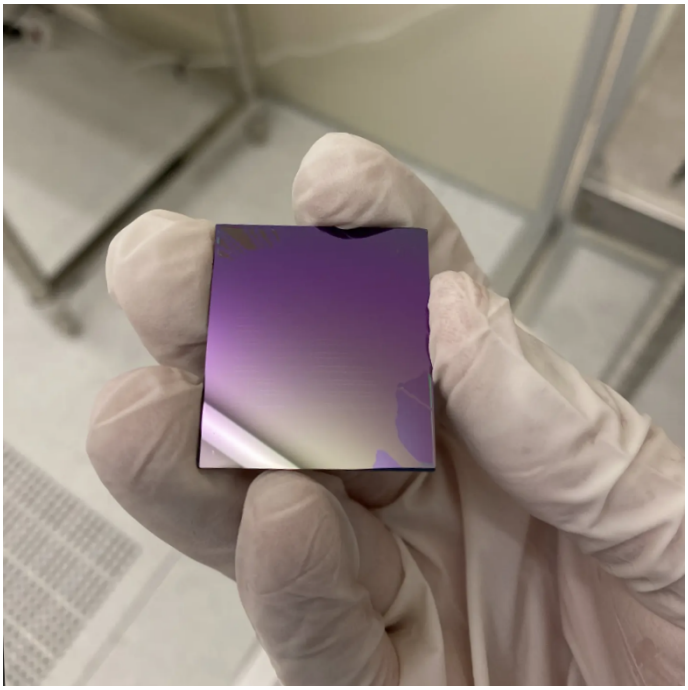
Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

Joakim Ekern

Development of a Nanophotonic Silicon Nitride Waveguide Sensor Platform for Use in SERS

Master's thesis in Nanotechnology
Supervisor: Astrid Aksnes
June 2023



NTNU
Norwegian University of
Science and Technology

Joakim Ekern

Development of a Nanophotonic Silicon Nitride Waveguide Sensor Platform for Use in SERS

Master's thesis in Nanotechnology
Supervisor: Astrid Aksnes
June 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Problem Description

In this master's thesis, slot waveguides will be investigated for use in Raman sensing. The waveguides are aimed at taking advantage of surface-enhanced Raman scattering (SERS) to amplify the Raman signal from a molecule of interest. The fabricated waveguides will consist of a silicon nitride (Si_3N_4) core on top of an insulator (SiO_2). The goal of the thesis is to develop and fabricate a waveguide platform from scratch.

Finite-difference eigenmode (FDE) computer simulations will be employed to predict the performance of the SiN waveguide, based on its cross-section. The goal is to determine the ideal cross-section design in order to achieve a large Raman enhancement. A range of different waveguide dimensions will be simulated, including the addition of a plasmonic gold layer, resulting in a nanoplasmonic slot waveguide platform.

The results of the computer simulations will be used to design and fabricate SiN slot waveguides at the NTNU NorFab cleanroom facilities. The waveguides will be patterned by employing electron beam lithography (EBL) and inductively coupled plasma-reactive ion etch (ICP-RIE), and subsequently characterised with a scanning electron microscope (SEM). The fabrication parameters (e.g. of the EBL and ICP-RIE) will be tuned to produce the desired SiN slot waveguides.

A Raman measurement lab environment will be constructed to test the waveguide's performance. A single mode 785 nm laser will be used to excite the waveguide samples, and the resulting Raman spectrum will be collected by a spectrometer. The waveguides will be tested with and without isopropyl alcohol (IPA). The goal is to determine the waveguide's ability to enhance the Raman signal from the IPA, which is a measure of the waveguide platform's performance.

List of Abbreviations

ALD atomic layer deposition

DUVL deep UV lithography

EBL electron beam lithography

EME eigenmode expansion

FDE finite-difference eigenmode

FDM finite-difference method

FDTD finite-difference time domain

FOM figure of merit

HIC high index contrast

ICP-RIE inductively coupled plasma-reactive ion etch

IPA isopropyl alcohol

LPCVD low pressure chemical vapour deposition

LSPR localized surface-plasmon resonance

MLA maskless aligner

NTNU Norwegian University of Science and Technology

NTP nitrothiophenol

PECVD plasma enhcanced chemical vapour deposition

PML perfectly matched layer

SEM scanning electron microscope

SERS surface-enhanced Raman scattering

SiN silicon nitride

SNR signal-to-noise ratio

SPP surface plasmon-polariton

List of Contents

Abstract

Surface-enhanced Raman scattering (SERS) is a powerful and label-free sensing technique which relies on probing the vibrational states of a molecule. This results in a unique fingerprint spectrum from the particular molecule, in the from of Stokes shifted light. Raman spectroscopy is limited by the small scattering-cross section of biological molecules, and free-space Raman microscopes are large and expensive. An alternative to free-space microscopes are waveguide sensors taking advantage of SERS. Such waveguides were the topic of this thesis. In a waveguide, the interaction length with the analyte can be increased, compared to free-space. With maturing technology, a waveguide-based sensing platform has the prospects of replacing conventional free-space Raman microscopes. This is due to an increased interaction length with the analyte, and the possibility of moving to a photonic sensing system fully integrated on a chip. Such a shift is analogous of the hugely beneficial shift towards integrated circuits in the electronics industry.

