynamics between the two waveguides. The agreement between the numerical and reconstructed data validates our analysis methods and provides insights into the coupling behavior in the DLSPPW system. A slight mismatch in the beating can be seen, indicating that the mode indices taken from the *Comsol* simulations are not exactly matching the experimental values. In order to extract the mode indices from the experimental data, we employ two different methods. First, a Fourier analysis of the real part data on the waveguide is performed and secondly, the numerical model is fitted to the experimental data.

The Fourier transform is displayed in Fig. 7.8. We expect two main spatial frequencies in the fourier transform, corresponding to the two modes. The peak position of the two main features are equivalent to the mode wavelengths and can be used to calculate the mode indices. The propagation constants extracted from the Fourier analysis of the s-SNOM measurements are $\beta_s = 2\pi/\lambda \cdot (1.125)$ and $\beta_a = 2\pi/\lambda \cdot (1.05)$, which are slightly larger than the values obtained from the *Comsol* simulations ($\beta_s = 2\pi/\lambda \cdot (1.12 + 0.005i)$ and $\beta_a = 2\pi/\lambda \cdot (1.04 + 0.0048i)$). This discrepancy could be due to fabrication imperfections,

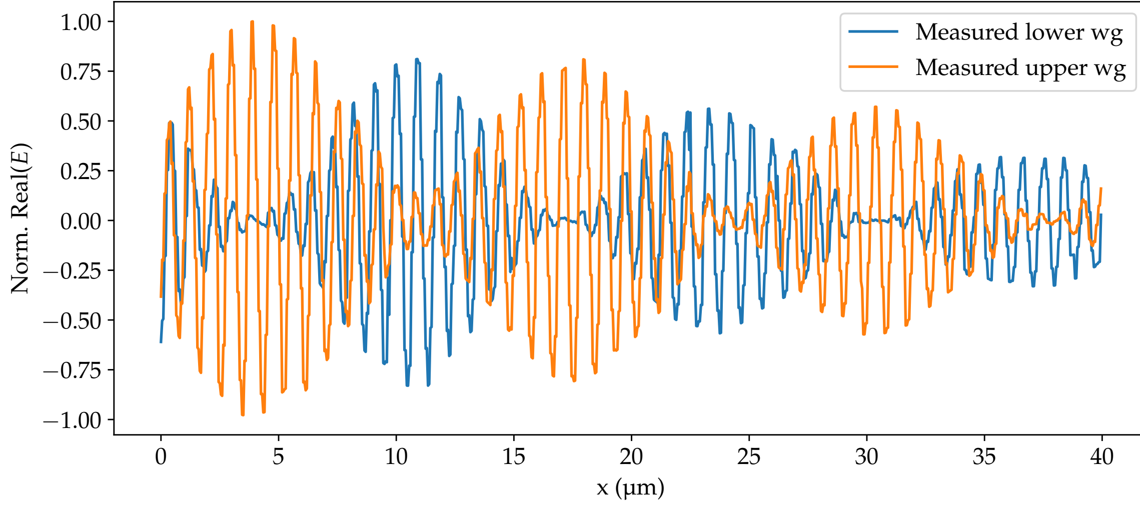


Figure 7.6: Real part of the electric field extracted from the s-SNOM measurements of two coupled DLSPPWs with a center to center distance of 600 nm. The electric field is reconstructed by combining the fitted amplitude data with the phase data. The periodic beating pattern resulting from the interference of the symmetric and antisymmetric modes is clearly visible.

material properties, or other experimental factors not accounted for in the simulations. Additionally, we can calculate the coupling length and coupling coefficient from the extracted mode indices, yielding $L_C = 6.7\,\mu\text{m}$ and $C = 0.234\,\mu\text{m}^{-1}$, which is, again, slightly larger than the value 6.25 μm obtained from the *Comsol* simulations. But overall the agreement between the measurements and simulations is quite good.

We now perform the second analysis method by fitting the numerical model to the reconstructed data from the s-SNOM measurements. The model is based on the sum of two propagating waves, corresponding to the symmetric and antisymmetric modes. The equtions for this are given by

$$E_\text{s}(x) = \exp\left(i\beta_\text{s}\left(x + x_0\right)\right)\,, \tag{7.6}$$

$$E_\text{a}(x) = \exp\left(i\beta_\text{a}\left(x + x_0\right)\right)\,, \tag{7.7}$$

$$E_\text{total}(x) = A\left(E_\text{s}(x) + E_\text{a}(x)\right)\,, \tag{7.8}$$

where A is the amplitude prefactor, $E_\text{s}(x)$ and $E_\text{a}(x)$ are the electric fields and β_s and β_a the mode indices of the symmetric and antisymmetric modes, respectively. The total electric field $E_\text{total}(x)$ is fitted to the reconstructed data from the s-SNOM measurements.

The fitting procedure yields propagation constants of $\beta_\text{s} = 2\pi/\lambda \cdot (1.131 + 0.0032i)$ and $\beta_\text{a} = 2\pi/\lambda \cdot (1.052 + 0.0044i)$, averaged over the two waveguides. The two individual fits are shown in Fig. 7.9. With this method we also get an estimate for the imaginary part of the mode indices, which are slightly lower than the values obtained from the *Comsol* simulations. This is unexpected, as usually experimental values show increased losses due to sample

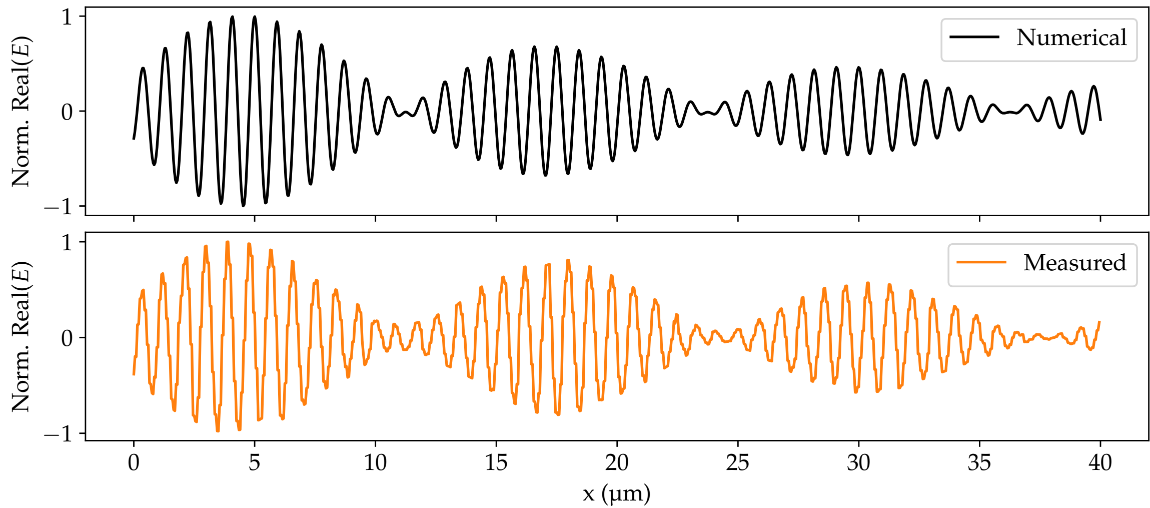


Figure 7.7: Comparison of the real part of the electric field of the upper waveguide, numerically calculated for the mode indices taken from *Comsol* simulations (top) and the real part of the electric field reconstructed from the s-SNOM measurements (bottom) of two coupled DLSPPWs with a center to center distance of 600 nm.

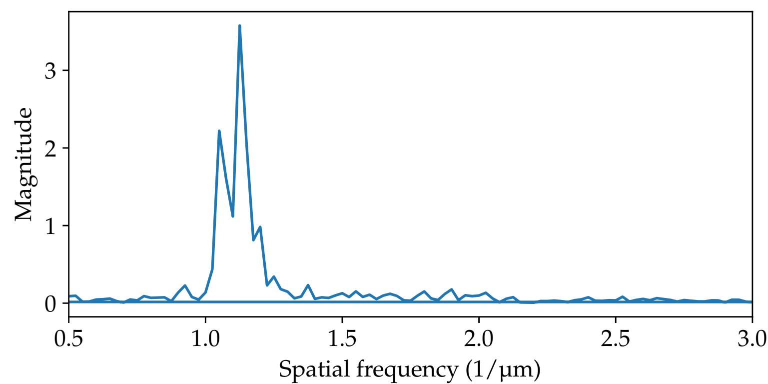


Figure 7.8: Fourier transform of the real part of the electric field extracted from the s-SNOM measurements of two coupled DLSPPWs with a center to center distance of 600 nm.

imperfections. The low losses could however be explained by a slightly lower coupling strength, leading to less losses for the asymmetric mode, and a different field distribution due to a slightly different waveguide shape compared to the simulations. The coupling length and coupling coefficient calculated from the fitted mode indices are $L_C = 6.33\,\mu\text{m}$ and $C = 0.248\,\mu\text{m}^{-1}$, which is in good agreement with the value obtained from the Fourier analysis ($L_C = 6.7\,\mu\text{m}$, $C = 0.234\,\mu\text{m}^{-1}$) and from the *Comsol* simulations ($L_C = 6.25\,\mu\text{m}$, $C = 0.251\,\mu\text{m}^{-1}$). Overall, both analysis methods provide consistent results and validate our understanding of the coupling dynamics between the two DLSPPWs. We forgo the extraction of the couling length and parameters for the other measurements, as this can and has been done in our group previously with leakage radiation microscopy (LRM). Having access to the near-field phase, we are especially interested in the coupling dynamics, which are not measureable with LRM.

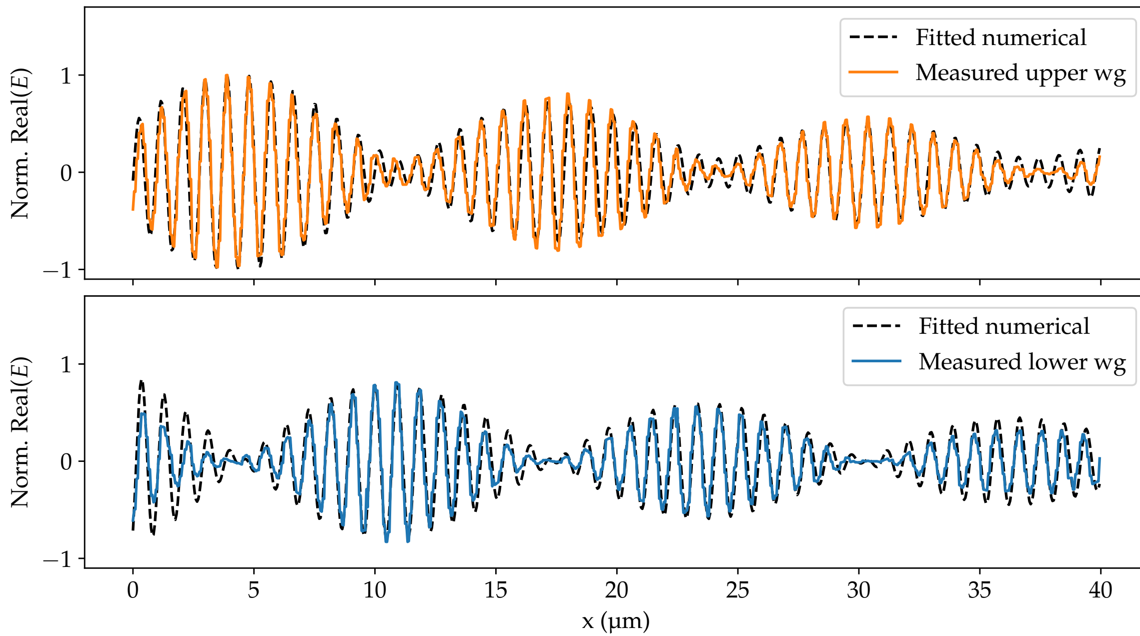


Figure 7.9: Comparison of the fitted numerical model to the real part of the two waveguides, recreated from the measured amplitude and phase data.

An interesting aspect of such a beating pattern is the relative phase difference between the two waveguides. From the measurements, we can directly calculate this along the propagation direction. We compare this to the phase difference calculated from the fitted numerical model in Fig. 7.10. In the numerical model the field in the second waveguide is also based on Eq. (7.8) as the first waveguide but the positional offset is shifted by half the beating period. The phase difference shows a periodic variation along the propagation direction, with a maximum phase difference of π. This behavior is expected, because one of the modes is asymmetric. The asymmetric mode has an opposite phase in the the two waveguides, and if both modes have equal strength this will lead to a total phase shift of

π. From the relative phase difference one could even estimate the relative strength of the two modes, e.g. if one mode is dominant it is less than π. In Fig. 7.10, one can actually see, that this phase shift decreases with the propagation distance because both modes have slightly different propagation lengths. This leads to a change in the relative strength of the two modes and thus to a weaker phase shift between the waveguides.

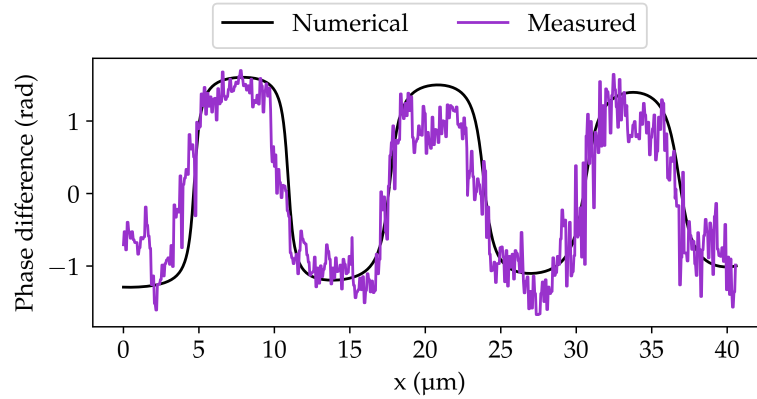


Figure 7.10: Phase difference extracted from the measurements shown in Fig. 7.9 and the numerical model for the averaged effective indices extracted from the fits displayed in Fig. 7.9.

The findings in this section demonstrate the coupling between two DLSPPWs using s-SNOM measurements. The periodic exchange of amplitude and the correlated phase variations along the propagation direction confirm the coupling dynamics predicted by the CMT.

7.2 Coupled DLSPP Waveguide Arrays

Coupled waveguide arrays are an interesting platform to simulate various quantum mechanical models [23–26]. One of the simplest models that can be simulated using coupled waveguide arrays is the SSH model. It exhibits topological properties, such as edge states and topological phase transitions [105, 133–136]. The SSH model is introduced theoretically and then demonstrated experimentally using a plasmonic implementation using an array of coupled DLSPPWs. One of the key features of the SSH model is the presence of topologically protected edge states, which are localized at the boundaries of the system and are robust against certain types of perturbations. We investigate these edge states using s-SNOM measurements and analyze the results.

7.2.1 Coupled waveguides as quantum simulators

Similar to two coupled waveguides, an array of coupled waveguides can be described using CMT. We consider only nearest neighbor coupling and identical waveguides ($C_{nm} = C_{mn} =$

C and $\beta_n = \beta$ for all $n, m \in \mathbb{N}$), meaning the coupling between two neigboring waveguides is symmetric and each waveguide has the same propagation constant. Analogous to the two waveguide case, the amplitude A_n of the mode in the n-th waveguide evolves along the propagation direction x according to [137]

$$\frac{dA_n(x)}{dx} = iC_{n-1}A_{n-1}(x) + i\beta A_n(x) + iC_{n+1}A_{n+1}(x)\,, \tag{7.9}$$

where C_{n-1} and C_{n+1} are the coupling coefficients between the n-th waveguide and its nearest neighbors.

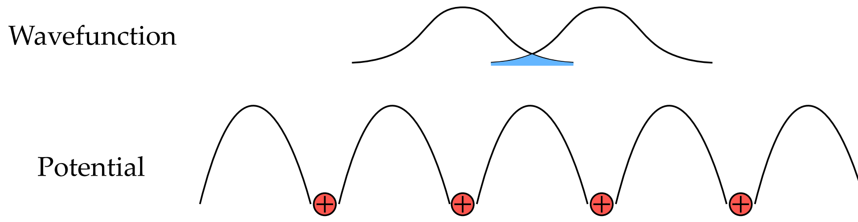


Figure 7.11: Sketch of the tight-binding model, including the electron wavefunction overlap of two neighboring sites and the periodic potenial of the one dimensional lattice.

According to the tight-binding model in solid state physics, the Hamiltonian of an electron hopping on a one-dimensional infinitly long lattice with identical and real hopping amplitudes can be written as [138]

$$\hat{H} = J \sum_{n=-\infty}^{\infty} \left(|n\rangle\langle n+1| + |n+1\rangle\langle n|\right) - V \sum_{n=-\infty}^{\infty} \left(|n\rangle\langle n|\right) \tag{7.10}$$

$$= J \sum_{n=-\infty}^{\infty} \left(|n\rangle\langle n+1| + h.c.\right) - V \sum_{n=-\infty}^{\infty} \left(|n\rangle\langle n|\right)\,, \tag{7.11}$$

where J is the hopping amplitude (overlap integral) between neighboring sites, V is an arbitrary on site potential and $|n\rangle$ is a Wannier state localized at site n. In this model only hopping (tunneling) between nearest neighbors (atoms) is considered. The particle state $|\psi\rangle$ can be expressed as a superposition of the Wannier states

$$|\psi\rangle = \sum_{n=-\infty}^{\infty} A_n |n\rangle\,, \tag{7.12}$$

where $A_n = \langle n|\psi\rangle$ are the probability amplitudes of finding the electron at site n. Inserting this expression into the time-dependent Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle\,, \tag{7.13}$$

allows us to derive the temporal evolution of the amplitudes A_n

$$
\begin{aligned}
\frac{d}{dt}A_n(t) &= \langle n| \frac{d}{dt} |\psi(t)\rangle &(7.14)\\
&= -i\hbar \langle n| \hat{H} |\psi(t)\rangle &(7.15)\\
&= -i\hbar \langle n| \left(J \sum_{m=-\infty}^{\infty} (|m\rangle\langle m+1| + |m+1\rangle\langle m|) - V \sum_{m=-\infty}^{\infty} (|m\rangle\langle m|) \right) |\psi(t)\rangle &(7.16)\\
&= -i\hbar \left(J\left(A_{n+1}(t) + A_{n-1}(t)\right) - VA_n(t)\right) . &(7.17)
\end{aligned}
$$

By identifying the time t with the propagation distance x and setting $\hbar = 1$, we see that the equations for the amplitudes A_n are mathematically equivalent to the coupled mode equations (Eq. (7.9)) for an array of coupled waveguides. This analogy allows us to simulate the behavior of electrons in a 1d-lattice using coupled waveguides, enabling the study of condensed matter systems, such as the SSH model [23–26, 41, 105].

7.2.2 Su-Schrieffer-Heeger Model

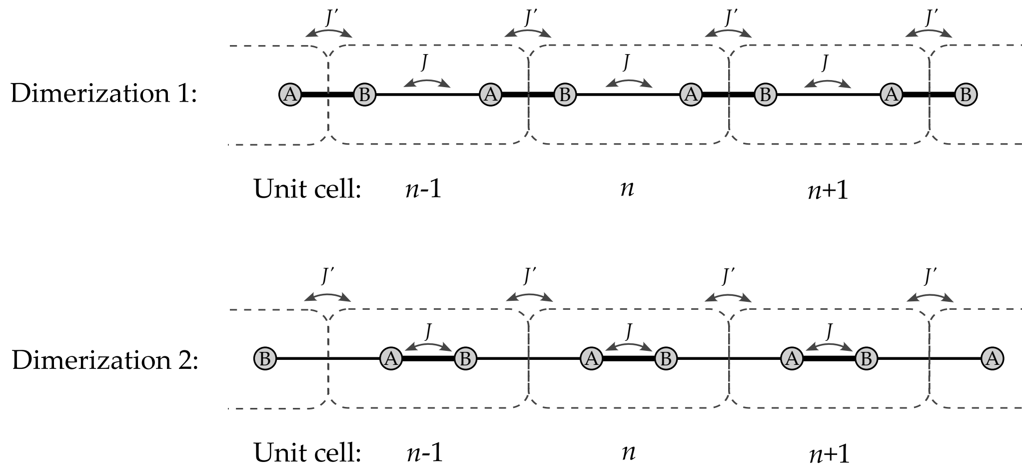


Figure 7.12: Sketch of the two dimerizations (choices of unit cell) of the one dimensional chain of atoms of the SSH model, with alternating intra cell (J) and inter cell (J') hopping amplitudes between neighboring sites. The unit cells are indicated by the dashed box and contains two atoms (sublattices) labeled A and B.

In the following, we will use this analogy to describe the SSH model using the tight-binding approach. The SSH model consists of a one-dimensional chain of atoms with alternating strong and weak hopping amplitudes between neighboring sites, as sketched in Fig. 7.12. Due to this alternating hopping structure, the chain can be divided into unit cells, each containing two atoms (sublattices) labeled A and B. The hopping amplitude between atoms within the same unit cell is denoted as J (intra cell hopping), while the hopping

amplitude between atoms in adjacent unit cells is denoted as J' (inter cell hopping). Due to the symmetry, two different unit cells can be chosen, such that the intra cell hopping is stronger than the inter cell hopping or vice versa. These two different unit cell definitions are called dimerizations, due to the dimer-like nature of the chain. The Hamiltonian of the SSH model can thus be expressed like before for the A and B sublattices at site n [133, 139]

$$\hat{H} = -J \sum_{n=1}^{N} \left(|n, B\rangle\langle n, A| + h.c. \right) - J' \sum_{n=1}^{N-1} \left(|n+1, A\rangle\langle n, B| + h.c. \right) , \tag{7.18}$$

where $h.c.$ denotes the Hermitian conjugate and $|n, A\rangle$ and $|n, B\rangle$ correspond to the state of the chain, where the electron is in the unit cell n on sublattice A or B respectively. Alternatively, the Hamiltonian can also be expressed in matrix form

$$\hat{H} = \begin{pmatrix} 0 & J & 0 & \cdots & 0 \\ J^* & 0 & J' & \cdots & 0 \\ 0 & J'^* & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & J' \\ 0 & 0 & 0 & J'^* & 0 \end{pmatrix} . \tag{7.19}$$

For an infinite chain, this Hamiltonian is translationally invariant. To exploit this symmetry, we apply periodic boundary conditions, effectively connecting the last unit cell back to the first one. This allows us to use Bloch's theorem to analyze the system in momentum space, by performing a Fourier transform [133]

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{-ikna} |n\rangle , \quad \text{where}\, k \in \{\delta_k, 2\delta_k, \cdots, N\delta_k\} , \quad \delta_k = \frac{2\pi}{N}. \tag{7.20}$$

The tight-binding Hamiltonian in momentum space then reads

$$\hat{H}(k) = \sum_k |k\rangle \, \hat{H}_k \, \langle k| , \tag{7.21}$$

with

$$\hat{H}_k = \begin{pmatrix} 0 & J + J'e^{-ika} \\ J + J'e^{ika} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \rho(k) \\ \rho^*(k) & 0 \end{pmatrix} . \tag{7.22}$$

This Hamiltonian exhibits a chiral symmetry and can also be written using the Pauli matrices $\vec{\sigma}$ as [139]

$$\hat{H}_k = \vec{d}(k) \cdot \vec{\sigma} , \tag{7.23}$$

with

$$\vec{d}(k) = \begin{pmatrix} J + J' \cos(ka) \\ J' \sin(ka) \\ 0 \end{pmatrix} . \tag{7.24}$$

Before topological properties are explained in detail, we first derive the energy bands of the bulk for a long but finite chain by diagonalizing the Hamiltonian, yielding the energy eigenvalues $E_\pm(k)$ and eigenstates $|\psi_\pm(k)\rangle$ [139]

$$E_\pm(k) = \pm|\vec{d}(k)| = \pm\sqrt{J^2 + J'^2 + 2JJ'\cos(ka)}\,, \tag{7.25}$$

$$|\psi_\pm(k)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} \pm e^{-i\phi(k)} \\ 1 \end{pmatrix}, \quad \phi(k) = \tan^{-1}\left(\frac{d_y}{d_x}\right). \tag{7.26}$$

The choice of the hopping coefficients J and J' can lead now to several special cases, which we will discuss now based on [133].

(i) $J' = 0$: In this case, the chain consists of isolated dimers with strong intra cell hopping and no inter cell hopping. The energy spectrum consists of two flat bands at $E = \pm J$, corresponding to bonding and antibonding states within each dimer. There are no edge states in this configuration.

(ii) $J = 0$: Here, the chain consists of isolated atoms with no intra cell hopping and strong inter cell hopping. The energy spectrum again consists of two flat bands at $E = \pm J'$. However, in this case, there are free edge states localized at the boundaries of the chain, with energies at $E = 0$. They are called free because they do not contribute to the energy of the system.

(iii) $J > J'$: In this regime, the intra cell hopping is stronger than the inter cell hopping. The energy spectrum consists of two bands separated by a band gap of size $\Delta E = 2(J - J')$. There are no edge states in this configuration, and the system is in a topologically trivial phase.

(iv) $J = J'$: At this critical point, the intra cell and inter cell hopping amplitudes are equal. The band gap closes, and the energy spectrum exhibits a Dirac point at $k = \pi/a$. The system undergoes a topological phase transition at this point.

(v) $J < J'$: In this regime, the inter cell hopping is stronger than the intra cell hopping. The energy spectrum again consists of two bands separated by a band gap of size $\Delta E = 2(J' - J)$. However, in this case, there are topologically protected edge states localized at the boundaries of the chain, with energies at $E = 0$. These free edge states are robust against certain types of perturbations and are a hallmark of the topologically non-trivial phase of the SSH model.

The energy eigenvalues for a variation of the hopping amplitudes J and J' are shown in Fig. 7.13 (a). One can see the closing of the band gap at the critical point $J = J'$ and the emergence of zero energy edge states for $J < J'$. In Fig. 7.13 (b) the corresponding eigenstates for three different modes are shown for the case $J' = 1.5$ and $J = 1$. The position of the modes in the energy spectrum are indicated in (a) by colored lines. In order to express the

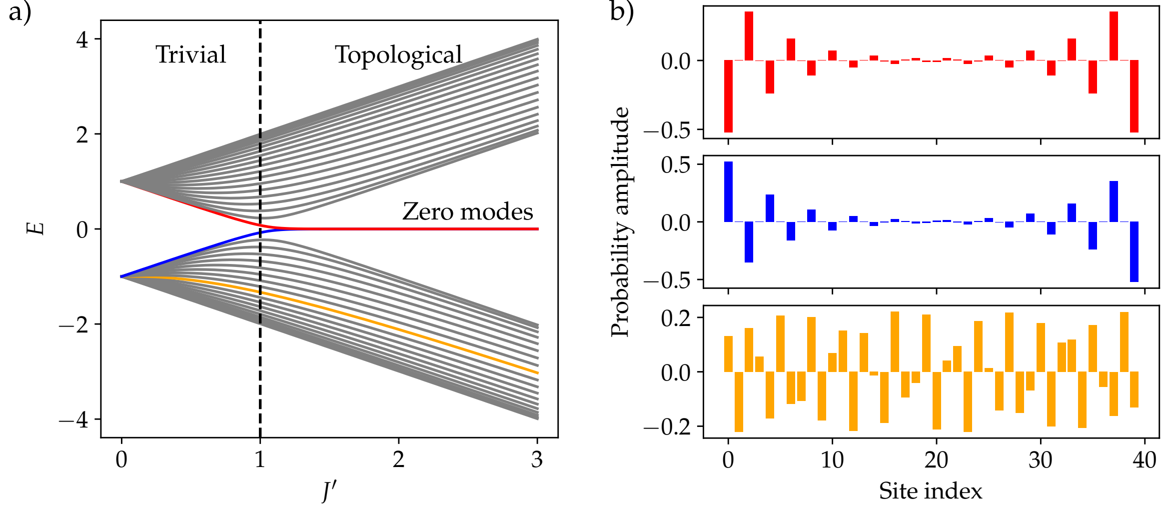


Figure 7.13: (a) Energy eigenvalues of a finite SSH bulk with 20 unit cells in dependence on the intercell hopping amplitude J' for a fixed intracell hopping amplitude of $J = 1$. The band gap closes at the critical point $J = J'$. For $J < J'$ zero energy edge states appear. The gray lines indicate the energy eigenvalues of the bulk states. (b) Corresponding eigenstates for three different modes for $J' = 1.5$ and $J = 1$.

topological properties of the SSH model more formally, a topological invariant, called the winding number, is introduced, which counts how many times the vector $\vec{d}(k)$ winds around the origin as k is varied over the Brillouin zone. The winding number is given by [133]

$$\nu = \frac{1}{2\pi} \int_{-\pi/a}^{\pi/a} \left(\hat{d}(k) \times \frac{d\hat{d}(k)}{dk} \right) dk\,, \quad \hat{d} = \frac{\vec{d}(k)}{|\vec{d}(k)|}\,. \tag{7.27}$$

In our case, the winding number can also be expressed in a more specific form using the $\rho(k)$ function defined above and the complex logarithm $log(|h|e^{i\,arg(h)}) = ln(|h|) + i\,arg(h)$ [133]

$$\nu = \frac{1}{2\pi i} \int_{-\pi/a}^{\pi/a} dk \frac{d}{dk} \log\left(\rho(k)\right) \tag{7.28}$$

$$= \frac{1}{2\pi} \left[arg(\rho(\pi/a)) - arg(\rho(-\pi/a)) \right]\,. \tag{7.29}$$

This will result in either $\nu = 0$ for the topologically trivial phase ($J > J'$) or $\nu = 1$ for the topologically non-trivial phase ($J < J'$). The winding of the vector $\vec{d}(k)$ is exemplified in Fig. 7.14 for the two different dimerization configurations and the transition point $J = J'$.

So far we only discussed the bulk properties of the SSH model and zero energy edge states located at the boundaries of the chain. Another practical way to create an edge state is to introduce a domain wall in the chain, where the dimerization configuration changes from the topological phase to the trivial phase or vice versa. This can be achieved by changing

Figure 7.14: Energy bands (left column) and winding of the vector $\vec{d}(k)$ (right column) for the two different dimerization configurations of the SSH model. (a, b) Topologically trivial phase with $J > J'$ and winding number $\nu = 0$. (c, d) Critical point with $J = J'$ where the band gap closes. (e, f) Topologically non-trivial phase with $J < J'$ and winding number $\nu = 1$.

the hopping amplitudes J and J' at a specific site in the chain. The presence of such a domain wall leads to the emergence of localized edge states at the interface between the two topological phases. These edge states are characterized by a zero energy and an exponential localization to the domain wall. The Hamiltonian matrix of such a system with a domain wall can be written as

$$\hat{H}_{\text{SSH}} = \begin{pmatrix} \ddots & & & & & & & \\ & 0 & J' & & & & & \\ & J' & 0 & J & & & \mathbf{0} & \\ & & J & 0 & J' & & & \\ & & & J' & 0 & J' & & \\ & & & & J' & 0 & J & \\ & & \mathbf{0} & & & J & 0 & J' \\ & & & & & & J' & 0 \\ & & & & & & & \ddots \end{pmatrix} . \tag{7.30}$$

Like before, we can diagonalize this Hamiltonian to obtain the energy eigenvalues and eigenstates of the system. The presence of the domain wall leads to the emergence of localized edge states at the interface between the two topological phases. This is illustrated in Fig. 7.15, where (a) shows the energy eigenvalues for a finite SSH model with a domain wall in the middle of the chain, and (b) displays the eigenstate of the zero energy edge state, which is exponentially localized at the domain wall. The existence of this edge state is a direct consequence of the topological properties of the SSH model and demonstrates the bulk-boundary correspondence principle, which states that the presence of edge states is determined by the topological invariants of the bulk system [133].

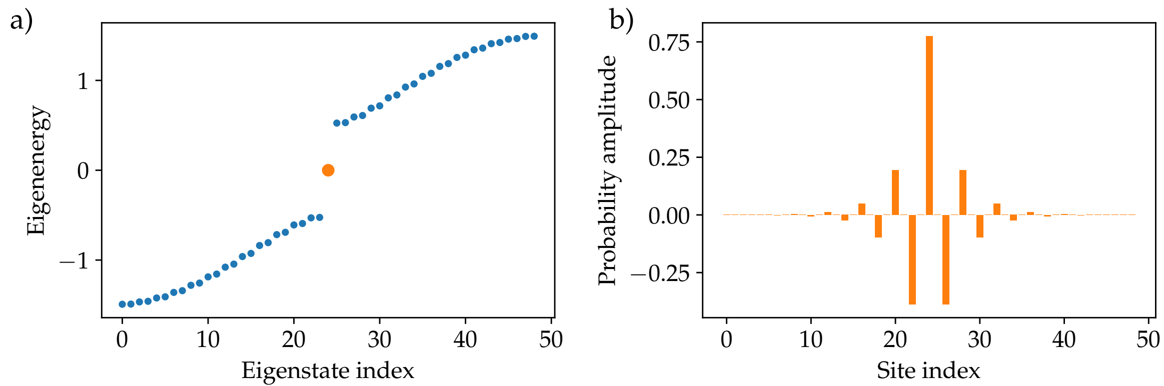


Figure 7.15: (a) Eigenvalues for a finite SSH model with a domain wall in the middle of the chain whith a length of 49 lattice sites. (b) Eigenstate of the zero energy edge state, indicated by the orange dot in (a), exponentially localized at the domain wall.

Furthermore, the chiral symmetry of the SSH Hamiltonian carries important implications for the nonzero energy eigenstates and zero energy eigenstates. The energy spectrum of the SSH Hamiltonian is symmetric, meaning for every state with energy E there exists a corresponding state with energy $-E$. Every nonzero energy eigenstate has an orthogonal counterpart and therefore equal support on both sublattices A and B. This symmetry can for example be seen in Fig. 7.13 (b, lowest panel). Zero energy eigenstates can be chosen to be supported by only one sublattice. They are then eigenstates of the chiral symmetry operator and thus chiral symmetric partners of themselves [133]. This symmetry leads to the distinct mode profile of the edge state shown in Fig. 7.15 (b), where the mode has nonzero amplitude only on every second site, corresponding to only one sublattice.

In order to visualize the dynamics of the SSH model with a domain wall, we perform numerical simulations of the amplitude evolution along the propagation direction. For this it is sufficent to calculate the overlap of the initial excitation with the different eigenmodes of the system and then evolve each mode according to its eigenenergy. The resulting evolution is shown in Fig. 7.16 for an array with a topological phase transition in the middle (left) and a simple bulk array without a domain wall (right). In the simulation, the waveguide at the domain wall (site 24) or in the bulk (site 24) is excited. The figure shows the localization of the edge state at the domain wall (left) and the ballistic spreading of the bulk modes. In

contrast, the right part of the figure shows only ballistic spreading of the excitation in the bulk array, as no edge state is present. This simulation illustrates the distinct dynamics of the SSH model with a domain wall compared to a simple bulk array.

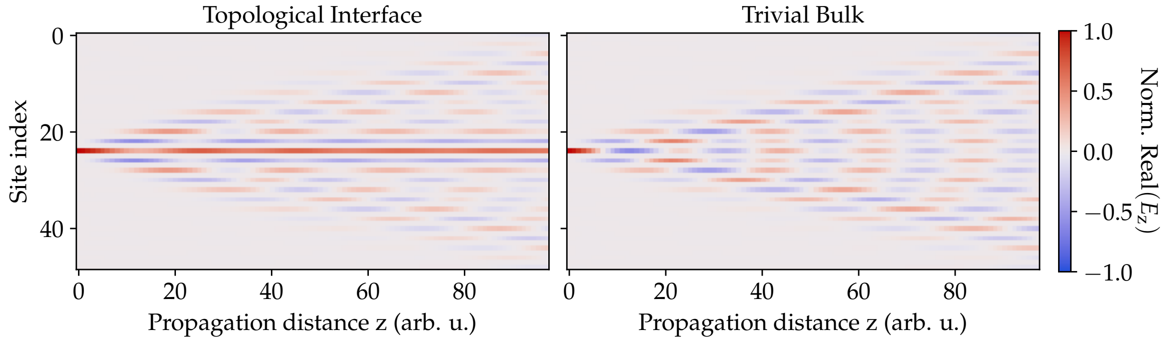


Figure 7.16: Numerical simulation of the spreading of an array with a topological phase transition in the middle (left panel) and a simple bulk array (right panel). Plotted is the normalized real part of the amplitude evolution along the propagation direction. This is not a full CMT calculation including losses and only shows the temporal evolution and interference of the different modes excited by the initial condition. This simulation is based on a script provided by S. Linden.

7.2.3 Plasmonic Implementation

In the previous section the theoretical background of the SSH model was discussed. In the following, an implementation of a topological phase transition in an array of coupled DLSPPWs is demonstrated. The hopping coefficients J and J' of the SSH model can be directly mapped to the coupling coefficients C_1 and C_2 between neighboring waveguides in the array. The coupling strength can be controlled by varying the center to center distance between adjacent waveguides (d_1 and d_2), which allows us to realize the two distinct dimerization configurations of the SSH model. A sketch of such a waveguide array is shown in Fig. 7.17. The left dimerization (purple) corresponds to the topologically trivial phase with strong intra cell coupling ($C_1 > C_2$), while the right dimerization (blue) represents the topologically non-trivial phase with strong inter cell coupling ($C_1 < C_2$). Like before, one can expect to observe an edge state localized at the interface between the two dimerization configurations at site 0, if the coupling coefficients are chosen accordingly. Such a plasmonic implementation of the SSH model was demonstrated by Bleckmann et al. [105], where the existence of the edge state was verified using far-field imaging with LRM. The mode distribution of the edge state is in good agreement with theoretical predictions. However, due to the far-field imaging technique, only the intensity distribution could be measured, while phase information was not accessible. In the following, we will use s-SNOM to measure both the amplitude and phase distribution of the electric field in a plasmonic SSH waveguide array, providing a more comprehensive characterization of the topological edge state.

Since the topological edge mode is only nonzero on the lattice sites 0 and even numbered sites, as indicated in Fig. 7.17, we can only couple to this mode by exciting one of these waveguides. Therefore, a grating coupler is added to the waveguide at site 0. However, multiple bulk mode also have nonzero amplitude at this site, so we expect to excite a superposition of the edge state and bulk states, as shown in Fig. 7.16. Due to the different eigenenergies, these modes will travel away from the edge, while the edge state remains localized at the interface, because it sits inside the band gap.

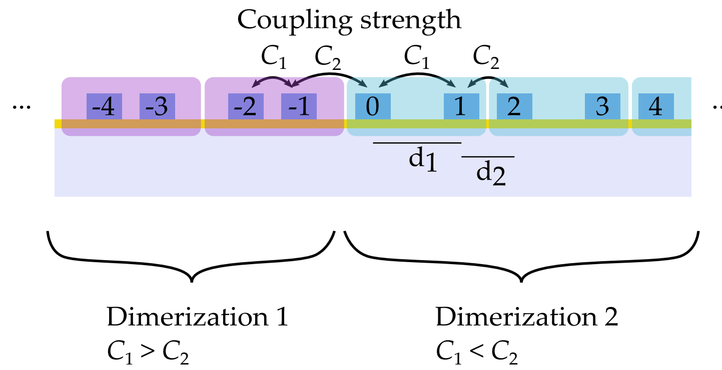


Figure 7.17: Schematic representation of a topological interface created by the two distinct dimerization configurations in the SSH model.

7.2.4 Fabrication

The fabrication of the SSH waveguide arrays is similar to the two coupled waveguides described above. The waveguides are again written from inside out in parallel with a linespacing of 5 nm. The waveguide at site 0 with the incoupler is extended by 5 µm. We will use similar parameters as Bleckmann et al. [105] to fabricate our waveguide array. Namely, a center to center distance of $d_2 = 600$ nm for the strong coupling (C_2) and $d_1 = 1\,000$ nm for the weak coupling (C_1). For the waveguide dimensions, a width of 300 nm and a height of 100 nm are chosen. For this a design width of 350 nm and a dose factor of 0.65 are used. The central waveguide is written with a slightly higher dose factor of 0.68 to account for the lower proximity effect due to the larger distance to the neighboring waveguides. The waveguides are written with a length of 170 µm to allow for sufficient propagation distance to observe the edge state. Surrounding the edge waveguide, 15 waveguide pairs are written on each side to limit the influence of the outer boundaries.

7.2.5 Measurements

As discussed before, we fabricated an array of coupled DLSPPWs implementing the SSH model with a defect at site 0. Similar to the single waveguide measurements, we operate

the s-SNOM in the transmission mode and lock the position of the lower parabola to the grating coupler at site 0. A large area scan of the waveguide array is shown in Fig. 7.18. The waveguide arrangement, sketched in Fig. 7.17, can be identified in the height map. The amplitude image shows ballistic spreading of the initial excitation, while the edge state remains exponentially localized at site 0, similar to the numerical simulation in Fig. 7.16. The phase image shows the excitation of free surface plasmon polaritons (SPPs) at the input grating coupler, and the propagation of the individual waveguide modes. Further analysis of the phase data is required to extract meaningful information like the sign of the electric field.

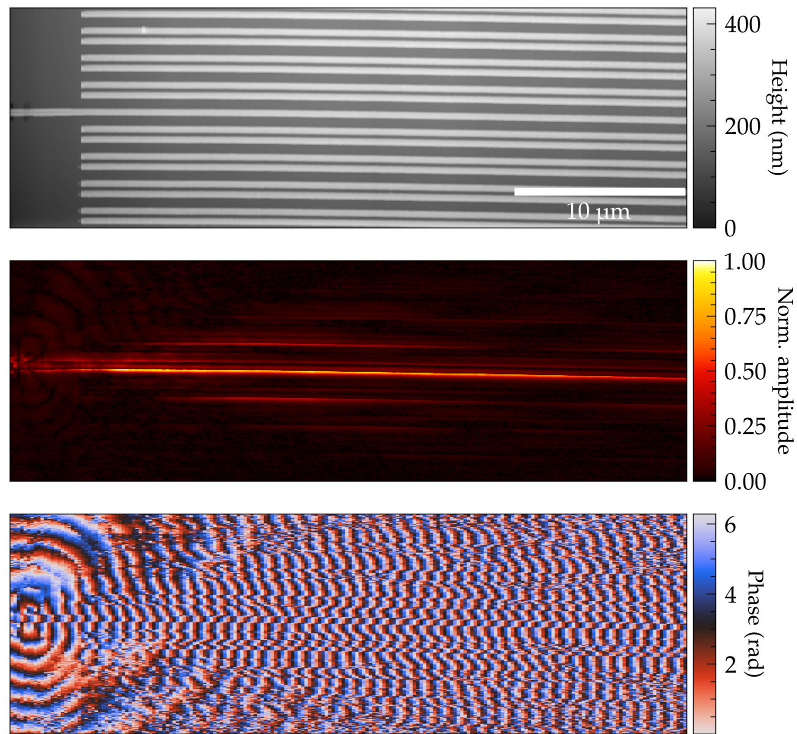


Figure 7.18: s-SNOM measurement at 970 nm of a coupled waveguide array in the SSH configuration with a center to center distance of 600 nm for the strong coupling and 1 000 nm for the weak coupling. The waveguides have a width of 380 nm and a height of 170 nm. This measurement is an example large area scan but not the same sample as used for the further analysis. Shown is the second demodulation order with only the synccorrection applied to the phase data.

Such a large scan gives a good overview of the evolution of the inital excitation, but it limits the resolution and quality of the data due to the long measurement time. Therefore, shorter cut-line scans perpendicular to the waveguides, with higher resolution and better signal to noise ratio are performed for further analysis. Since the temporal evolution along the propagation direction is slowly varying and not at the focus of this work, these individual cut-line scans are spaced 10 µm apart along the propagation direction. Each cut-line has a

width of 1 µm to allow for line-averaging and a varying length to include all waveguides with sufficient signal-to-noise ratio (SNR). An example cut-line measurement performed 60 µm away from the incoupler on a SSH waveguide array with a topological phase transition in the middle is shown in Fig. 7.19.

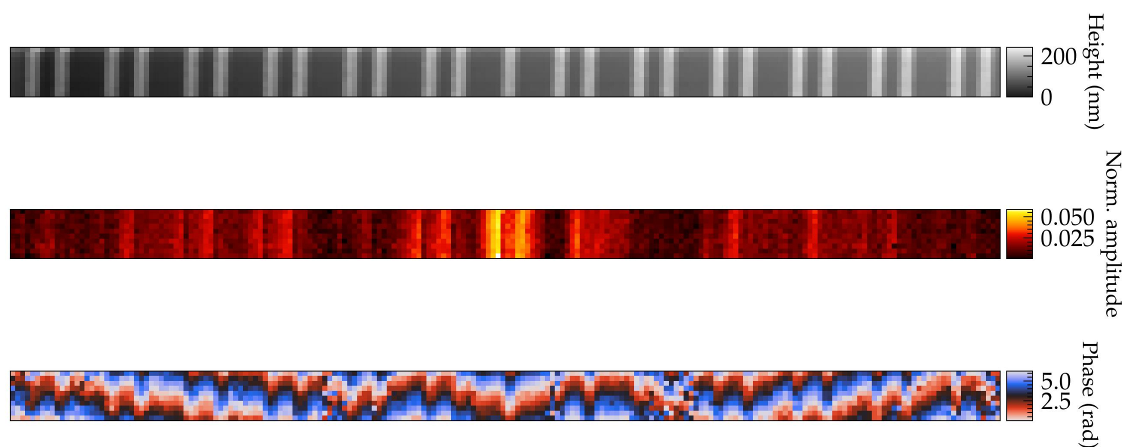


Figure 7.19: Example of a cut-line measurement at a distance of 60 µm from the incoupler. Shown is the height data and the amplitude and phase demodulated at the second order, with only the synccorrection applied to the phase data.

The raw amplitude an phase data in Fig. 7.19 are noisy and cluttered. Especially the phase data is hard to interprete, as each waveguide exhibits an additional phase shift in the middle compared to the field on the gold film next to it. In order to improve the SNR of each cutline measurement, the following modifications are applied. First, the forward and backward scan are overlaid, effectively doubling the integration time. This works well for the height and amplitude data, but not for the phase data. Since the two directions are often shifted by some tens of nanometers the two scans have to be aligned. For a higher alignment precision, the data is scaled up by a factor of 4 and smoothed with a Gaussian filter with a sigma of 2 pixels. The Gaussian filter has only a minor influence on the information in the scans and it further improves the SNR. After the filtering, a line-averaging is performed on the amplitude data. Since the width of the scan is with 1 µm much shorter than variations along the propagation direction, this averaging does not remove any relevant information.

The resulting cut-line data sets for the measurements at different distances to the incoupler are shown in Fig. 7.20. In (a) the configuration with a domain wall between the two dimerization configurations is shown, while in (b) the trivial bulk configuration without a domain wall is displayed. The individual cut-lines are shifted by an incremental offset of 0.1 for better visibility. In both cases the cut-lines are centered around the excited waveguide at $y = 0$ µm, which is the edge waveguide for the topological configuration (a). The positions

of the waveguides are extracted from the overlaid and averaged height data by their mean value along the cut-line. These positions are indicated by the gray bars in the amplitude plots. The detected amplitudes are higher next to the waveguides compared to on top of them similar to the single waveguide measurements in Chapter 6. This makes it harder to directly read out the amplitudes per waveguide. The temporal evolution along the propagation direction is very different for the two configurations. In the nontrivial configuration (a) a large portion of the amplitude remains localized at the excited waveguide and uneven neighbors, while in the trivial bulk configuration (b) the amplitude spreads whithout any localization. This is a first indication of the presence of the topological edge state in the domain wall configuration. In the following, a more quantitative analysis is performed. The presence of the large amount of bulk modes stems from the fact that an excitation of just the waveguide at site 0 allows for coupling to all modes with nonzero amplitude at this site, including multiple bulk modes. This corresponds to the numerical simulation shown in Fig. 7.16 (left panel).

To quantitatively compare the measured mode distribution of the edge state with theoretical predictions, the amplitude at each waveguide position is extracted from the line-averaged cut-line data. This is done by fitting a double Gaussian function to each waveguide in the amplitude data. The two Gaussians account for the increased amplitude next to the waveguide compared to on top of it, as already introduced in Section 7.1.3. With this method we get an amlitude value per waveguide for each cut-line measurement. In order to calculate the real part, a phase difference per waveguide relative to the central waveguide is extracted. This is done by comparing phase profiles taken directly on top of the waveguides. The real part of the electric field per waveguide can then be reconstructed by multiplying the fitted amplitude with the cosine of the extracted phase difference. This operation is only performed for the last two cut-line measurements in a distance of 60 µm and 70 µm to the incoupler, where the bulk modes have already spread sufficently far away from the edge state. In order to get a more robust result, the reconstructed real part of the electric field is averaged from these two cut-lines. The bulk modes, which are still present in the measurements, interfere and lead to a variation along the propagation direction, while the edge mode stays constant [2]. Therefore, averaging leads to a reduction of the bulkmodes, while the edge mode is not affected. The resulting mode distribution is shown in Fig. 7.21 in comparison with the theoretically expected edge mode distribution shown earlier in Fig. 7.15 (b). Up to the sixth neighboring waveguide, the overlap of the experimental data with the theoretical prediction is very good. Beyond that, the topological edge mode amplitude is very small and the bulk modes start to dominate the measured signal again. This confirms the existence of the topological edge state in our plasmonic SSH waveguide array and demonstrates the capability of s-SNOM to measure both amplitude and phase of such modes. The characteristic oscillation of the sign of the electric field amplitude can not not be observed in LRM measurements alone.

[2] The variations can be seen in Fig. 7.16 close to the central waveguide.

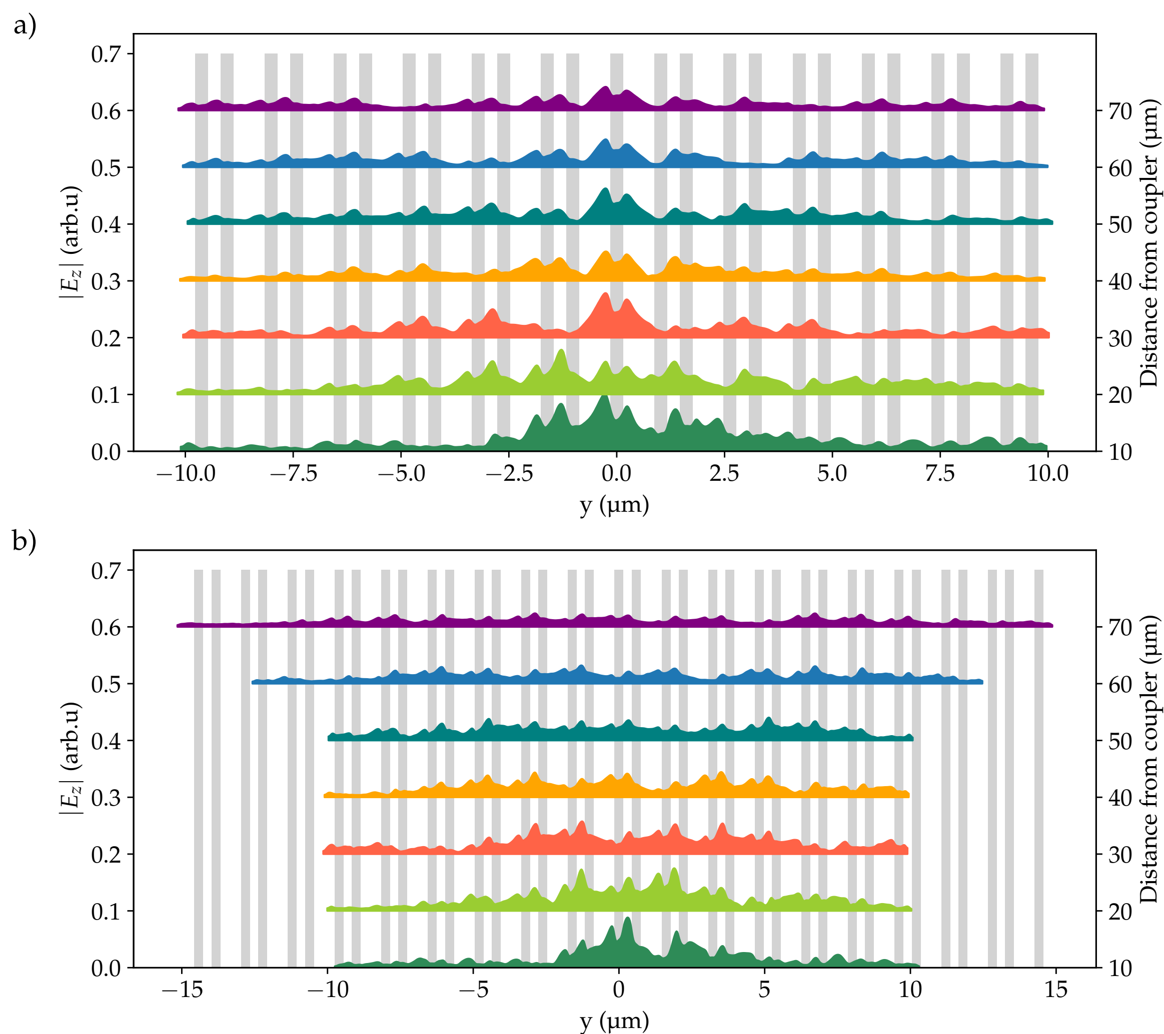


Figure 7.20: s-SNOM amplitude cut-line measurements of a coupled waveguide array. In (a) the configuration with a domain wall between the two dimerization configurations is shown, while in (b) the trivial bulk configuration without a domain wall is displayed. The individual cut-lines are spaced 10 µm apart along the propagation direction (x). The waveguides are indicated by the gray bars and all cut-lines are shifted relative to the excitation waveguide, which is the edge waveguide for the topological configuration (a).

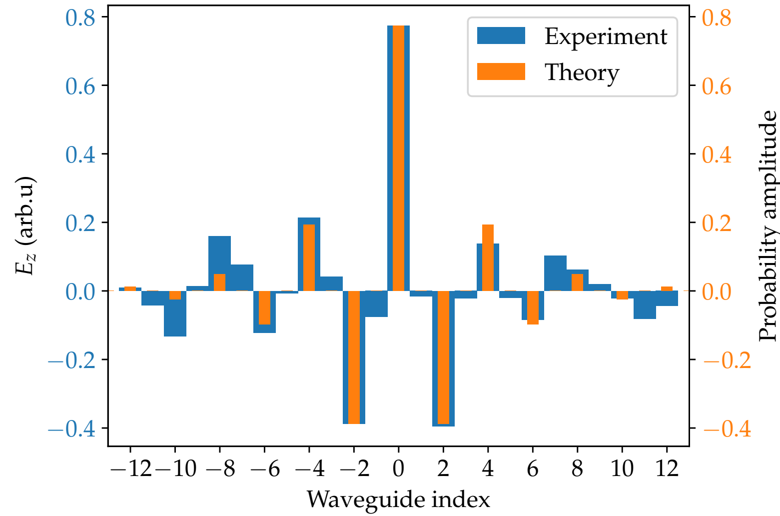


Figure 7.21: Comparison of the expected edge mode distribution with the experimental results. The experimental data is normalized to the maximum amplitude of the theoretical prediction. The experimental data is the average of the reconstructed real part of the electric field from the cut-line measurements performed in a distance of 60 µm and 70 µm from the incoupler.

7.3 Conclusion

In this chapter, s-SNOM measurements of coupled DLSPPWs were presented. First, measurements of two coupled waveguides with varying separation distance were shown. The characteristic periodic amplitude and phase modulation due to the beating of the symmetric and antisymmetric eigenmodes was observed. From these measurements, the coupling length was extracted and found to be in good agreement with *Comsol* simulations.

In the second part, the theoretical background of coupled waveguide arrays and their mathematical similarity to the one dimensional tight-binding model for electrons was discussed. As an example, the SSH model exhibiting a topological phase transition was introduced. A plasmonic implementation of the SSH model using coupled DLSPPWs was presented, along with s-SNOM measurements of the mode distribution of the topological edge state. The experimental results were compared with theoretical predictions, showing good agreement.

CHAPTER 8

Conclusion

In this thesis, various aspects of plasmonic waveguides have been explored, with a focus on their fabrication, characterization using s-SNOM and the investigation of specific waveguide systems such as chiral couplers and coupled DLSPPWs. The fabrication of nanophotonic structures and techniques to create high-quality plasmonic waveguides were discussed and refined. This included EBL, FIB milling and metal deposition. Additionally, large single-crystalline silver plates for the coupling of exciton-polaritons to surface plasmon polaritons were chemically grown, which are used in the studies of Yuhao Zhang. The various sample fabrication processes were crucial for the fabrication plasmonic waveguides, and took up a significant portion of the work presented in this thesis. A new two-step positive tone EBL process was developed for dielectric-loaded surface plasmon polariton (DLSPP) waveguides, improving shape and reproducibility. A python-based script was created in this context to design nanostructures, which is now widely used in the research group.

The s-SNOM technique was introduced, detailing its working principles, background suppression methods, amplitude and phase retrieval, and the experimental setup. The tip-sample interactions that are important for understanding material contrast were discussed, as well as experimental considerations such as sample orientation and input polarization to achieve different excitation and collection schemes. Data corrections were also addressed. A custom python toolbox was developed for data analysis, which is publicly available on *GitHub* and documented on *ReadTheDocs*. Additionally, a python interface was created to extract and convert simulation data from *COMSOL* to the format used in the s-SNOM toolbox, allowing for consistent analysis of both experimental and simulation data.

The core of this thesis presented three main experimental studies using plasmonic waveguides. The first study focused on the photonic spin Hall effect using strip metal waveguides with chiral couplers. A strong beating pattern was observed in both measurements and simulations, which could be explained by the interference of the waveguide mode with dipolar background excitation from the broad excitation laser. The observed directionality did not follow the simple expected resonance condition, instead exhibiting broader and more robust behavior with only weak influence from radius, width and bending.

The second study provided a detailed investigation of DLSPP waveguides, with a particular focus on the in-plane electric field component. Mode profiles, mode index, and propagation lengths were extracted from measurements and simulations. The in-plane component was discussed, and its ratio was calculated from simulations. An unexpectedly low amplitude on the waveguide was observed in measurements, which could be explained by the polarization sensitivity of the probing tip. Additionally, a strong phase shift was observed on the waveguide, though no exact explanation was found, possible causes were discussed. To directly measure the in-plane component, a reflection grating was implemented to spatially separate the two components. Polarizability-adapted simulations matched the measurements well and confirmed this approach.

The third study examined coupled DLSPP waveguides, demonstrating the periodic energy transfer between two waveguides and the influence of the gap size on the coupling strength. A detailed description of the coupling mechanism via beating of two modes was provided. Coupled waveguide arrays were explored as quantum simulators, drawing an analogy to the tight-binding Hamiltonian of electrons in a one-dimensional lattice. The SSH model was implemented using a coupled DLSPPW array, and topological edge states were directly observed using s-SNOM, along with the demonstration of the expected phase relation.

In conclusion, this thesis has advanced our understanding of plasmonic waveguides through the development of fabrication techniques, the application of s-SNOM for characterization, and the investigation of specific waveguide systems.

Outlook

In the future, several promising directions emerge from the work presented in this thesis. In the area of chiral couplers, exploring the use of orbital angular momentum instead of spin angular momentum could be an interesting avenue. Initial tests with vortex beams have been conducted, but conclusive results have yet to be achieved. This approach could offer new possibilities for controlling light in plasmonic circuits, although challenges related to the momentum matching and the polarization state when using high numerical aperture objectives or mirrors need to be addressed.

For DLSPP waveguides, further studies on the in-plane electric field component are instrumental. Estimating the polarization sensitivity of the probing tip could enhance our understanding of the measurements and the tip-sample interaction. DLSPP waveguides provide a favorable platform due to their relatively high in-plane field component and reduced suppression compared to metal films. The use of dielectric tips could be explored to achieve even lower suppression, albeit with a trade-off in signal strength.

In the field of coupled DLSPP waveguides, future research could focus on exploring different array configurations, such as implementing various topological models, investigating disorder effects, and studying non-Hermitian effects by introducing gain or loss to specific waveguides. One could also implement a spatial light modulator to selectively excite certain modes by tailoring the input amplitude and phase profiles of waveguide arrays.

Additionally, waveguide splitters and interferometers could be examined with s-SNOM to further the development of plasmonic circuitry.

Overall, the work presented in this thesis lays a solid foundation for continued exploration and innovation in the field of plasmonic waveguides.

APPENDIX A

Additional Information

A.1 s-SNOM Calculations

A.1.1 Lock-in Detection

The demodulated detector signal at the n-th harmonic can be approximated by considering only components involving the 0-order background term, as this is usually the dominant term:

$$U_{D,n} \approx k_D \left(\sigma_{NF,n}\sigma^*_{BG,0} + \sigma_{BG,0}\sigma^*_{NF,n} + \sigma_{BG,n}\sigma^*_{BG,0} + \sigma_{BG,0}\sigma^*_{BG,n}\right) . \tag{A.1}$$

Similar to before, we can express the complex scattering coefficients as:

$$\sigma_{NF,n} = s_{NF,n}e^{i\phi_{NF,n}} , \tag{A.2}$$

$$\sigma_{BG,n} = s_{BG,n}e^{i\phi_{BG,n}} . \tag{A.3}$$

Therefore, the demodulated detector signal can be expressed as:

$$U_{D,n} \approx k_D b_0 \Big(s_{NF,n}e^{i\left(\phi_{NF,n}-\phi_{BG,0}\right)} + s_{NF,n}e^{-i\left(\phi_{NF,n}-\phi_{BG,0}\right)} \tag{A.4}$$

$$+s_{BG,n}e^{i\left(\phi_{BG,n}-\phi_{BG,0}\right)} + s_{BG,n}e^{-i\left(\phi_{BG,n}-\phi_{BG,0}\right)}\Big) , \tag{A.5}$$

which can be simplified to:

$$U_{D,n} \approx 2k_D b_0 \left(s_{NF,n}\cos\left(\phi_{NF,n} - \phi_{BG,0}\right) + s_{BG,n}\cos\left(\phi_{BG,n} - \phi_{BG,0}\right)\right) . \tag{A.6}$$

For a sufficiently high harmonic n, the background term $s_{BG,n}$ is typically much smaller than the near-field term $s_{NF,n}$ and we get:

$$U_{D,n} \approx 2k_D s_{BG,0} s_{NF,n}\cos\left(\phi_{NF,n} - \phi_{BG,0}\right) . \tag{A.7}$$

A.1.2 Polarization Projection in s-SNOM Setup

In the follwing we calculate the projection of the polarization vector of the light scattered by the tip after reflection on the upper parabolic mirror (M_1) of the scanning near-field optical microscopy (SNOM) setup. The situation is sketched in Fig. A.1. The tip is located at T and is polarized with a polarization vector $\boldsymbol{p}_0$. The light is scattered along vector $\boldsymbol{a}$ towards the upper parabolic mirror M_1, where it is reflected towards the second mirror M_2 along vector $\boldsymbol{b}$. The reflection from mirror M_2 can be neglected, as the light stays in the same plane. The angles $\alpha = 45°$ and $\beta = 30°$ define the emission direction of the tip. In order to calculate the polarization state at the position of M_2, we start by projecting the initial polarization vector $\boldsymbol{p}_0$ onto the plane perpendicular to the emission direction $\boldsymbol{a}$. The projected polarization vector $\boldsymbol{p}'$ is then rotated around the normal vector of the area spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ by an angle of $-\theta$, where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$. This rotation accounts for the reflection on mirror M_1. Uppon reflection on the mirror, the Fresnel coefficients for p- and s-polarization should be considered, but this is neglected here for simplicity.

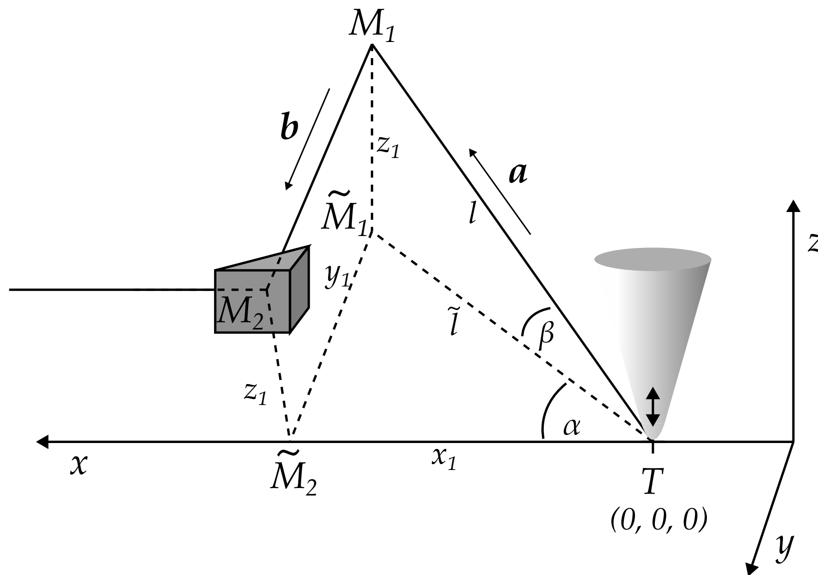


Figure A.1: Sketch of the tip emission and reflection on the upper parabola M_1. The coordinate system is defined from a point of view of someone standing in front of the SNOM

We start the derivation by writing out the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in the x-y-z-coordinate

system

$$\boldsymbol{a} = \overline{TM_1} = \begin{pmatrix} \tilde{l}\cos(\alpha) \\ -\tilde{l}\sin(\alpha) \\ l\sin(\beta) \end{pmatrix} = \begin{pmatrix} \cos(\beta)\cos(\alpha) \\ -\cos(\beta)\sin(\alpha) \\ \sin(\beta) \end{pmatrix}, \tag{A.8}$$

$$\boldsymbol{b} = \overline{M_1 M_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{A.9}$$

The next step is to calculate the polarization unit vectors in the x'-y'-z' coordinate system, denoted as $\hat{\boldsymbol{x}}'$ and $\hat{\boldsymbol{z}}'$ as indicated in Fig. A.2

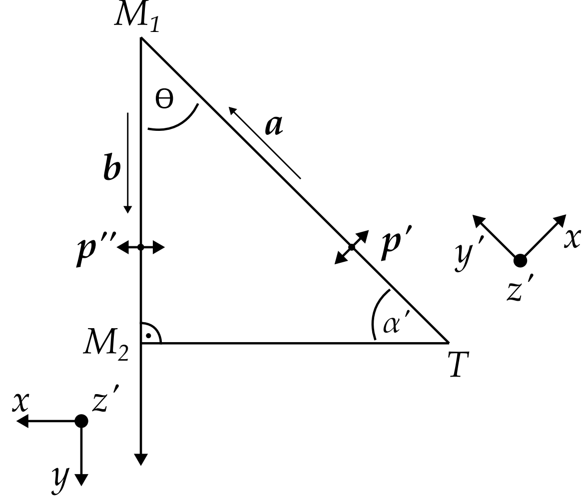


Figure A.2: Projection of the polarization vector onto the x'-y'-z'-coordinate system.

$$\hat{\boldsymbol{x}}' = \frac{\boldsymbol{a}\times\boldsymbol{b}}{|\boldsymbol{a}\times\boldsymbol{b}|} = \frac{1}{l\sqrt{\sin^2(\beta)+\cos^2(\beta)\cos^2(\alpha)}} \begin{pmatrix} -l\sin(\beta) \\ 0 \\ l\cos(\beta)\cos(\alpha) \end{pmatrix}, \tag{A.10}$$

$$\hat{\boldsymbol{z}}' = \frac{\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{a})}{|\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{a})|} = \frac{-1}{\sqrt{\sin^2(\beta)+\cos^2(\beta)\cos^2(\alpha)}} \begin{pmatrix} l^2\cos^2(\beta)\sin(\alpha)\cos(\alpha) \\ l^2\sin^2(\beta)+l^2\cos^2(\beta)\cos^2(\alpha) \\ l\cos(\beta)\,l\sin(\alpha)\sin(\beta) \end{pmatrix}. \tag{A.11}$$

The projection of the initial polarization vector $\boldsymbol{p}_0$ on the x'-z'-plane is then given by

$$\boldsymbol{p}' = (\boldsymbol{p}_0\cdot\hat{\boldsymbol{x}}')\,\hat{\boldsymbol{x}}' + (\boldsymbol{p}_0\cdot\hat{\boldsymbol{z}}')\,\hat{\boldsymbol{z}}'. \tag{A.12}$$

(A.13)

The next step is to rotate the polarization vector $\boldsymbol{p}'$ around the axis $\hat{\boldsymbol{z}}'$ by an angle of $-\theta$. The angle θ can be calculated using the dot product of the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$

$$\cos(\theta) = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}|\,|\boldsymbol{b}|}, \tag{A.14}$$

$$= \frac{\left(\tilde{l}\cos(\alpha)\right)(0) + \left(\tilde{l}\sin(\alpha)\right)(1) + \left(l\sin(\beta)\right)(0)}{\sqrt{\left(-\tilde{l}\cos(\alpha)\right)^2 + \left(\tilde{l}\sin(\alpha)\right)^2 + \left(-l\sin(\beta)\right)^2}}, \tag{A.15}$$

$$= \frac{\tilde{l}\sin(\alpha)}{\sqrt{\tilde{l}^2 + l^2\sin(\beta)}}, \tag{A.16}$$

$$\theta = \arccos\left(\frac{\tilde{l}\sin(\alpha)}{\sqrt{\tilde{l}^2 + l^2\sin(\beta)}}\right). \tag{A.17}$$

We can use the relation $\tilde{l} = l\cos(\beta)$, which follows from Fig. A.3, to simplify the expression

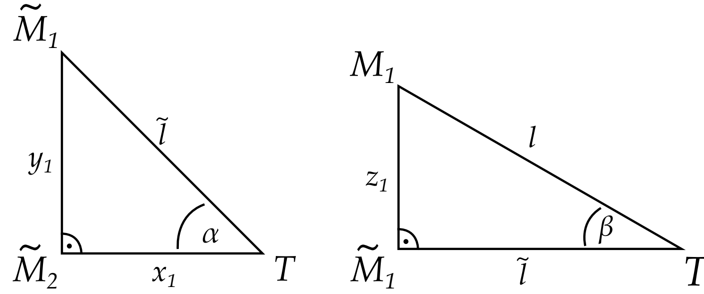


Figure A.3: Sketch of the relations of the angles and pathways.

$$\theta = \arccos\left(\frac{l\cos(\beta)\sin(\alpha)}{\sqrt{l^2\cos^2(\beta) + l^2\sin(\beta)}}\right), \tag{A.18}$$

$$= \arccos\left(\frac{l\cos(\beta)\sin(\alpha)}{l}\right), \tag{A.19}$$

$$= \arccos\left(\cos(\beta)\sin(\alpha)\right). \tag{A.20}$$

Using the angles from the s-SNOM setup we get

$$\boxed{\theta \approx 52.24^\circ}. \tag{A.21}$$

The rotation around an arbitrary axis $\hat{\boldsymbol{k}}$ by an angle γ can be expressed using Rodrigues' rotation formula

$$\mathbf{R} = \mathbf{I} + \sin(\gamma)\,\mathbf{K} + (1 - \cos(\gamma))\,\mathbf{K}^2, \tag{A.22}$$

where **I** is the identity matrix, **K** is the cross-product matrix of the rotation axis and γ is the rotation angle. The cross-product matrix of the rotation axis is given by

$$\mathbf{K} = \begin{pmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{pmatrix}, \tag{A.23}$$

where $\hat{\boldsymbol{k}} = (k_1, k_2, k_3)$ is the unit vector pointing along the rotation axis.

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + s \begin{pmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{pmatrix} + (1-c) \begin{pmatrix} -k_2^2 - k_3^2 & k_1 k_2 & k_1 k_3 \\ k_1 k_2 & -k_1^2 - k_3^2 & k_2 k_3 \\ k_1 k_3 & k_2 k_3 & -k_1^2 - k_2^2 \end{pmatrix}, \tag{A.24}$$

$$= \begin{pmatrix} 1 + (1-c)\left(-k_2^2 - k_3^2\right) & -s \cdot k_3 + (1-c)\, k_1 k_2 & s \cdot k_2 + (1-c)\, k_1 k_3 \\ s \cdot k_3 + (1-c)\, k_1 k_2 & 1 + (1-c)\left(-k_1^2 - k_3^2\right) & -s \cdot k_1 + (1-c)\, k_2 k_3 \\ -s \cdot k_2 + (1-c)\, k_1 k_3 & s \cdot k_1 + (1-c)\, k_2 k_3 & 1 + (1-c)\left(-k_1^2 - k_2^2\right) \end{pmatrix}, \tag{A.25}$$

where $c = \cos(\gamma)$ and $s = \sin(\gamma)$. Using the relation $k_1^2 + k_2^2 + k_3^2 = 1$ we can simplify the expression

$$\mathbf{R} = \begin{pmatrix} k_1^2 (1-c) + c & k_1 k_2 (1-c) - k_3 s & k_1 k_3 (1-c) + k_2 s \\ k_1 k_2 (1-c) + k_3 s & k_2^2 (1-c) + c & k_2 k_3 (1-c) - k_1 s \\ k_1 k_3 (1-c) - k_2 s & k_2 k_3 (1-c) + k_1 s & k_3^2 (1-c) + c \end{pmatrix}. \tag{A.26}$$

With this we can calculate the rotation matrix for the rotation around the z'-axis by an angle of $\gamma = -\theta$

$$\boldsymbol{p}'' = \mathbf{R} \boldsymbol{p}'. \tag{A.27}$$

If we assume for example an initial polarization vector along the z-axis, we find that the final polarization vector after reflection on mirror M_1 is rotated by 12.7° relative to the z-axis.

A.2 DLSPP Waveguides

A.2.1 Single DLSPP Waveguide Mode Simulations

In the following we show a some additional mode simulations of single DLSPPWs and discuss the estimation of the mode confinement.

Mode Profiles

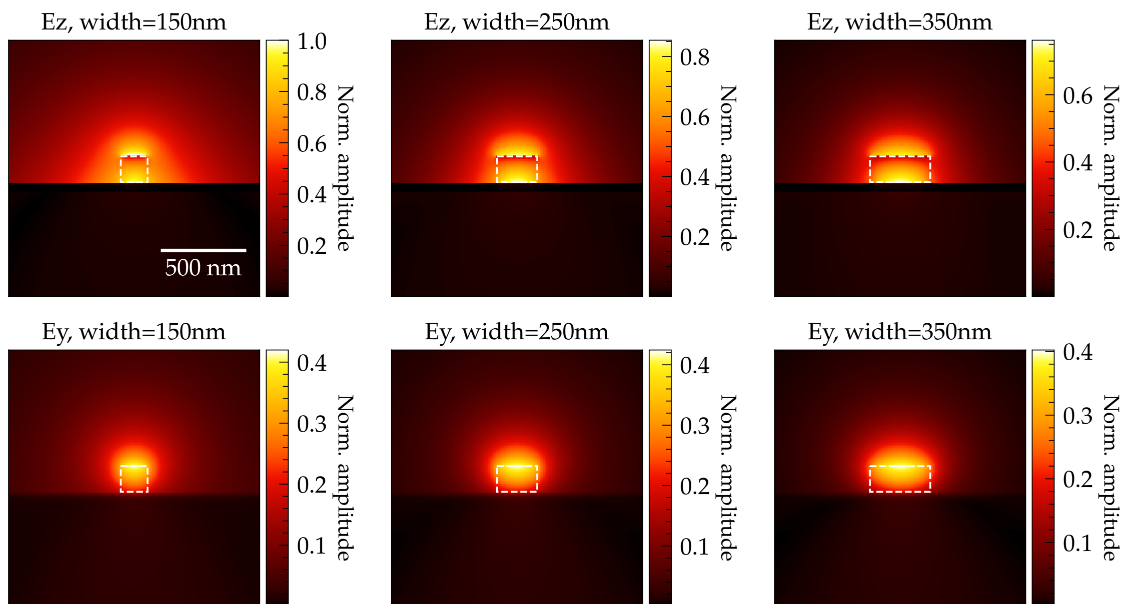


Figure A.4: Mode profile of a DLSPPW with a width between 150 nm and 350 nm and a height of 150 nm at a wavelength of 970 nm. The absolute value of the z-component of the electric field is shown in the upper half of the image, while the lower half shows the x-component (longitudinal).

In Figs. A.4 and A.5 two examples of mode profiles for different wavelengths and waveguide widths are shown. The first example in Fig. A.4 shows a wavelength of 970 nm and waveguide widths between 150 nm and 350 nm. The second example in Fig. A.5 shows a wavelength of 1 500 nm and waveguide widths between 300 nm and 500 nm. The height is kept constant at 150 nm. In both cases the out-of-plane component (upper half labeled 'Ez') is dominant, but the in-plane component (lower half labeled 'Ey') can still be quite significant.

Confinement Factors

The mode confinement can be defined in various ways, here we simply define it as the ratio of the illumination wavelength to the decay length of the mode in the two transversal

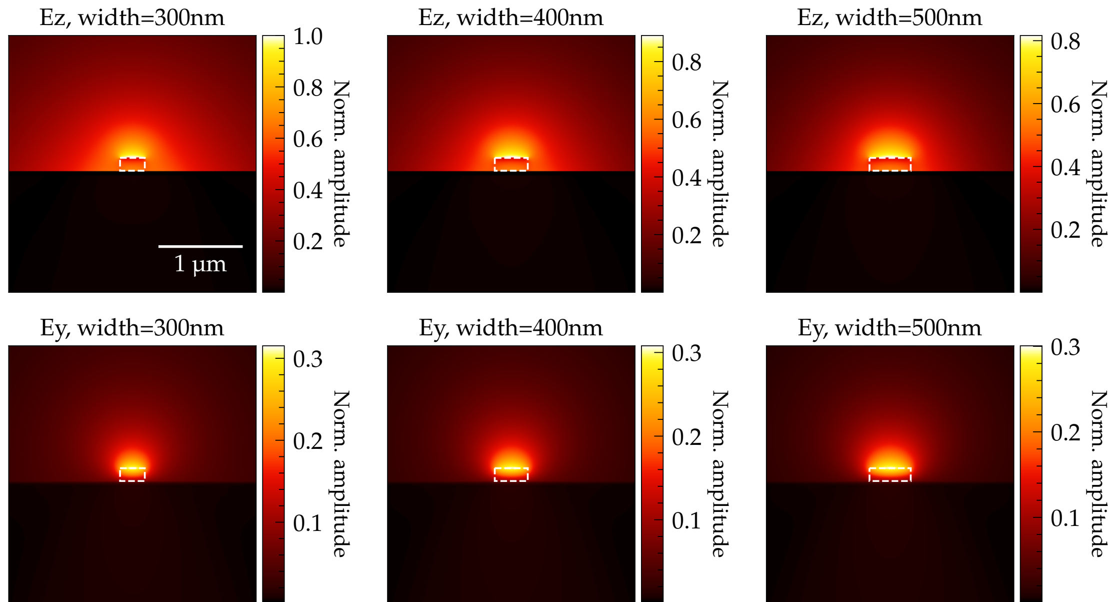


Figure A.5: Mode profile of a DLSPPW with a width between 300 nm and 500 nm and a height of 150 nm at a wavelength of 1 500 nm. The absolute value of the z-component of the electric field is shown in the upper half of the image, while the lower half shows the x-component (longitudinal).

directions. The decay length is calculated by fitting an exponential function to a cutline across the waveguide. The cutlines are parallel and vertical to the metal surface through the center of the waveguide. In Fig. A.6 and Fig. A.7 we show the horizontal and vertical confinement factors for different waveguide widths and heights at the two wavelengths of interest. This shows, that the confinement increases with increasing waveguide dimensions, as expected. The confinement is generally higher for the shorter wavelength of 970 nm, as the mode is more tightly bound to the waveguide structure.

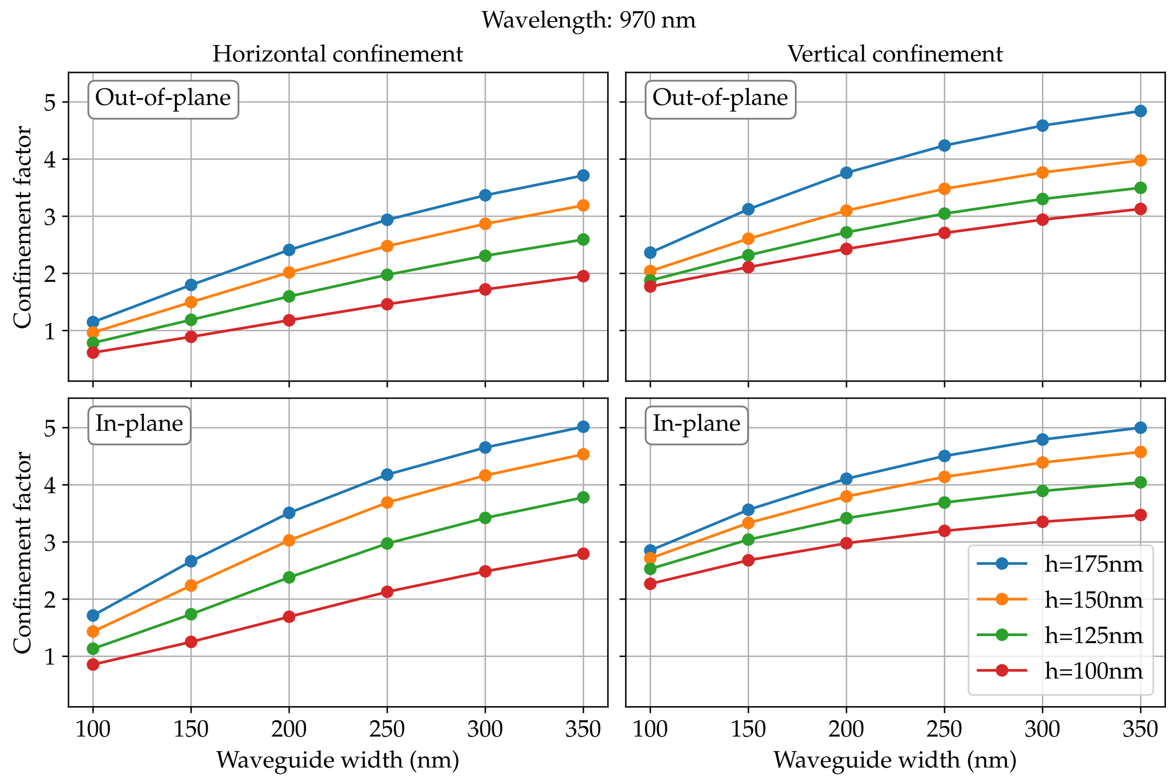


Figure A.6: Confinement factors for different waveguide widths and heights at a wavelength of 970 nm. The confinement factor is defined as the ratio of the free space wavelength to the average decay lengths in transversal directions.

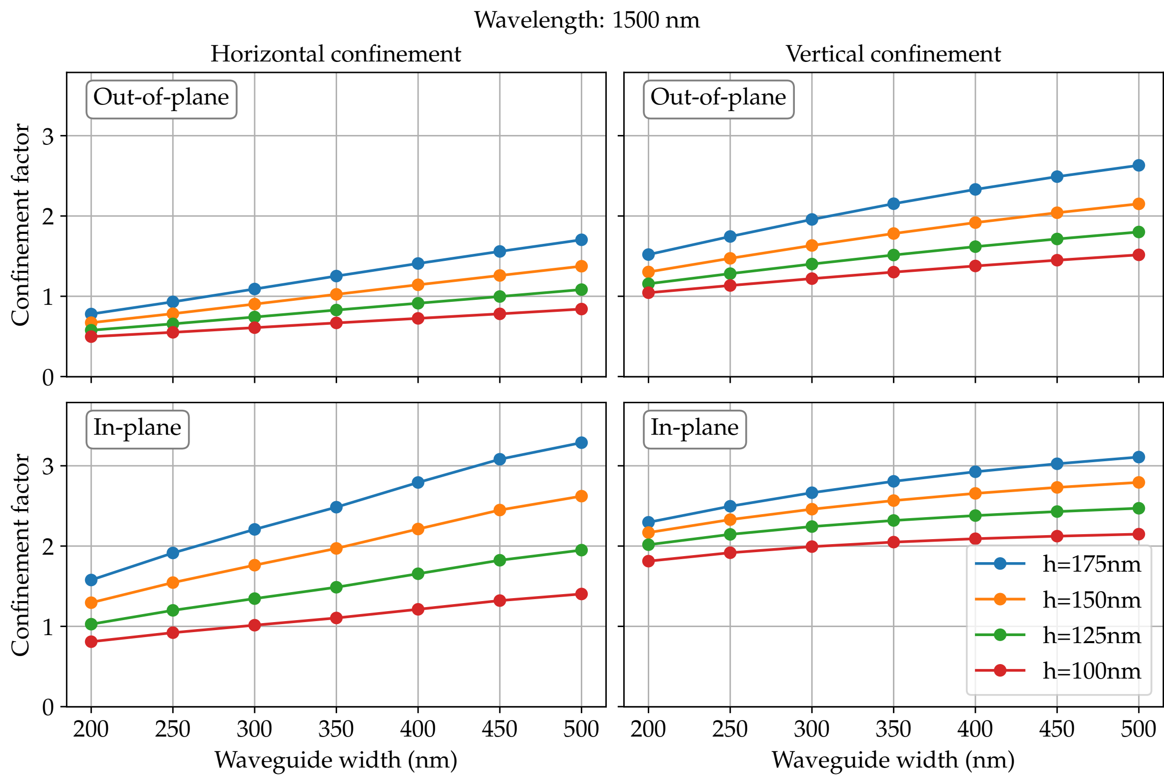


Figure A.7: Confinement factors for different waveguide widths and heights at a wavelength of 1 500 nm. The confinement factor is defined as the ratio of the free space wavelength to the average decay lengths in transversal directions.

A.2.2 PDM for the Adaptation of the Simulations

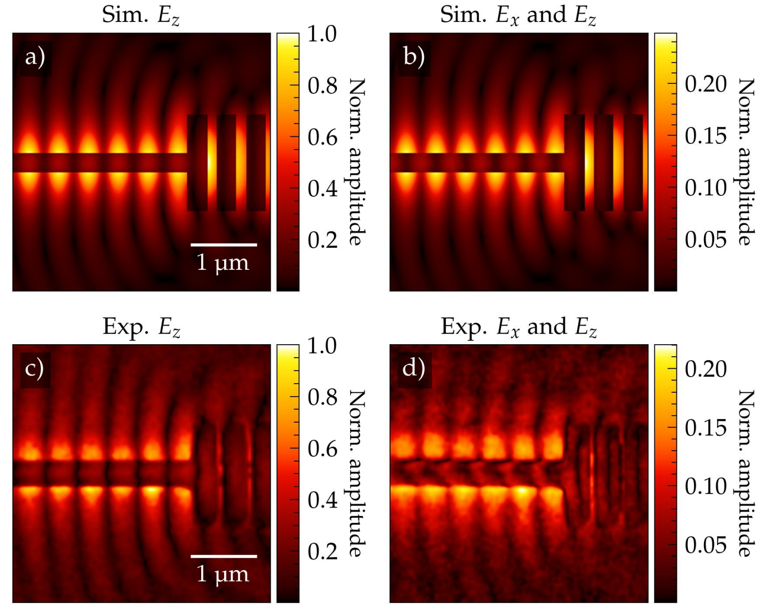


Figure A.8: Direct comparison of the measurements and simulations of the reflection grating. The first row (a, b) shows the modified simulation data, and the second row (c, d) the measurements. The simulation data in (a) is the combination of the out of plane component above the gold film and waveguide respectively, corrected by the relative polarizabilities. The simulation data in (b) is a combination of the out-of-plane component above the gold film combined with the in-plane-component above the waveguide corrected by the relative polarizabilities. The out-of-plane component above the gold film is further multiplied with a factor to emulate the reduction of this component due to the analyzer. (c, d) are the respective measurements of the out-of-plane and in-plane component.

The adaptation of the simulation data to match the measurements shown in Fig. A.8 is purely qualitative and was already introduced in the main text, see Section 6.2.3. A more straightforward quantitative approach would be to use the point dipole model (PDM) as described in Section 4.2.4 to calculate a superposition of the two electric field components. The main idea is, that the measurements do not represent the pure field components, possibly caused by a slight misalignment of the analyzer. This is apparent especially for the measurement of the in-plane component. Here, we do not expect to measure any significant field on the gold film next to the waveguide because this component is suppressed via the tip-sample interaction, as we know from the PDM. However, we do see a relatively strong signal on the gold film next to the waveguide with the same distribution as the out-of-plane component. We therefore assume that a small portion of the out-of-plane component is mixed into the in-plane measurement. As an example, let us assume a 10% mixing of the out-of-plane component into the in-plane measurement. We can then calculate a resulting superposition by firstly scaling the simulation electric field amplitudes with the ratios calculated using

Eq. (4.31) and Eq. (4.32). This gives us the polarizability corrected simulation results. From these we can then calculate the superposition by adding 10% of the out-of-plane component to the in-plane component. Please note, that the superposition has to be calculated using the complex electric fields including amplitude and phase, not just the amplitudes. The results are shown in Fig. A.9, where (a) shows the out-of-plane component, and (b) the superposition of the in-plane component with a 10% mixing of the out-of-plane component.

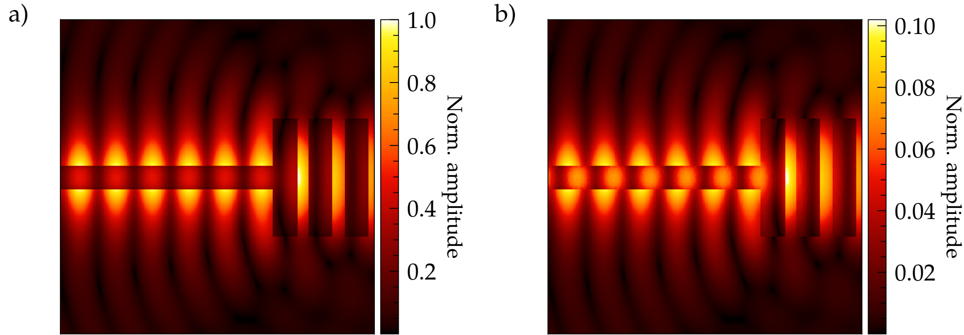


Figure A.9: (a) Modified simulation data of the out-of-plane component above the gold film and waveguide respectively, corrected by the relative polarizabilities using the PDM. Superposition of the out-of-plane component and the in-plane component corrected by the relative polarizabilities using the PDM, with a 10% mixing of the out-of-plane component into the in-plane component.

Compared to the qualitative approach shown in Fig. A.8 (b), the superposition shown in Fig. A.9 (b) looks similar, but not exactly the same. The mixing of the two components leads to a shift of the amplitude maxima on the waveguide towards the positions of the out-of-plane component. Next to the waveguide the amplitude on gold is dominated by the out-of-plane component, as the in-plane component is strongly suppressed. This validates our initial qualitative approach, but it also shows that the model is not sufficient to fully reconstruct the measurements from the simulations.

Additionally, the applied PDM is strictly speaking only valid for reflection type s-SNOM measurements. The factor of $(1 + c_r r)^2$ in Eq. (4.31) and Eq. (4.32) is to take into account the excitation and emission of the tip. The incident light can be divided in two beam paths, one directly incident on the tip and one reflected from the sample towards the tip. The same applies for the scattered light from the tip towards the detector. Therefore, the factor $(1 + c_r r)^2$ appears in the equations. However, in the current case of a transmission type s-SNOM setup this factor is not strictly valid anymore, as there is no reflected beam path for the incident light. Instead the incident light creates a surface wave, which then excites the tip. The factor should thus only appear for the emission path, leading to a factor of $(1 + c_r r)$ instead of $(1 + c_r r)^2$. If we apply this change to the PDM and recalculate the superposition, we get the results shown in Fig. A.10. One can directly see that the results differ significantly from the previous approach and the measurements. We can therefore not reliably use the PDM to reconstruct the measurements from the simulations, as the model is not fully valid

for our setup configuration. So far no better model could be found to describe the tip-sample interaction in transmission type s-SNOM setups.

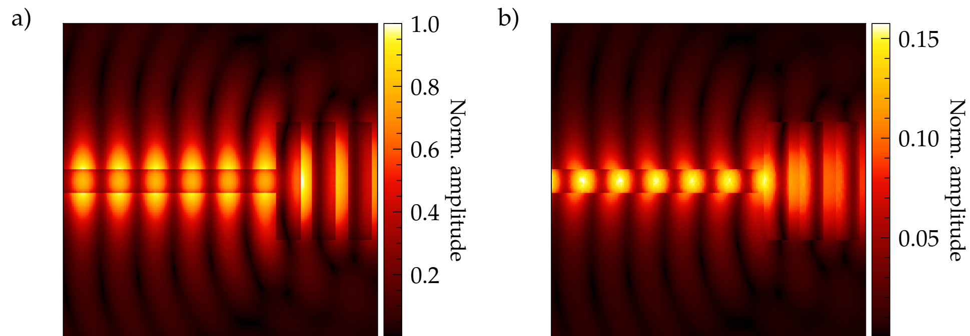


Figure A.10: (a) Modified simulation data of the out-of-plane component above the gold film and waveguide respectively, corrected by the relative polarizabilities using the PDM. Superposition of the out-of-plane component and the in-plane component corrected by the relative polarizabilities using the PDM, with a 10% mixing of the out-of-plane component into the in-plane component. In both cases the factor $(1 + c_r r)$ is used instead of $(1 + c_r r)^2$ to account for the transmission type s-SNOM setup.

A.2.3 Three Dimensional Field Mapping

A further, less common technique is the three dimensional scanning. Here, the tip is raster scanned over the sample, as in the default scanning mode, but at each pixel an approach (or deproach) curve is taken. This allows us to get more insight into the fields above the sample rather than just on the surface. However, the measurements need to be treated carefully as they do not correspond to the actual field amplitude and phase, as the scattered field from the tip is not only dependent on the local field but also on the tip-sample coupling. The coupling is strongly dependent on the tip-sample distance and can lead to a strong modification of the measured amplitude and phase. Furthermore, background induced effects might also play a role. Approach (or deproach) curves are often used to identify if the signal has a significant background contribution. Therefore, we will not try to extract quantitative values from these measurements but rather use them to get a qualitative understanding of the fields above the waveguide. An example of such a measurement is shown in Fig. A.11. Here, a DLSPPW at 970 nm was measured. The fields besides the waveguide seem to decay quite fast, while above the waveguide the fields seem to extend further into the z-direction. This is not completely unexpected, as we know from the mode profile simulations that a significant portion of the field leaks out at the top of the waveguide. It also has to do with the coupling of the tip to the underlying material and possibly background effects. Especially the strong variations in the phase images are unexpected and hard to explain.

A.2.4 Fabrication Problems

During the various fabrications of DLSPPW samples, several problems occurred. The most important problems are discussed in the following.

Single Waveguides with Negative Tone Lithography

An example of unsuccsessfully fabricated waveguides is shown in Fig. A.12. Here, many waveguides are fragmented, which could be caused by multiple things. One possibility is that the resist (stored in ethyl lactate) did not crosslink properly or became brittle and got removed in the developement process. Another possibility is that the scanning electron microscope (SEM) had a malfunction during the writing process, for example a timing issue with the blanker or the deflection unit. Another reason could be a too high dose, which could lead to stress or burning of the resist. Especially in the shown example image the upper waveguides, which are written with a higher dose, show stronger fragmentation. Switching to a resist dissolved in anisol lead to better results. But for this resist, fine atomic force microscopy (AFM) scans revealed that the waveguides often have a rough surface like shown in Fig. A.13. The roughness is most likely caused by non optimal crosslinking of the resist, especially for lower doses. This drastically limits the applicability of this fabrication method to fabricate thin and narrow waveguides. Up to this point many samples have been imaged with the light microscope to verify the writing process. This is useful, because it allows to see if the structrues are written and if the dose was chosen correctly. If structures

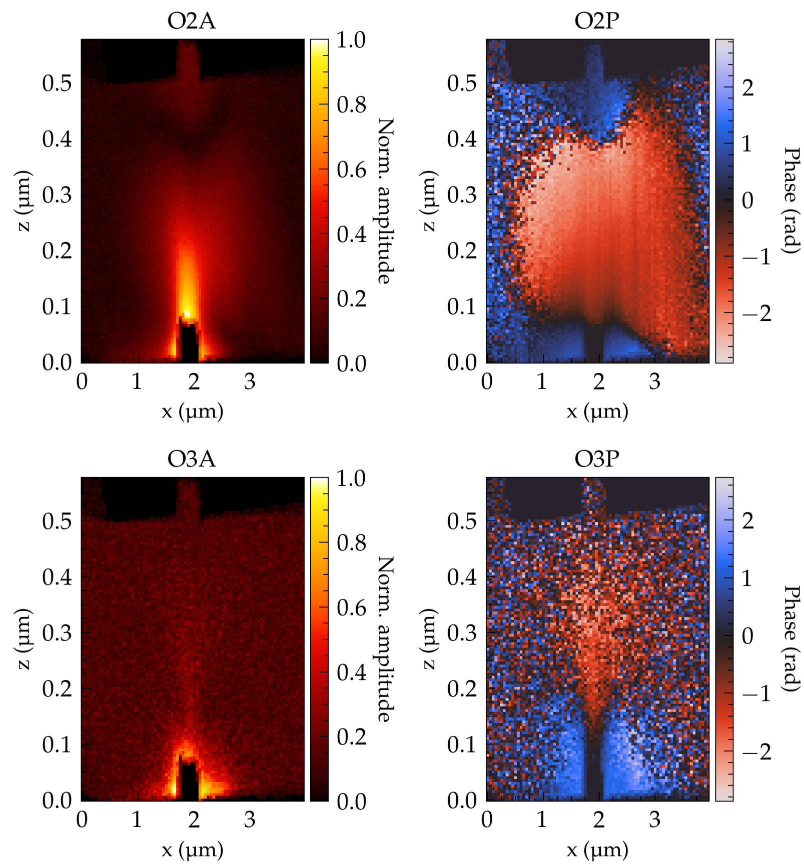


Figure A.11: Three dimensional s-SNOM measurement of a DLSPPW at a wavelength of 970 nm. The tip is raster scanned over the sample while recording approach curves at each pixel. Shown is an averaged cutplane along the y-axis (perpendicular to the waveguide) for four different optical channels. The lower z-coordinate of each approach curve was used to align the data, leading to a more physical representation.

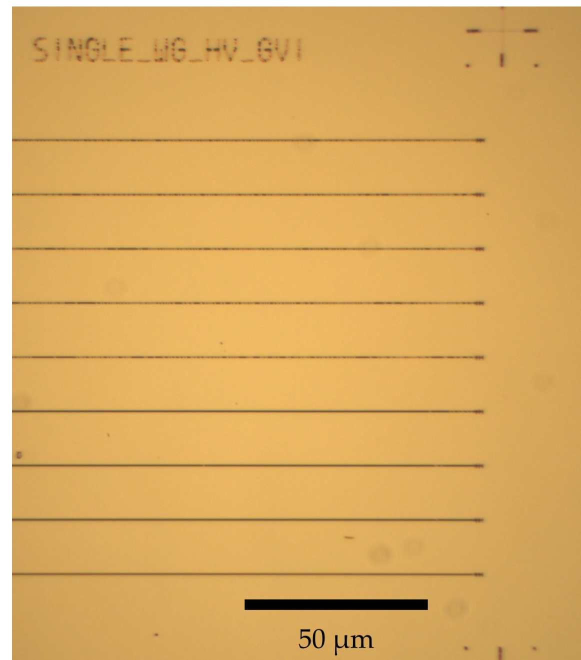


Figure A.12: Example of an array of single waveguides fabricated with poly-methyl-methacrylate (PMMA) negative tone lithography. A large portion of the waveguides is fragmented. The reason for this is most likely linked to the resist but an SEM malfunction cannot be ruled out. The font shares the same issue. The dose in this array is increased from the bottom to the top.

are under exposed they can still be seen in the light microscope, but they will not remain after the developement process. But that raises the question whether the imaging before the developement could have an influence on the final result. Thus, a test was performed where the sample was not imaged before the developement process and was kept as dark as possible. This sample did show a slightly better surface quality, but the waveguides were still quite rough. But it seems as if the resist is sensitive to light after electron beam exposure, especially in regions close to crosslinked areas. Thus, all following samples were kept in the dark as much as possible and are not imaged before the developement process.

As a test, waveguides were also fabricated on glass substrates instead of gold films. These structures did not suffer from the grainy surface, but they did show clustered dirt particles similar to the grainy surface but only locally. An example image of a waveguide fabricated on glass is shown in Fig. A.14. This raised the question if the PMMA bottle was contaminated with small undesolved particles. But a surface study before and after coating and removing PMMA did not show agglomerations of particles like in the previous examles. The acceleration voltage was also varied to see if this has an influence on the surface quality. The results are shown in Fig. A.15. The acceleration voltage does not seem to have a strong influence on the surface quality, but in general a higher dose is required for higher acceleration voltages. This is also discussed in Section 3.1.4.

As an alternative to PMMA as a negative tone resist, the resist Medusa was also tried. This

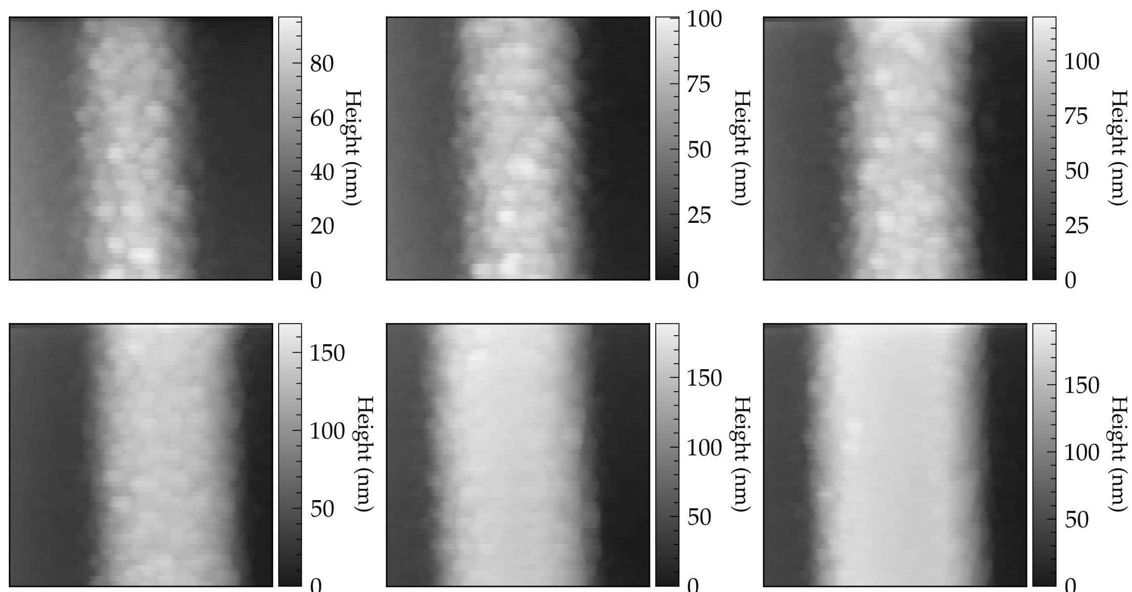


Figure A.13: Example AFM scans of waveguides fabricated with PMMA negative tone lithography. Each scan is 1 µm x 1 µm. The surface of the waveguides is very rough and grainy. For an increasing width the center of the waveguides becomes smoother, but the edges remain rough.

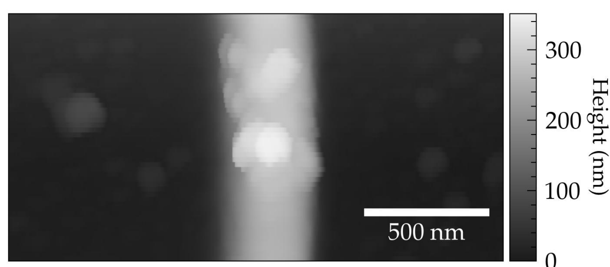


Figure A.14: Example of a waveguide fabricated with PMMA negative tone lithography on a glass substrate. Apart from some clustered dirt particles the surface is very smooth.

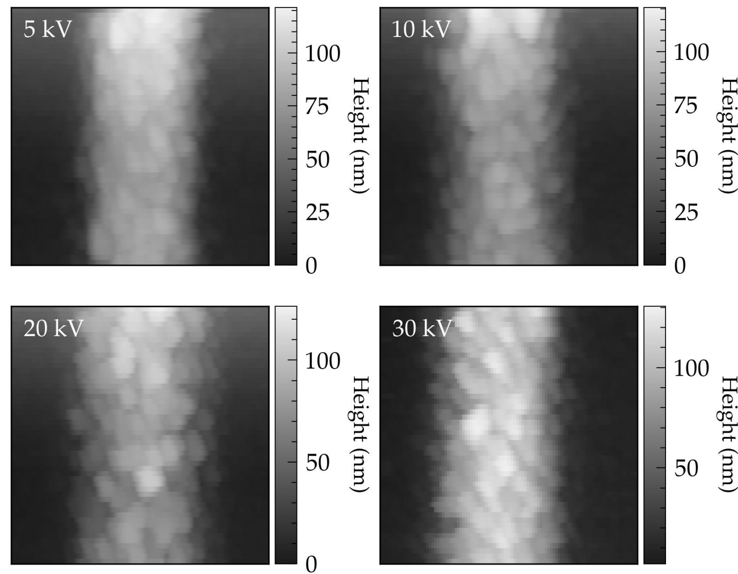


Figure A.15: Example AFM scans of waveguides fabricated with different acceleration voltages. Each scan is 1 µm x 1 µm. The surface quality of the waveguides does not seem to depend strongly on the acceleration voltage.

resist is especially made for negative tone lithography and should thus be more suitable. The resulting waveguides were indeed much smoother, but the minimal width was limited to approximately 700 nm at a height of only 120 nm. This is most probably due to the electron backscattering from the gold film, as much narrower structrues are possible on glass substrates. The resist also shows detachment issues for lower doses, which makes it less suitable for this application. A similar behavior is observed for PMMA negative tone lithography as well. On glass substrates much thinner and higher waveguides are possible. For example a height of 280 nm at a width of only 270 nm was measured.

Coupled Waveguides with Negative Tone Lithography

When fabricating coupled waveguides with negative tone lithography, similar problems occured. An example is shown in Fig. A.16. Here, the waveguides are written with a center to center distance of 500 nm and a width of 200 nm each. Due to proximity effects in EBL, the waveguides are partially merged together. This effect could not be completely avoided by dose adjustments or using proximity effect correction algorithms. Thus, a new method based on positive tone EBL was developed, as described in Appendix A.2.5. It shoud be noted that the grainy surface texture and the problems in general did not occure every time, but often enough to make this fabrication method unreliable.

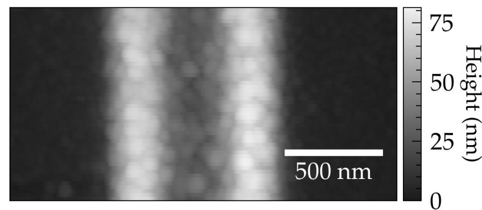


Figure A.16: Example of closely spaced waveguides fabricated with PMMA negative tone lithography showing proximity effects and a grainy texture. The waveguides are written with a center to center distance of 600 nm. Due to proximity effects in EBL, the waveguides are partially merged together.

A.2.5 Fabrication of coupled waveguides with positive tone EBL

The positive tone lithography process using PMMA is described in Section 3.1.3. This process was found to be much more suitable for fabricating closely spaced coupled waveguides. An example image is shown in Fig. A.17. Here, two waveguides are written with a center to center distance of 600 nm and a width of 200 nm. In order to achieve this the areas surrounding the waveguides need to be exposed. Due to the proximity effect we need to choose different doses for the thin area between the waveguides and the larger area surrounding them. For a better sample edge quality we also expose the areas surrounding the waveguides as individual lines with a linedistance of 5 nm. The required doses need to be fine tuned for each specific distance and waveguide width.

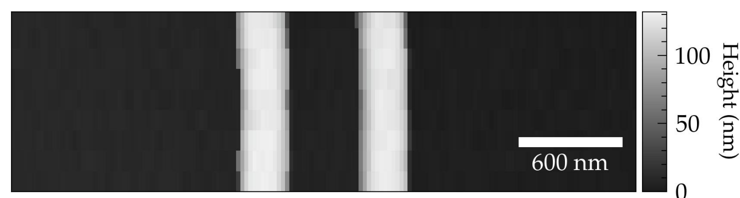


Figure A.17: Example of closely spaced waveguides fabricated with PMMA positive tone lithography. The waveguides are written with a center to center distance of 600 nm and have a width of 200 nm. No proximity effects are visible and the surface quality is very good.

This applies not only to two coupled waveguides but also to larger arrays of waveguides with varying distances. Three example AFM profiles of a topological SSH waveguide arrays fabricated with this process are shown in Fig. A.18. The dose factors for these arrays are (45 and 50, 27 and 30, 22 and 25) for a design width of 350 nm, where a higher dose leads to more proximity effect and narrower structures. Ideally not only the dose is varied but the design width as well to achieve the best possible result. Overall the dose factor should be kept as low as possible, close to the polymerization threshold, to reduce the proximity effect.

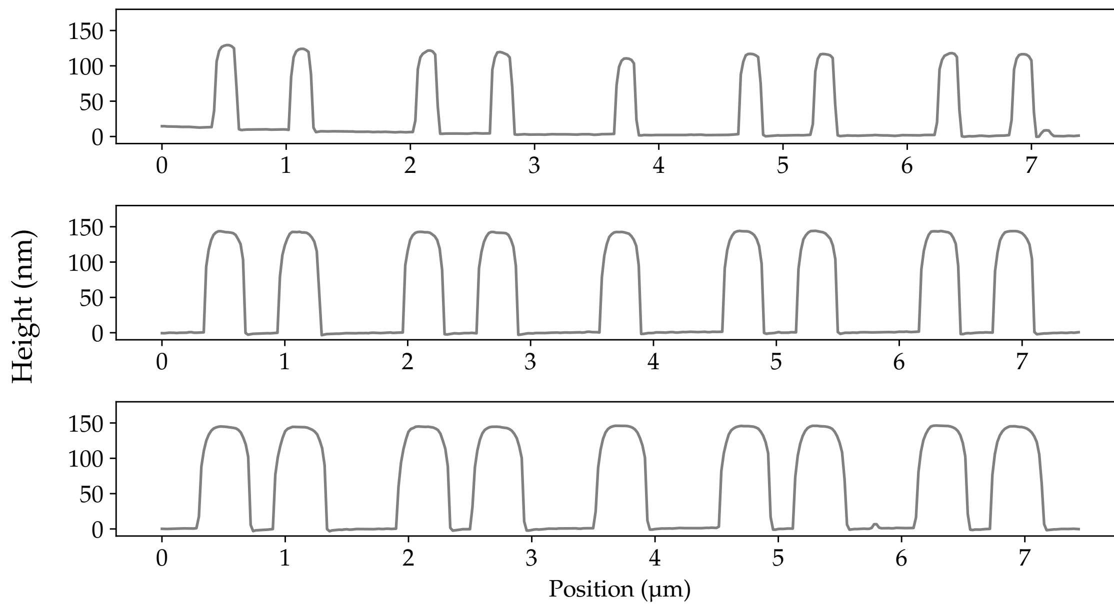


Figure A.18: Example AFM profiles of a topological SSH waveguide array fabricated with PMMA positive tone lithography. The waveguides are written with a center to center distance of 600 nm and 1 000 nm respectively. The overall quality and uniformity of the waveguides is very good. The shape and minimal and maximal waveguide width are much better defined compared to the negative tone lithography process.

A.3 Silver Crystals

Single-crystalline metal films are a superior material to their thermally evaporated counterparts, as they have a much lower surface roughness and a lower density of grain boundaries and defects. This leads to lower optical losses and longer propagation lengths of SPPs [140–142]. Additionally, the smooth surface makes them an ideal platform for FIB milling of nanostructures, such as plasmonic gratings. Polycrystalline metal films on the other hand have a much higher surface roughness, which makes it difficult to mill fine structures with a FIB, because the grain boundaries scatter the ions and lead to an uneven milling depth [143]. This can be seen in the SEM images of structures milled into polycrystalline gold films in Fig. 3.3 and Fig. 3.6.

Silver crystals are not used in this thesis, but have been prepared in collaboration with Yuhao Zhang, a former PhD student in our group. The silver crystals have been used as a metallic platform for the creation of plasmonic gratings, which were then coupled to excitons in tungsten disulfide monolayers. For more details on the experiments please refer to [144]. Here, we briefly describe the fabrication process of the silver crystals and the subsequent milling of the grating structures with the FIB.

The growth of the silver crystals is based on the work of Chun-Yuan Wang et al. [145]. They use a wet chemical synthesis to grow large single crystalline silver plates from silver nitrate in solution. The growth process is based on a platinum-catalyzed, ammonium hydroxide (NH_4OH) controlled polyol reduction method [146]. The difference to the older method is mainly the high concentration of ammonium hydroxide, which helps to control the reduction rate of the silver ions and thus the crystal growth.

A.3.1 Growth Recipe

In the following we briefly describe our adapted growth recipe for the silver crystals.

1. 15 mL of ethylene glycol are poured in a clean glass beaker. For a more precise volume control the ethylene glycol is weighted with a scale to $16.7\,g \pm 0.2\,g$.
2. 433 mg of silver nitrate ($AgNO_3$) are added to the ethylene glycol and stirred using a magnetic stirrer until fully dissolved.
3. Add 0.5 g of poly-vinyl-pyrrolidone (PVP, molecular weight $55.000\,g\,mol^{-1}$) and stir until fully dissolved, but not longer than 5 min.
4. Add 1.85 mL of ammonium hydroxide (NH_4OH, 28% wt) and stir for another 30 s.
5. Add 0.54 mL of a $0.02\,mol\,L^{-1}$ chloroplatinic acid (H_2PtCl_6) solution in water and stir again shortly.
6. Add 1.3 mL of hydrogen peroxide (H_2O_2, 30% wt) and stir shortly again. The reaction will start immediately, and disturbances should be avoided.

7. Pour the solution into a clean glass petri dish and cover it to avoid dust contamination. Let it sit undisturbed for several days at room temperature.

8. In order to grow the crystals directly on a substrate, such as silicon wafer pieces, the substrate can be placed in the petri dish prior to pouring the solution. The substrates should be cleaned properly.

During and after the reaction, the solution should remain clear, otherwise the reaction might have started prematurely or the chemicals were contaminated. After several days, large silver crystals should have formed in the solution. The samples are then removed from the solution and rinsed with deionized water. One should be careful though as most crystals are only weakly attached to the substrate and can easily be washed away. After that the still wet samples are rinsed with ultrapure acetone and dried with nitrogen. Many crystals will only adhere properly after the drying step. The samples are covered by losely attached crystals, which can be picked up with a polydimethylsiloxane (PDMS) stamp, leaving only the well attached crystals on the substrate. This step is important, because it reduces the amount of weakly attached crystals. In a later procedure, transition metal dichalcogenide (TMDC) monolayers are transferred onto the crystals using a PDMS stamping technique. If the crystals do not adhere well enough, they may lift of later during the transfer. To protect the silver surface from sulfidation, a second PDMS stamp is used, which is only removed right before the FIB milling.

An example AFM scan of the surface of a silver crystal is shown in Fig. A.19. This measurement was selected to show that even atomic terraces are visible on the surface. The surface roughness in the center area is around 120 pm root mean square.

A.3.2 Fabrication of Nanogroove Arrays

The nanogroove arrays are milled as vertical rectangles from left to right. Each rectangle consists of multiple lines, typically around 7, with a spacing of 5 nm. The angle of each successive rectangle is rotated to create the desired groove spacing gradient. Depending on the spectral needs, the lower and upper groove spacing can be adjusted. A typical example would be a distance variation of 400 nm over a length of 40 µm. The shape and depth of the grooves can be controlled by the width of the rectangles and the line dose. As discussed before, the depth of the milled structrues depends on the milling procedure. We found that a single line with a high dose leads to a significant depth variation on the first few micro meters, but afterwards the groove depth is quite uniform. As a compromise, to ensure an overall uniform depth, while maintaining sharp features, we use a moderate dose and a line count of around 7. Writing several lines close to each other is essentially the same as looping, but the width of the grooves can be adjusted as well. An example SEM image of the milled nanogroove arrays is shown in Fig. A.20.

The grooves are written with a beam current of 30 to 50 pA and a line dose of roughly $5\,\mathrm{nC\,cm^{-1}}$, depending on the desired groove depth. The dose values have to be calibrated each time, especially because the FIB system was quite unstable during the time of this

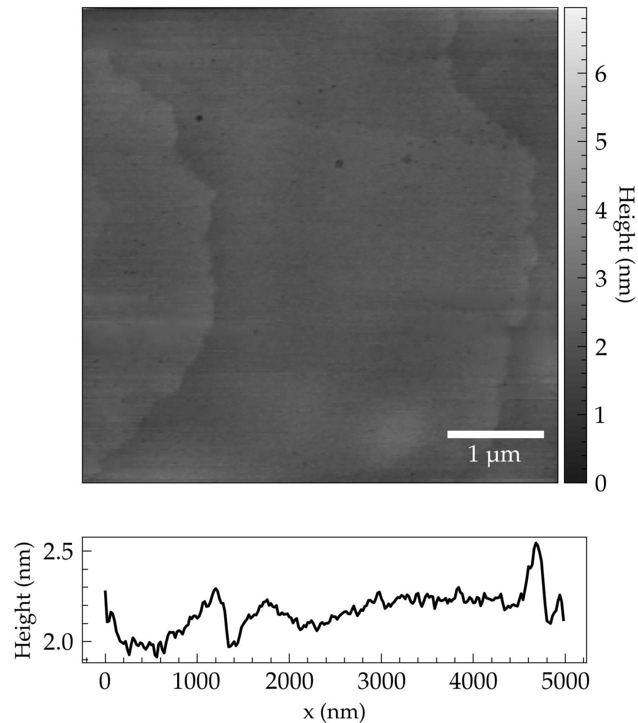


Figure A.19: Leveled AFM roughness scan of the silver crystal surface. The lower plot shows a linecut of the surface roughness in the middle averaged over 10 lines.

thesis, leading to significant current drifts over time. In order to identify the right dose, a test pattern consisting of single rectangles with varying dose or line count are milled next to the actual structures. A horizontal crosscut of the test pattern is then imaged with the SEM at an angle of 54°, which allows to measure the groove depth and shape.

After the milling process, the silver crystal samples are covered with a PDMS stamp to protect the silver surface from sulfidation. The samples can be stored like this for several days before degradation effects become noticeable. While it has to be considered that the PDMS could lead to residues affecting the optical quality of the sample, practice showed that small bubbles form around the silver crystals, leaving most of the crystal surface untouched. Thus, the influence on the optical experiments is minimal and degradation effects are usually the most important factor. A better approach to protect the silver crystals is to cover them with a thin layer of alumina via atomic layer deposition or thermal evaporation. The aluminum layer protects the silver from sulfidation and can be removed with potassium hydroxide (KOH), as this does not dissolve silver. As a sidenote, the aluminum can also act as a sacrificial layer during FIB milling, allowing to achieve finer features [147]. Other metals, such as chromium, can also be used for this purpose, especially in the context of wet etching [148]. But this process was not used for the nanogrooves, only for test structures.

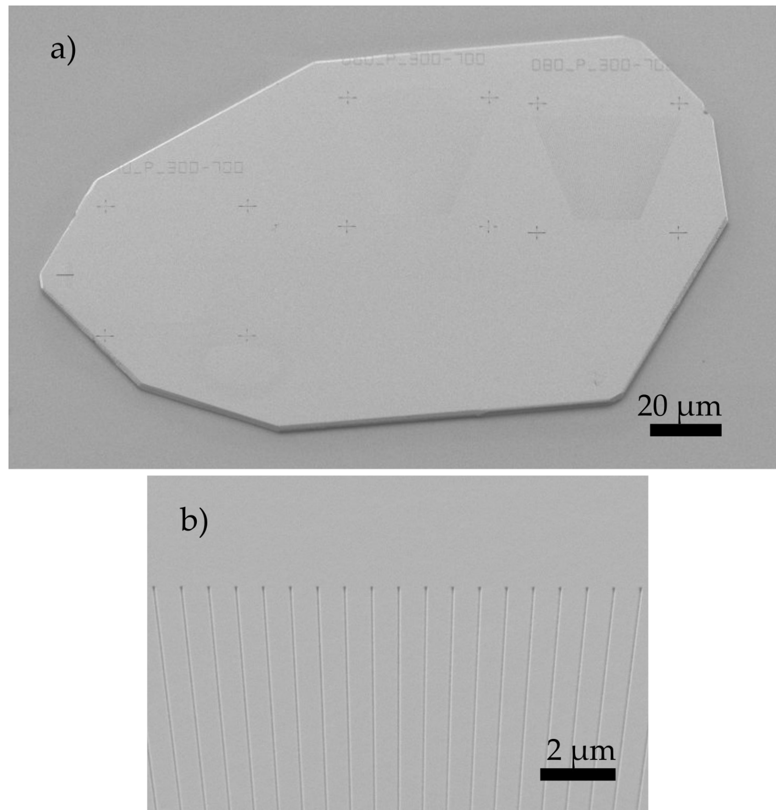


Figure A.20: SEM images of nanogrooves milled into silver crystals with a focused ion beam. The separation of the grooves is varied continuously from 700 nm to 300 nm from top to bottom. The grooves have a width of around 80 nm and a depth of around 50 nm. (a) Large silver crystal with multiple nanogroove arrays milled into it. (b) Close-up of one of the nanogroove arrays.

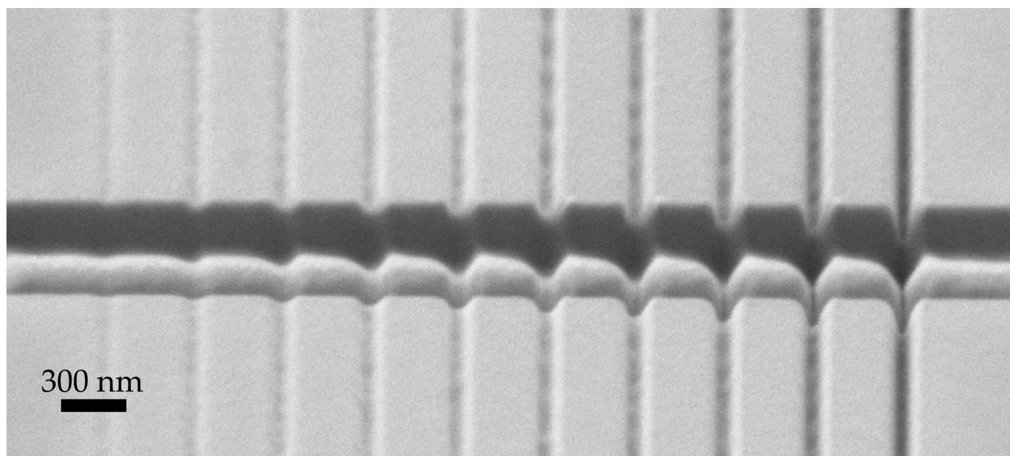


Figure A.21: SEM image of a nanogroove dosetest milled into a silver crystal with a focused ion beam. The dose increases from left to right, leading to deeper grooves. The SEM images the sample at an angle of 54°, the image is stretched to display the depth of the grooves correctly.

A.4 Incorporated Software

During this work several software programms and packages have been developed, the most important ones are described briefly in the following. All programms are accessible in the working group or can be recieved on reasonable request.

A.4.1 Elphy/Raith Software

EBL is based on the exposure of an electron sensitive resist using a focused electron beam. The beam is controlled using deflection coils, which allow to move the beam in x and y direction over the sample surface. The scannable area, also called the writefield, of the beam is limited by the deflection coils and depends on the magnification used. Typical writefield sizes are 100 µm x 100 µm or 200 µm x 200 µm. To expose larger areas, or place more structures, multiple writefields can be stitched together in a position list. The accuracy of the stitching depends on the SEM used, but is typically in the range of a few micro meters. To define positions in multiple writefields, an overall u-v coordinate system is used, where each writefield has its own origin at the center of the writefield. The position list for multiple writefields is then defined in this overall u-v coordinate system.

In this work, a SEM with the *Elphy* (*Raith*) software installed on the SEM pc is used for EBL. The *Elphy* software allows to create designs directly on the SEM pc or load external files. For example, ascii files can be loaded, which contain a list of elements to be exposed. Each element has a header containing the element type (dot, line, area, font, ellipsoid ...), the dose factor and the layer. The element type is important for the *Elphy* software to interpret the following data correctly. For example, areas are automatically filled with lines, while lines are exposed as a series of dots. The dose factor is given relative to the base dose defined in the *Elphy* software. The layer is used to separate different exposure layers, for example multiple steps of a fabrication process. The header is followed by a list of x-y coordinates defining the element. All coordinates are given in micrometer relative to the center of the writefield. Around a two dimensional map of all coordinates a bounding box is defined, which is then shifted to the lower left corner of the writefield. Markers should be added in the corners of the writefield, if the coordinates should be defined relative to the writefield center and neighboring writefields. The bounding box is then automatically equal to the writefield. These markers do not need to be exposed, they only need to be present in the ascii file. Some layers are reserved for internal functions, such as layer 63, which is used for alignment of markers and is important for multi step exposures.

To create these ascii files, a custom python script was developed. The working principle is based on object oriented programming. The main class is the object class containing all main commands for creating the ascii file and writing into it. Then, subclasses for specific geometric objects such as rectangles, circular discs, rings, triangles ... are defined. Even a custom font was developed based on a simple 7 line digital clock interface with additional features for string characters. Each subclass then creates a geometric object based on input parameters such as position, size, dose factor, layer ... The object is then, depending on the

line type, filled with lines or added as a polygon area to the ascii file. The filling with lines is done by adding connected lines from inside out with the defined line spacing until the defined shape is filled. If the area line type is used, this will be done automatically by the *Elphy* software during exposure, but this leads to less control over the exposure. The scripts also contains an object type for various marker types, such as those used in the multi step exposure process.

The created ascii files can also be visualized using matplotlib. Here, two plotting modes are implemented. The default will plot the dose factors color coded, while the second mode plots the individual elements with incremental colors to visualize the temporal exposure sequence.

A.4.2 SNOM/AFM Analysis Software

AFM and s-SNOM measurements generate manny different channels containing binary data, which need to be analyzed and visualized. For this purpose, a custom python package was developed to handle and analyze the data. The software is specifically tailored to work with data from the *NeaSNOM* system used in this work, but can easily be extended to work with other file formats as well.

The *NeaSNOM* system allows for different measurement modes such as two dimensional scans, one dimensional approach curves or three dimensional scans (two dimensional maps of approach curves). Each measurement mode has its own file format and structure. The most commonly used measurement mode in this work is the two dimensional scan mode, where the tip is raster scanned over the sample surface while recording multiple channels such as topography, optical amplitude and phase at different demodulation orders of the tip tapping frequency. The data is saved in binary ".gsf" files, which contain a header with essential parameters such as scan size, pixel resolution and other metadata. Additional parameters are saved in accompanying ".html" and ".txt" files. For a comprehensive analysis, it is often not sufficient to analyze only one individual channel, but rather a combination of multiple channels.

The developed python package (SNOM_AFM_analysis, *Github*) is based on object oriented programming principles. A loader class is implemented to handle the loading of .gsf files and to extract information from the header and additional files. This class contains many low level functions that are shared across all three implemented measurement modes. Then, three subclasses are defined for the different measurement modes, inheriting from the loader class. Each subclass has its own functions to initialize the data, plot the data and apply modifications. For example, functions for flattening the topography, cropping the size, and saving the modified data are included. A detailed description of the software can be found in the code documentation (Documentation, *ReadTheDocs*).

The software provides a versatile toolbox for s-SNOM and AFM data analysis, and is easily extendable by adding additional functionality. The software is still in a beta state, being operational but certain aspects need improvement to allow easy access for a wider audience. A graphical user interface (GUI) was also developed (GUI, *GitHub*) to make the

software more user friendly and to quickly generate images. The software and GUI have been indispensable for the analysis of all s-SNOM and AFM data in this work.

A.4.3 Comsol Software

The *Comsol Multiphysics* software is used to perform finite element method (FEM) simulations of the plasmonic waveguide structures. The software allows to set up complex geometries, define material properties, and simulate the electromagnetic fields using different solvers. The simulation results can be visualized directly in the software, but for further analysis and comparison with experimental data, the raw data needs to be exported and processed. For this purpose a custom python script was developed to handle the exported data from *Comsol*. The simulation data can be exported as a text file. Depending on the study type, the data format can vary, for example between 2d and 3d studies. The developed python script is able to load these text files and convert them into a format compatible with the s-SNOM analysis software described in Appendix A.4.2. This allows to directly compare simulation results with experimental s-SNOM data. The script also includes functionality to cut 3d data sets into 2d slices, which is useful for visualizing cross sections of the simulated fields. Additionally, the script can handle complex data and automatically generates amplitude and phase images from the real and imaginary parts of the fields.

Bibliography

[1] I. Koren and C. M. Krishna, *Temperature-Aware Computing*, Sustainable Computing: Informatics and Systems **1** (2011) 46, ISSN: 2210-5379 (cit. on p. 1).

[2] H. Kawaura, T. Sakamoto and T. Baba, *Observation of Source-to-Drain Direct Tunneling Current in 8 Nm Gate Electrically Variable Shallow Junction Metal–Oxide–Semiconductor Field-Effect Transistors*, Applied Physics Letters **76** (2000) 3810, ISSN: 0003-6951 (cit. on p. 1).

[3] B. J. Shastri et al., *Photonics for Artificial Intelligence and Neuromorphic Computing*, Nature Photonics **15** (2021) 102, ISSN: 1749-4893 (cit. on p. 1).

[4] N. L. Kazanskiy, M. A. Butt and S. N. Khonina, *Optical Computing: Status and Perspectives*, Nanomaterials **12** (2022) 2171, ISSN: 2079-4991 (cit. on p. 1).

[5] I. Bente et al., *The Potential of Multidimensional Photonic Computing*, Nature Reviews Physics **7** (2025) 439, ISSN: 2522-5820 (cit. on p. 1).

[6] D. Pérez-López and L. Torrijos-Morán, *Large-Scale Photonic Processors and Their Applications*, npj Nanophotonics **2** (2025) 32, ISSN: 2948-216X (cit. on p. 1).

[7] S. Hua et al., *An Integrated Large-Scale Photonic Accelerator with Ultralow Latency*, Nature **640** (2025) 361, ISSN: 1476-4687 (cit. on p. 1).

[8] S. R. Ahmed et al., *Universal Photonic Artificial Intelligence Acceleration*, Nature **640** (2025) 368, ISSN: 1476-4687 (cit. on p. 1).

[9] P. L. McMahon, *The Physics of Optical Computing*, Nature Reviews Physics **5** (2023) 717, ISSN: 2522-5820 (cit. on p. 1).

[10] Z. Ying et al., *Electronic-Photonic Arithmetic Logic Unit for High-Speed Computing*, Nature Communications **11** (2020) 2154, ISSN: 2041-1723 (cit. on p. 1).

[11] E. Abbe, *Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung*, Archiv für Mikroskopische Anatomie **9** (1873) 413, ISSN: 0176-7364 (cit. on pp. 1, 2, 27).

[12] L. R. F.R.S., *XXXI. Investigations in Optics, with Special Reference to the Spectroscope*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **8** (1879) 261, ISSN: 1941-5982, 1941-5990 (cit. on pp. 1, 2, 27).

[13] R. Zia, J. A. Schuller, A. Chandran and M. L. Brongersma, *Plasmonics: The next Chip-Scale Technology*, Materials Today **9** (2006) 20, ISSN: 1369-7021 (cit. on p. 2).

[14] D. K. Gramotnev and S. I. Bozhevolnyi, *Plasmonics beyond the Diffraction Limit*, Nature Photonics **4** (2010) 83, ISSN: 1749-4893 (cit. on p. 2).

[15] W. L. Barnes, *Surface Plasmon–Polariton Length Scales: A Route to Sub-Wavelength Optics*, Journal of Optics A: Pure and Applied Optics **8** (2006) S87, ISSN: 1464-4258 (cit. on p. 2).

[16] M. Hojjat Jodaylami, J.-F. Masson and A. Badia, *Surface Plasmon Resonance Sensing*, Nature Reviews Methods Primers **5** (2025) 47, ISSN: 2662-8449 (cit. on p. 2).

[17] T. M. S. Ashrafi and G. Mohanty, *Surface Plasmon Resonance Sensors: A Critical Review of Recent Advances, Market Analysis, and Future Directions*, Plasmonics **20** (2025) 6825, ISSN: 1557-1963 (cit. on p. 2).

[18] R. Ron, T. Zar and A. Salomon, *Linear and Nonlinear Optical Properties of Well-Defined and Disordered Plasmonic Systems: A Review*, Advanced Optical Materials **11** (2023) 2201475, ISSN: 2195-1071 (cit. on p. 2).

[19] R. Rong et al., *Beam Steering of Nonlinear Optical Vortices with Phase Gradient Plasmonic Metasurfaces*, ACS Photonics **10** (2023) 3248 (cit. on p. 2).

[20] T. Sharma, Z. Zhang, J. Wang, Z. Cheng and K. Yu, *Past, Present, and Future of Hybrid Plasmonic Waveguides for Photonics Integrated Circuits*, Nanotechnology and Precision Engineering **7** (2024) 045001, ISSN: 1672-6030 (cit. on p. 2).

[21] O. Bitton and G. Haran, *Plasmonic Cavities and Individual Quantum Emitters in the Strong Coupling Limit*, Accounts of Chemical Research **55** (2022) 1659, ISSN: 0001-4842 (cit. on p. 2).

[22] F. Badshah, Y. Zhou, M. Idrees, Ziauddin and Rahmatullah, *Vacuum-Induced Excitation of Surface Plasmon Polaritons*, Physical Review A **112** (2025) 043706 (cit. on p. 2).

[23] A. Aspuru-Guzik and P. Walther, *Photonic Quantum Simulators*, Nature Physics **8** (2012) 285, ISSN: 1745-2481 (cit. on pp. 2, 112, 114).

[24] A. A. Sukhorukov, *Classical Simulation of Squeezed Light in Optical Waveguide Arrays*, Physical Review A **87** (2013) (cit. on pp. 2, 112, 114).

[25] K. Poulios,
Quantum Walks of Correlated Photon Pairs in Two-Dimensional Waveguide Arrays,
Physical Review Letters **112** (2014) (cit. on pp. 2, 112, 114).

[26] Y. Dong, *Waveguide-QED Platform for Synthetic Quantum Matter*,
Physical Review A **104** (2021) (cit. on pp. 2, 112, 114).

[27] M. J. Maleki and M. Soroosh, *A Low-Loss Subwavelength Plasmonic Waveguide for Surface Plasmon Polariton Transmission in Optical Circuits*,
Optical and Quantum Electronics **55** (2023) 1266, ISSN: 1572-817X (cit. on p. 2).

[28] L. Huang et al., *Sub-Wavelength Patterned Pulse Laser Lithography for Efficient Fabrication of Large-Area Metasurfaces*, Nature Communications **13** (2022) 5823,
ISSN: 2041-1723 (cit. on p. 2).

[29] S. Pres et al., *Detection of a Plasmon-Polariton Quantum Wave Packet*,
Nature Physics **19** (2023) 656, ISSN: 1745-2481 (cit. on p. 2).

[30] Q. Ma et al.,
Visualization of Plasmon-Enhanced Electric Fields in Silver Dendritic Fractal Structures,
Optics Communications **591** (2025) 132109, ISSN: 0030-4018 (cit. on p. 2).

[31] F. Zenhausern, Y. Martin and H. K. Wickramasinghe, *Scanning Interferometric Apertureless Microscopy: Optical Imaging at 10 Angstrom Resolution*,
Science **269** (1995) 1083 (cit. on pp. 2, 27).

[32] F. Keilmann and R. Hillenbrand,
Near-Field Microscopy by Elastic Light Scattering from a Tip, Philosophical Transactions: Mathematical, Physical and Engineering Sciences **362** (2004) 787, ISSN: 1364-503X,
JSTOR: `4142390` (cit. on pp. 2, 40).

[33] Y. Lefier, R. Salut, M. A. Suarez and T. Grosjean, *Directing Nanoscale Optical Flows by Coupling Photon Spin to Plasmon Extrinsic Angular Momentum*,
Nano Letters **18** (2018) 38, ISSN: 1530-6984, 1530-6992 (cit. on pp. 2, 57–59).

[34] M. Thomaschewski, Y. Yang, C. Wolff, A. S. Roberts and S. I. Bozhevolnyi,
On-Chip Detection of Optical Spin–Orbit Interactions in Plasmonic Nanocircuits,
Nano Letters **19** (2019) 1166, ISSN: 1530-6984, 1530-6992 (cit. on pp. 2, 57).

[35] Y. Kan, S. Kumar, F. Ding, C. Zhao and S. I. Bozhevolnyi,
Spin–Orbit Controlled Excitation of Quantum Emitters in Hybrid Plasmonic Nanocircuits,
Advanced Optical Materials **8** (2020) 2000854, ISSN: 2195-1071 (cit. on pp. 2, 57).

[36] M. Thomaschewski et al., *Near-Field Observation of the Photonic Spin Hall Effect*,
Nano Letters **23** (2023) 11447, ISSN: 1530-6984 (cit. on pp. 2, 57).

[37] T. Holmgaard and S. I. Bozhevolnyi,
Theoretical Analysis of Dielectric-Loaded Surface Plasmon-Polariton Waveguides,
Physical Review B **75** (2007) 245405, ISSN: 1098-0121, 1550-235X (cit. on pp. 3, 73).

[38] T. Holmgaard et al., *Efficient excitation of dielectric-loaded surface plasmon-polariton waveguide modes at telecommunication wavelengths*, Phys. Rev. B **78** (16 2008) 165431 (cit. on pp. 3, 73, 80).

[39] H. Siampour et al., *Ultrabright single-photon emission from germanium-vacancy zero-phonon lines: deterministic emitter-waveguide interfacing at plasmonic hot spots*, Nanophotonics **9** (2020) 953 (cit. on pp. 3, 73, 84, 94).

[40] M. Z. Hasan and C. L. Kane, *Colloquium: Topological Insulators*, Reviews of Modern Physics **82** (2010) 3045 (cit. on p. 3).

[41] Z. Fedorova (Cherpakova) et al., *Limits of Topological Protection under Local Periodic Driving*, Light: Science & Applications **8** (2019) 63, ISSN: 2047-7538 (cit. on pp. 3, 114).

[42] S. A. Maier, *Plasmonics: Fundametals and Applications*, Springer Science+Business Media, LLC, 2007 (cit. on pp. 5–7, 9).

[43] P. B. Johnson and R. W. Christy, *Optical Constants of the Noble Metals*, Phys. Rev. B **6** (12 1972) 4370 (cit. on pp. 6, 49, 61, 74).

[44] F. von Cube et al., *Spatio-spectral characterization of photonic meta-atoms with electron energy-loss spectroscopy*, Opt. Mater. Express **1** (2011) 1009 (cit. on p. 6).

[45] K. Y. Bliokh and F. Nori, *Transverse spin of a surface polariton*, Phys. Rev. A **85** (6 2012) 061801 (cit. on p. 7).

[46] K. Y. Bliokh, A. Y. Bekshaev and F. Nori, *Extraordinary Momentum and Spin in Evanescent Waves*, Nature Communications **5** (2014) 3300, ISSN: 2041-1723 (cit. on p. 7).

[47] K. Y. Bliokh, D. Smirnova and F. Nori, *Quantum Spin Hall Effect of Light*, Science **348** (2015) 1448 (cit. on p. 7).

[48] D. O'Connor, P. Ginzburg, F. J. Rodríguez-Fortuño, G. A. Wurtz and A. V. Zayats, *Spin–Orbit Coupling in Surface Plasmon Scattering by Nanostructures*, Nature Communications **5** (2014) 5327, ISSN: 2041-1723 (cit. on p. 7).

[49] I. Zailer, J. E. F. Frost, V. Chabasseur-Molyneux, C. J. B. Ford and M. Pepper, *Crosslinked PMMA as a High-Resolution Negative Resist for Electron Beam Lithography and Applications for Physics of Low-Dimensional Structures*, Semiconductor Science and Technology **11** (1996) 1235, ISSN: 0268-1242 (cit. on p. 11).

[50] F. Rahman et al., *A Review of Polymethyl Methacrylate (PMMA) as a Versatile Lithographic Resist – With Emphasis on UV Exposure*, Microelectronic Engineering **224** (2020) 111238, ISSN: 0167-9317 (cit. on p. 11).

[51] M. Bolorizadeh and D. Joy, *Low Voltage Electron Beam Lithography in PMMA*, 2005 NSTI Nanotechnology Conference and Trade Show - NSTI Nanotech 2005 Technical Proceedings (2005) (cit. on p. 12).

[52] W. S. M. Werner et al., *Energy Dissipation of Fast Electrons in Polymethylmethacrylate*, Physical Review Letters **132** (2024) 186203 (cit. on p. 12).

[53] T. Borzenko, C. Gould, G. Schmidt and L. W. Molenkamp, *Metallic Air-Bridges Fabricated by Multiple Acceleration Voltage Electron Beam Lithography*, Microelectronic Engineering **75** (2004) 210, ISSN: 0167-9317 (cit. on p. 12).

[54] E. Boere, E. van der Drift, J. Romijn and B. Rousseeuw, *Experimental Study on Proximity Effects in High Voltage E-Beam Lithography*, Microelectronic Engineering **11** (1990) 351, ISSN: 0167-9317 (cit. on p. 12).

[55] B. A. Miller-Chou and J. L. Koenig, *A Review of Polymer Dissolution*, Progress in Polymer Science **28** (2003) 1223, ISSN: 00796700 (cit. on p. 15).

[56] A. Kokkinis et al., *Molecular Weight and Processing Effects on the Dissolution Properties of Thin Poly(Methyl Methacrylate) Films*, Microelectronic Engineering **85** (2008) 93, ISSN: 01679317 (cit. on p. 15).

[57] Y. Sakurabayashi, T. Masaki, T. Iwao and M. Yumoto, *Surface Hardness Improvement of PMMA by Low-Energy Ion Irradiation and Electron Irradiation*, Electronics and Communications in Japan **94** (2011) 19, ISSN: 1942-9541 (cit. on p. 15).

[58] F. T. Z. Toma, M. S. Rahman, K. M. A. Hussain and S. Ahmed, *Thin Film Deposition Techniques: A Comprehensive Review*, https://article.innovationforever.com/JMN/20240268.html (2024) (cit. on p. 17).

[59] D. W. Pohl, W. Denk and M. Lanz, *Optical Stethoscopy: Image Recording with Resolution λ/20*, Applied Physics Letters **44** (1984) 651, ISSN: 0003-6951 (cit. on p. 27).

[60] E. Betzig and J. K. Trautman, *Near-Field Optics: Microscopy, Spectroscopy, and Surface Modification Beyond the Diffraction Limit*, Science **257** (1992) 189 (cit. on p. 27).

[61] L. Novotny, E. J. Sánchez and X. S. Xie, *Near-Field Optical Imaging Using Metal Tips Illuminated by Higher-Order Hermite–Gaussian Beams*, Ultramicroscopy **71** (1998) 21, ISSN: 0304-3991 (cit. on p. 27).

[62] G. Binnig, H. Rohrer, C. Gerber and E. Weibel, *Surface Studies by Scanning Tunneling Microscopy*, Physical Review Letters **49** (1982) 57, ISSN: 0031-9007 (cit. on p. 28).

[63] *The Nobel Prize in Physics 1986*, URL: `https://www.nobelprize.org/prizes/physics/1986/summary/` (cit. on p. 28).

[64] G. Binnig, C. F. Quate and C. Gerber, *Atomic Force Microscope*, Phys. Rev. Lett. **56** (9 1986) 930 (cit. on p. 28).

[65] B. Voigtländer, *Atomic Force Microscopy*, 2nd, Springer Nature Switzerland AG, 2019, ISBN: 978-3-030-13653-6 (cit. on p. 28).

[66] S. Maghsoudy-Louyeh, M. Kropf and B. R. Tittmann, *Review of Progress in Atomic Force Microscopy*, The Open Neuroimaging Journal (2018) (cit. on p. 28).

[67] J. M. Larson, H. A. Bechtel and R. Kostecki, *Detection and Signal Processing for Near-Field Nanoscale Fourier Transform Infrared Spectroscopy*, Advanced Functional Materials **34** (2024) 2406643, ISSN: 1616-301X, 1616-3028 (cit. on pp. 31, 36).

[68] S. Mastel, A. A. Govyadinov, T. V. A. G. de Oliveira, I. Amenabar and R. Hillenbrand, *Nanoscale-Resolved Chemical Identification of Thin Organic Films Using Infrared near-Field Spectroscopy and Standard Fourier Transform Infrared References*, Applied Physics Letters **106** (2015) 023113, ISSN: 0003-6951 (cit. on p. 31).

[69] A. Hartschuh, N. Anderson and L. Novotny, *Near-Field Raman Spectroscopy Using a Sharp Metal Tip*, Journal of Microscopy **210** (2003) 234, ISSN: 1365-2818 (cit. on p. 31).

[70] J. Z. Lapin, R. Beams, L. G. Cançado and L. Novotny, *Near-Field Raman Spectroscopy of Nanocarbon Materials*, Faraday Discussions **184** (2015) 193 (cit. on p. 31).

[71] Y. Lee et al., *Characterization of the Structural Defects in CVD-grown Monolayered MoS 2 Using near-Field Photoluminescence Imaging*, Nanoscale **7** (2015) 11909 (cit. on p. 31).

[72] R. Hillenbrand and F. Keilmann, *Complex Optical Constants on a Subwavelength Scale*, Physical Review Letters **85** (2000) 3029, ISSN: 0031-9007, 1079-7114 (cit. on pp. 31, 40).

[73] M. Schnell et al., *Amplitude- and Phase-Resolved Near-Field Mapping of Infrared Antenna Modes by Transmission-Mode Scattering-Type Near-Field Microscopy*, The Journal of Physical Chemistry C **114** (2010) 7341, ISSN: 1932-7447 (cit. on p. 31).

[74] A. V. Ermushev, B. V. Mchedlishvili, V. A. Oleĭnikov and A. V. Petukhov, *Surface Enhancement of Local Optical Fields and the Lightning-Rod Effect*, Quantum Electronics **23** (1993) 435, ISSN: 1063-7818 (cit. on pp. 33, 46).

[75] N. Ocelic, A. Huber and R. Hillenbrand, *Pseudoheterodyne Detection for Background-Free near-Field Spectroscopy*, Applied Physics Letters **89** (2006) 101124, ISSN: 0003-6951 (cit. on pp. 36, 40).

[76] N. Ocelic, *Quantitative Near-field Phonon-polariton Spectroscopy*, PhD thesis: Technischen Universität München, 2007, URL: `https://mengkun.weebly.com/uploads/3/9/6/5/39652852/5._ocelic-phd_thesis-final.pdf` (cit. on pp. 36, 43).

[77] L. Mester, *Substrate-enhanced and subsurface infrared near-field spectroscopy of organic layers*, PhD thesis: CIC nanoGUNE, 2020, URL: `https://www.nanogune.eu/en/research/nanooptics/publications/phd-theses` (cit. on pp. 38, 47).

[78] L. Gomez et al., *Apertureless Scanning Near-Field Optical Microscopy: A Comparison between Homodyne and Heterodyne Approaches*, Journal of the Optical Society of America B **23** (2006) 823, ISSN: 0740-3224, 1520-8540 (cit. on p. 40).

[79] T. Taubner, R. Hillenbrand and F. Keilmann, *Performance of Visible and Mid-Infrared Scattering-Type near-Field Optical Microscopes*, Journal of Microscopy **210** (2003) 311, ISSN: 1365-2818 (cit. on p. 46).

[80] B. Hecht, H. Bielefeldt, Y. Inouye, D. W. Pohl and L. Novotny, *Facts and Artifacts in Near-Field Optical Microscopy*, Journal of Applied Physics **81** (1997) 2492, ISSN: 0021-8979 (cit. on p. 46).

[81] G. Kaupp, A. Herrmann and M. Haak, *Artifacts in Scanning Near-Field Optical Microscopy (SNOM) Due to Deficient Tips*, Journal of Physical Organic Chemistry **12** (1999) 797, ISSN: 1099-1395 (cit. on p. 46).

[82] V. E. Babicheva, S. Gamage, M. I. Stockman and Y. Abate, *Near-Field Edge Fringes at Sharp Material Boundaries*, Optics Express **25** (2017) 23935, ISSN: 1094-4087 (cit. on p. 46).

[83] B. Knoll and F. Keilmann, *Enhanced dielectric contrast in scattering-type scanning near-field optical microscopy*, Optics Communications **182** (2000) 321 (cit. on p. 46).

[84] A. Cvitkovic, N. Ocelic and R. Hillenbrand, *Analytical model for quantitative prediction of material contrasts in scattering-type near-field optical microscopy*, Optics Express **15** (2007) 8550, ISSN: 1094-4087 (cit. on pp. 46, 50).

[85] T. Vincent et al., *snompy: a package for modelling scattering-type scanning near-field optical microscopy*, 2024, arXiv: `2405.20948 [cond-mat.mtrl-sci]` (cit. on pp. 48, 50).

[86] W. S. M. Werner, K. Glantschnig and C. Ambrosch-Draxl, *Optical Constants and Inelastic Electron-Scattering Data for 17 Elemental Metals*, Journal of Physical and Chemical Reference Data **38** (2009) 1013, ISSN: 0047-2689 (cit. on p. 49).

[87] N. Sultanova, S. Kasarova and I. Nikolov, *Dispersion properties of optical polymers*, ACTA PHYSICA POLONICA A **116** (2009) (cit. on pp. 49, 74).

[88] K. Y. Bliokh, F. J. Rodríguez-Fortuño, F. Nori and A. V. Zayats, *Spin–orbit interactions of light*, Nature Photonics **9** (2015) 796 (cit. on p. 57).

[89] H. Mathur, *Thomas precession, spin-orbit interaction, and Berry's phase*, Physical Review Letters **67** (1991) 3325 (cit. on p. 57).

[90] A. Bérard and H. Mohrbach, *Spin Hall effect and Berry phase of spinning particles*, Physics Letters A **352** (2006) 190, ISSN: 0375-9601 (cit. on p. 57).

[91] E. I. Rashba, *Spin–orbit coupling and spin transport*,
Physica E: Low-dimensional Systems and Nanostructures **34** (2006) 31,
ISSN: 1386-9477 (cit. on p. 57).

[92] J. Fabian, A. Matos-Abiague, C. Ertler, P. Stano and I. Zutic, *Semiconductor Spintronics*,
Acta Physica Slovaca. Reviews and Tutorials **57** (2007) (cit. on p. 57).

[93] F. I. Fedorov, *To the theory of total reflection**,
Journal of Optics **15** (2013) 014002, Publisher: IOP Publishing, ISSN: 2040-8986
(cit. on p. 57).

[94] C. Imbert, *Calculation and Experimental Proof of the Transverse Shift Induced by Total Internal Reflection of a Circularly Polarized Light Beam*, Physical Review D **5** (1972) 787,
ISSN: 0556-2821 (cit. on p. 57).

[95] F. Pillon, H. Gilles and S. Girard, *Experimental observation of the Imbert–Fedorov transverse displacement after a single total reflection*, Appl. Opt. **43** (2004) 1863
(cit. on p. 57).

[96] F. Goos and H. Hänchen, *Ein neuer und fundamentaler Versuch zur Totalreflexion*,
Annalen der Physik **436** (1947) 333, _eprint:
https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19474360704,
ISSN: 1521-3889 (cit. on p. 57).

[97] Y. Zhao, J. S. Edgar, G. D. M. Jeffries, D. McGloin and D. T. Chiu,
Spin-to-Orbital Angular Momentum Conversion in a Strongly Focused Optical Beam,
Physical Review Letters **99** (2007) 073901, ISSN: 0031-9007, 1079-7114 (cit. on p. 57).

[98] K. Y. Bliokh et al.,
Spin-to-orbital angular momentum conversion in focusing, scattering, and imaging systems,
Optics Express **19** (2011) 26132, ISSN: 1094-4087 (cit. on p. 57).

[99] F. J. Rodríguez-Fortuño, N. Engheta, A. Martínez and A. V. Zayats,
Lateral forces on circularly polarizable particles near a surface,
Nature Communications **6** (2015) 8799, ISSN: 2041-1723 (cit. on p. 58).

[100] J. Petersen, J. Volz and A. Rauschenbeutel,
Chiral nanophotonic waveguide interface based on spin-orbit interaction of light,
Science **346** (2014) 67, ISSN: 0036-8075, 1095-9203 (cit. on p. 58).

[101] V. A. Zenin, R. Malureanu, I. P. Radko, A. V. Lavrinenko and S. I. Bozhevolnyi,
Near-field characterization of bound plasmonic modes in metal strip waveguides,
Opt. Express **24** (2016) 4582 (cit. on pp. 60, 73, 80, 84).

[102] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw and J. P. Woerdman, *Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes*,
Physical Review A **45** (1992) 8185, ISSN: 1050-2947, 1094-1622 (cit. on p. 70).

[103] G. Veronis and S. Fan,
Guided subwavelength plasmonic mode supported by a slot in a thin metal film,
Opt. Lett. **30** (2005) 3359 (cit. on p. 73).

[104] A. Andryieuski et al.,
Direct Characterization of Plasmonic Slot Waveguides and Nanocouplers,
Nano Letters **14** (2014) 3925 (cit. on p. 73).

[105] F. Bleckmann, Z. Cherpakova, S. Linden and A. Alberti,
Spectral imaging of topological edge states in plasmonic waveguide arrays,
Phys. Rev. B **96** (4 2017) 045417 (cit. on pp. 73, 112, 114, 120, 121).

[106] B. Steinberger et al.,
Dielectric stripes on gold as surface plasmon waveguides: Bends and directional couplers,
Applied Physics Letters **91** (2007) 081111 (cit. on p. 73).

[107] A. V. Krasavin and A. V. Zayats, *Three-dimensional numerical modeling of photonic integration with dielectric-loaded SPP waveguides*, Physical Review B **78** (2008) 045425 (cit. on pp. 73, 92, 101).

[108] T. Holmgaard et al.,
Bend- and splitting loss of dielectric-loaded surface plasmon-polariton waveguides,
Optics Express **16** (2008) 13585 (cit. on p. 73).

[109] M. Sumetsky, *Nanophotonics of Optical Fibers*, Nanophotonics **2** (2013) 393,
ISSN: 2192-8614 (cit. on p. 76).

[110] E. Verhagen, M. Spasenović, A. Polman and L. (Kuipers,
Nanowire Plasmon Excitation by Adiabatic Mode Transformation,
Physical Review Letters **102** (2009) 203904 (cit. on p. 80).

[111] V. A. Zenin et al., *Boosting Local Field Enhancement by on-Chip Nanofocusing and Impedance-Matched Plasmonic Antennas*, Nano Letters **15** (2015) 8148 (cit. on p. 80).

[112] Y. Chen et al.,
Efficient interfacing photonic and long-range dielectric-loaded plasmonic waveguides,
Opt. Express **23** (2015) 9100 (cit. on p. 84).

[113] V. A. Zenin et al., *Hybrid plasmonic waveguides formed by metal coating of dielectric ridges*,
Optics Express **25** (2017) 12295 (cit. on p. 84).

[114] F. Bleckmann, *Controlling surface plasmon polaritons with dielectric nanostructures*,
PhD thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2016,
URL: `https://hdl.handle.net/20.500.11811/6947` (cit. on p. 85).

[115] Z. Fedorova, *Simulation of time-periodic and topological tight-binding systems with plasmonic waveguide arrays*,
PhD thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2022,
URL: `https://hdl.handle.net/20.500.11811/10259` (cit. on p. 85).

[116] T. Taubner, F. Keilmann and R. Hillenbrand,
Nanomechanical Resonance Tuning and Phase Effects in Optical Near-Field Interaction,
Nano Letters **4** (2004) 1669 (cit. on pp. 87, 89).

[117] R. Krutokhvostov et al., *Enhanced resolution in subsurface near-field optical microscopy*, Opt. Express **20** (2012) 593 (cit. on p. 88).

[118] A. P. Engelhardt, B. Hauer and T. Taubner, *Visibility of weak contrasts in subsurface scattering near-field microscopy*, Ultramicroscopy **126** (2013) 40 (cit. on p. 88).

[119] A. A. Govyadinov et al., *Recovery of Permittivity and Depth from Near-Field Data as a Step toward Infrared Nanotomography*, ACS Nano **8** (2014) 6911 (cit. on p. 88).

[120] S. Zhang et al., *Nano-spectroscopy of excitons in atomically thin transition metal dichalcogenides*, Nature Communications **13** (2022) 542 (cit. on p. 89).

[121] S. Jetté-Charbonneau, R. Charbonneau, N. Lahoud, G. Mattiussi and P. Berini, *Demonstration of Bragg gratings based on long ranging surface plasmon polariton waveguides*, Opt. Express **13** (2005) 4674 (cit. on p. 94).

[122] J.-C. Weeber et al., *Near-field characterization of Bragg mirrors engraved in surface plasmon waveguides*, Phys. Rev. B **70** (23 2004) 235406 (cit. on p. 94).

[123] M. Ota et al., *Dielectric-loaded surface plasmon polariton crossing waveguides using multimode interference*, Optics Letters **40** (2015) 2269 (cit. on p. 101).

[124] G. Yuan, P. Wang, Y. Lu and H. Ming, *Multimode interference splitter based on dielectric-loaded surface plasmon polariton waveguides*, Optics Express **17** (2009) 12594 (cit. on p. 101).

[125] Z. Chen et al., *Wavelength-selective directional coupling with dielectric-loaded plasmonic waveguides*, Optics Letters **34** (2009) 310 (cit. on p. 101).

[126] T. Holmgaard et al., *Wavelength selection by dielectric-loaded plasmonic components*, Applied Physics Letters **94** (2009) 051111 (cit. on p. 101).

[127] S. Randhawa et al., *Performance of electro-optical plasmonic ring resonators at telecom wavelengths*, Optics Express **20** (2012) 2354 (cit. on p. 101).

[128] J. Gosciniak et al., *Thermo-optic control of dielectric-loaded plasmonic waveguide components*, Optics Express **18** (2010) 1207 (cit. on p. 101).

[129] R. M. Briggs, J. Grandidier, S. P. Burgos, E. Feigenbaum and H. A. Atwater, *Efficient Coupling between Dielectric-Loaded Plasmonic and Silicon Photonic Waveguides*, Nano Letters **10** (2010) 4851 (cit. on p. 101).

[130] A. Ma, Y. Li and X. Zhang, *Coupled Mode Theory for Surface Plasmon Polariton Waveguides*, Plasmonics **8** (2013) 769 (cit. on p. 102).

[131] Y. Suematsu and K. Kishino,
Coupling coefficient in strongly coupled dielectric waveguides, Radio Science **12** (1977) 587
(cit. on p. 102).

[132] A. W. Snyder, *Coupled-Mode Theory for Optical Fibers*,
Journal of the Optical Society of America **62** (1972) 1267 (cit. on p. 103).

[133] J. K. Asbóth, L. Oroszlány and A. Pályi, *A Short Course on Topological Insulators: Band-Structure Topology and Edge States in One and Two Dimensions*, vol. 919,
Lecture Notes in Physics, Cham: Springer International Publishing, 2016
(cit. on pp. 112, 115–117, 119).

[134] H. Schomerus, *Topologically Protected Midgap States in Complex Photonic Lattices*,
Optics Letters **38** (2013) 1912, ISSN: 0146-9592, 1539-4794 (cit. on p. 112).

[135] L. Li, Z. Xu and S. Chen, *Topological Phases of Generalized Su-Schrieffer-Heeger Models*,
Physical Review B **89** (2014) 085111 (cit. on p. 112).

[136] E. J. Meier, F. A. An and B. Gadway,
Observation of the Topological Soliton State in the Su–Schrieffer–Heeger Model,
Nature Communications **7** (2016) 13986, ISSN: 2041-1723 (cit. on p. 112).

[137] A. Block et al., *Bloch oscillations in plasmonic waveguide arrays*,
Nature Communications **5** (2014) 3843 (cit. on p. 113).

[138] D. H. Dunlap and V. M. Kenkre,
Dynamic localization of a charged particle moving under the influence of an electric field,
Physical Review B **34** (1986) 3625 (cit. on p. 113).

[139] P. Delplace, *Zak phase and the existence of edge states in graphene*,
Physical Review B **84** (2011) (cit. on pp. 115, 116).

[140] M. Prämassing, M. Liebtrau, H. J. Schill, S. Irsen and S. Linden,
Interferometric Near-Field Characterization of Plasmonic Slot Waveguides in Single- and Poly-Crystalline Gold Films, Optics Express **28** (2020) 12998, ISSN: 1094-4087
(cit. on p. 150).

[141] C. Schörner, S. Adhikari and M. Lippitz,
A Single-Crystalline Silver Plasmonic Circuit for Visible Quantum Emitters,
Nano Letters **19** (2019) 3238, ISSN: 1530-6984, 1530-6992 (cit. on p. 150).

[142] L. Liu et al.,
Atomically Smooth Single-Crystalline Platform for Low-Loss Plasmonic Nanocavities,
Nano Letters **22** (2022) 1786, ISSN: 1530-6984 (cit. on p. 150).

[143] T. Mori, T. Mori, Y. Tanaka, Y. Suzaki and K. Yamaguchi,
Fabrication of Single-Crystalline Plasmonic Nanostructures on Transparent and Flexible Amorphous Substrates, Scientific Reports **7** (2017) 42859, ISSN: 2045-2322
(cit. on p. 150).

[144] Y. Zhang, H.-J. Schill, S. Irsen and S. Linden, *Strong Coupling between WS2 Monolayer Excitons and a Hybrid Plasmon Polariton at Room Temperature*, Nanophotonics **13** (2024) 2847, ISSN: 2192-8614 (cit. on p. 150).

[145] C.-Y. Wang et al., *Giant Colloidal Silver Crystals for Low-Loss Linear and Nonlinear Plasmonics*, Nature Communications **6** (2015) 7734, ISSN: 2041-1723 (cit. on p. 150).

[146] M. Rycenga et al., *Controlling the Synthesis and Assembly of Silver Nanostructures for Plasmonic Applications*, Chemical Reviews **111** (2011) 3669, ISSN: 0009-2665 (cit. on p. 150).

[147] A. Kannegulla and L.-J. Cheng, *Metal Assisted Focused-Ion Beam Nanopatterning*, Nanotechnology **27** (2016) 36LT01, ISSN: 0957-4484 (cit. on p. 152).

[148] A. Kanwal, B. R. Ilic, C. H. Ray, K. Siebein and J. A. Liddle, *Chromium Oxide – A Novel Sacrificial Layer Material for MEMS/NEMS and Micro/Nanofluidic Device Fabrication*, Micro and Nano Engineering **16** (2022) 100145, ISSN: 2590-0072 (cit. on p. 152).

List of Figures

2.1 Dielectric function of gold and dispersion of SPPs at an air/gold interface. . . 6
2.2 Propagating spp sketch. 7
2.3 Sketch of the two used waveguide types. 8
2.4 Sketch of grating coupling. 9

3.1 Schematic of positive and negative tone lithography. 12
3.2 Gold alignment marker and disc array on glass substrate. 17
3.3 FIB milled aligment marker in gold film. 19
3.4 Grating coupler excitation sketch. 20
3.5 Fabrication of grating couplers sketch. 21
3.6 FIB milled incoupling gratings. 22
3.7 Design of the first step of the two step (positive tone) EBL process for DLSppWGs. 23
3.8 Example dose test of incoupling gratings for a wavelength of 970 nm with PMMA positive tone. 24
3.9 Example of an incoupling grating for a wavelength of 970 nm with PMMA positive tone. 25
3.10 Example of multiple incoupling gratings for a wavelength of 970 nm with PMMA positive tone. 25

4.1 Schematic of the forces acting on the tip in the different modes of operation of an AFM. 29
4.2 Schematic of the AFM setup. 30
4.3 AFM scan of TGQ1 test sample 31
4.4 Schematic of the s-SNOM setup. 33
4.5 Schematic of the tip emission directions in the transmission s-SNOM setup. . 34
4.6 Visualization of the lock-in detection principle. 38
4.7 Scattering amplitudes of the near-field, background and their interference. . . 39
4.8 Demodulated detector signal at the n-th harmonic for different background phases. 39
4.9 Schematic of the pseudo-heterodyne detection. 41
4.10 Numerical visualization of the demodulated detector signal. 42
4.11 Analytic result for the second harmonic of the tip oscillation frequency Ω. . . 44

4.12 Example of experimentally measured approach curves on a gold substrate, demonstrating the impact of pseudo-heterodyne detection. 44
4.13 Sketch of the tip-sample interaction. 45
4.14 Sketch of the tip-sample interaction and formation of the image dipole. 47
4.15 Example approach curves for a gold and PMMA sample, calculated with the PDM for the in- and out-of plane polarization. 49
4.16 Sample orientation for transmission mode measurements. 51
4.17 Sketch of the influence of the direction of wave propagation in transmission mode. 52
4.18 Transmission mode s-SNOM measurement of an isolated grating coupler. . . 53

5.1 Sketch of the excitation mechanism for the chiral couplers. 58
5.2 Design of the chiral couplers. 60
5.3 *Comsol* simulation of chiral couplers for different illumination polarizations. . 62
5.4 Numerical model of the interference of a waveguide mode with a dipolar excitation along the waveguide. 63
5.5 SEM images of the samples for different coupler radii. 64
5.6 Near-field measurements of the chiral coupler for different excitation conditions. 64
5.7 Measurements for different waveguide separations. 66
5.8 Measurements for different waveguide widths. 66
5.9 Measurements for different coupler radii. 67
5.10 Example cutlines of the averaged real part along two waveguides of a chiral coupler. 68
5.11 Comparison of the directionality calculated from simulations and measurements. 69
5.12 3D-printed adapter for vortex retarder plate. 70
5.13 Measurement of a disc array excited with a donut beam in transmission mode. 71
5.14 Measurement of a ring of discs excited with a donut beam in transmission mode. 72

6.1 Mode profile simulation of a DLSPPW. 75
6.2 Effective mode index and propagation length calculated from the simulations. 76
6.3 Ratio of the in- and out-of-plane electric field component above the waveguide. 77
6.4 AFM height profile of a dose test for PMMA negative tone lithography. 79
6.5 Height and width extracted from an AFM scan of a dose test for PMMA negative tone lithography. 79
6.6 Design of the second step of the two step positive tone EBL process for the fabrication of DLSPPWs. 81
6.7 Transmission mode s-SNOM measurement of a single DLSPPW at a wavelength of 970 nm. 83
6.8 s-SNOM measurement of a single DLSPPW with a tapered incoupler at 1 600 nm. 85

6.9 Effective mode index for different waveguide geometries, experimental and simulated values. 86
6.10 Phase map of two DLSPPWs, showing symmetric and asymmetric phase shifts. 86
6.11 Reflection mode s-SNOM measurement of a PMMA wedge on a gold film. . . 87
6.12 Experimental phase shift for different excitation wavelengths. 89
6.13 Indication of the in-plane component using an automated rotation of the analyzer. 91
6.14 Simulation of the reflection grating with Comsol. 93
6.15 Amplitude profiles on the waveguide from simulation. 94
6.16 Design of the reflection grating using positive tone EBL. 95
6.17 Example AFM scan of reflection gratings. 95
6.18 Large s-SNOM measurement of a DLSPPW with a reflection grating. 97
6.19 s-SNOM measurements of a reflection grating with different illumination polarizations and geometries. 98
6.20 Comparison of measurements and tip-interation adapted simulations. 99
6.21 Amplitude profiles from the measurements on the waveguide. 99

7.1 Simulated mode profiles of two coupled DLSPPWs. 103
7.2 s-SNOM measurements of two coupled DLSPP waveguides. 105
7.3 s-SNOM measurements of two coupled DLSPPWs with varying distance between the waveguides. 106
7.4 Fit procedure of height and amplitude data for two coupled DLSPP waveguides.107
7.5 Extracted amplitudes of two coupled DLSPP waveguides using two methods. 108
7.6 Real part of the electric field extracted from s-SNOM measurements of two coupled DLSPPWs. 109
7.7 Comparison of the beating pattern of the measurements to the numerical model using the simulation results. 110
7.8 Fourier transform of the beating to extract the mode indices. 110
7.9 Comparison of the fitted numerical model to the real part of the two waveguides.111
7.10 Phase difference extracted from the measurements and fitted numerical model.112
7.11 Sketch of the tight-binding model. 113
7.12 Sketch of the one dimensional chain of atoms of the SSH model. 114
7.13 Energy eigenvalues of a finite SSH bulk and corresponding eigenstates. . . . 117
7.14 Energy bands and winding of the vector $\vec{d}(k)$ for the two different dimerization configurations of the SSH model. 118
7.15 Eigenvalues for a finite SSH model with a domain wall and eigenstate of the zero energy edge state. 119
7.16 Numerical simulation of the spreading of bulk modes and the edge state. . . 120
7.17 Schematic representation of a topological interface in the SSH model. 121
7.18 Large s-SNOM measurement of a coupled waveguide array in the topological configuration. 122
7.19 Example of a cut-line measurement at a distance of 60 µm from the incoupler. 123

7.20 s-SNOM amplitude cut-line measurements of two coupled waveguide arrays (topological, trivial). 125
7.21 Comparison of the expected edge mode distribution with the experimental results. 126

A.1 Sketch of the tip emission and reflection on the upper parabola. 132
A.2 Projection of the polarization vector onto the x'-y'-z'-coordinate system. . . . 133
A.3 Sketch of the relations of the angles and pathways. 134
A.4 Mode profile simulations of DLSPPWs with different geometries at 970 nm . 136
A.5 Mode profile simulations of DLSPPWs with different geometries at 1 500 nm . 137
A.6 Confinement factors for different waveguide widths and heights at 970 nm . . 138
A.7 Confinement factors for different waveguide widths and heights at 1 500 nm . 139
A.8 Comparison of the measurements and tip-interaction adapted simulations of the reflection grating. 140
A.9 Polarizability corrected simulation data using the PDM. 141
A.10 Polarizability corrected simulation data using the PDM adapted for the transmission mode. 142
A.11 Three dimensional s-SNOM measurement of a DLSPPW at 970 nm. 144
A.12 Example of an array of fragmented single waveguides fabricated with PMMA negative tone lithography. 145
A.13 Example AFM scans of grainy waveguides fabricated with PMMA negative tone lithography. 146
A.14 Example of a waveguide fabricated with PMMA negative tone lithography on a glass substrate. 146
A.15 Example AFM scans of waveguides fabricated with different acceleration voltages. 147
A.16 Example of coupled waveguides fabricated with PMMA negative tone lithography showing proximity effects. 148
A.17 Example of coupled waveguides fabricated with PMMA positive tone lithography. 148
A.18 Example AFM profiles of a SSH waveguide array fabricated with positive tone EBL. 149
A.19 AFM roughness scan of the silver crystal surface. 152
A.20 SEM images of nanogrooves in silver crystals. 153
A.21 SEM image of a nanogroove dosetest. 153

Glossary

AC alternating current

AFM atomic force microscopy

CMT coupled mode theory

DC direct current

DLSPP dielectric-loaded surface plasmon polariton

DLSPPW dielectric-loaded surface plasmon polariton waveguide

EOAM extrinsic orbital angular momentum

EBL electron beam lithography

FDM finite dipole model

FEM finite element method

FIB focused ion beam

IOAM intrinsic orbital angular momentum

ITO indium tin oxide

KOH potassium hydroxide

LRM leakage radiation microscopy

MIBK methyl isobutyl ketone

NA numerical aperture

ND neutral density

NIR near-infrared

NMP n-methyl-2-pyrrolidone

OAM orbital angular momentum

PDM point dipole model

PDMS polydimethylsiloxane

PID proportional-integral-differential

PML perfectly matched layer

PMMA poly-methyl-methacrylate

SAM spin angular momentum

SNOM scanning near-field optical microscopy

SEM scanning electron microscope

SOI spin-orbit interaction

SSH Su-Schrieffer-Heeger

s-SNOM scattering-type scanning near-field optical microscope

SPP surface plasmon polariton

SNR signal-to-noise ratio

STM scanning tunneling microscopy

TM transverse magnetic

TMDC transition metal dichalcogenide

VIS visible

VPP vortex phase plate

Acknowledgements

I would like to express my sincere appreciation to all those who supported me during the course of this work, as well as to the members of the examination committee for the evaluation of this thesis.

First and foremost, I thank my first supervisor, Prof. Stefan Linden, for giving me the opportunity to work as a doctoral student and for his continuous support and encouragement. Our many fruitful scientific discussions greatly contributed to my intuitive understanding of near-field microscopy and nanophotonics. I especially fondly remember the annual barbecues and hope to join them again in the future.

Secondly, I thank Priv.-Doz. Elisabeth Soergel for agreeing to act as the second referee. I always enjoyed tutoring her seminar and the occasional drink together in the beer garden. I also thank Prof. Corinna Kollath and Prof. Moritz Sokolowski for serving on the examination committee.

I would like to acknowledge all current and former members of my research group, whether for the pleasant working atmosphere, stimulating scientific discussions, or shared social activities. In particular, I thank Mike Prämassing, who introduced me to near-field microscopy early on, when I was still a bachelor student.

I am grateful to Stephan Irsen and all colleagues at Caesar for granting me access to the focused ion beam system and for their support during my work there. I always greatly enjoyed my visits.

I would also like to thank the many proofreaders of this thesis, especially Lukas Rieland, Paul Steinmann, Max Wegerhoff, Gian-Marco Schnüriger, Felix Begemann, and Anna Sidorenko. I deeply appreciate their comments and suggestions, which significantly improved the quality and clarity of this thesis.

Finally, I would like to express my deepest gratitude to my family and close friends for their constant emotional support throughout this journey.