A silicon nitride slot waveguide with a 150 nm narrow gap was investigated in this thesis. Finite-difference eigenmode (FDE) calculations were employed to investigated the value of the Raman conversion efficiency η, which is proportional to the ratio of Stokes shifted power to pump power in a waveguide $P_s/P_p \propto \eta$. For the selected SiN waveguide platform, η yielded a maximum value of on the order of 10^4 for the fundamental TE mode. The silicon nitride waveguides were successfully fabricated by means of electron beam lithography (EBL) and inductively coupled plasma-reactive ion etch (ICP-RIE). Process optimization on the in-house EBL-system, identified an area dose of 170 $\mu C/cm^2$ at a beam current of 2 nA to be ideal for the purposes of this thesis. The waveguides were fabricated on SiN-on-insulator wafers provided by SINTEF.

The efforts in this thesis were concluded by the construction of a Raman measurement setup. The Raman excitation laser was a fiber-coupled 785 nm single mode laser. The waveguide's Raman spectra were collected, both with and without isopropyl alcohol (IPA). All of the resulting spectra were noisy, and there were no distinguishable differences between the measured background spectra and the IPA spectra.

SAMMENDRAG

Overflateforsterket Ramanspredning (SERS) er en kraftfull og ikke-ødeleggende sensorteknikk basert på å eksitere vibrasjonsnivåene til et molekyl. Dette resulterer i et unikt emisjonsspekter — et slags molekykært fingeravtrykk — som følge av at molekylet avgir Stokesspredt lys når det absorberer lys fra en eksitasjonslaser. To begrensende faktorer for Ramanspektroskopi er at biologiske molekyler har et lite spredningstverrsnitt, og at Ramanmikroskoper ofte er store og dyre. Bruk av bølgeledersensorer er en potensiell løsning på disse utfordringene. Da vil interaksjonlengden mellom en prøve og en lysstråle kunne økes. Dersom et fullverdig optisk sensorsystem kan integreres på en brikke, kan det erstatte de tradisjonelle Raman mikroskopene. Det vil være en analogi til den enorme nytteverdien elektronikkbransjen har hatt av å gå over til å bruke integrerte kretser på små brikker.

Denne masteroppgaven fokuserte på å undersøke en kløftbølgeleder av silisumnitrid med en 150 nm bred kløft. Teoretiske beregninger av bølgelederens virkningsgrad η ble utført ved hjelp av endelige differenser av eigenmoder-modellen (FDE). Virkningsgraden er en størrelse som er proporsjonal med forholdet mellom Stokesspredt lys og laserpumpen i bølgelederen, etter $P_s/P_p \propto \eta$. Den fundamentale TE moden i denne oppgavens bølgeleder, ble beregnet til å ha en maksimal virkningsgrad i størrelsesorden 10^4. Bølgelederen ble fabrikert gjennom elektronstrålelitografi (EBL) og spenningsledet ioneplasmaetsing (ICP-RIE). En parameteroptimalisering ble gjennomført på EBL-en som ble brukt, og den optimale arealdosen ble undersøkt til å være 170 $\mu C/cm^2$, med en strålesrøm på 2 nA. Bølgelederne ble fabrikert på silisiumnitrid-på-glass-brikker produsert av SINTEF.

Oppgaven ble fullført ved å bygge og teste et optisk Ramanspektroskopi-oppsett. Eksitasjonslaseren som ble brukt var en fiberlaser som opererte i enkelmodus, med en bølgelengde på 785 nm. Ramanspekteret fra flere forskjellige bølgeledere ble målt, både med og uten isopropanol (IPA). Formålet med dette var å måle både bølgeledernes bakgrunnsspekter, og det forsterkede Ramanspekteret fra IPA. Det var mye støy i spekterene som ble målt, som førte til at det ikke var mulig å se forskjell på målingene med og uten IPA.

PREFACE

This thesis expands upon the author's term project, which is an unpublished project written in the fall of 2022. Similarly to in this thesis, silicon nitride waveguides optimized for SERS sensing were considered. However, that project did not involve any fabrication, and slot waveguides were not considered. Therefore, the majority of this thesis is completely new. The chapters 1, 2, and 3 are reproduced here, but reworked to fit the objective of this thesis.

The topic of this thesis was partially proposed by SINTEF, which had no control of the project's development. To our knowledge, no research at NTNU has previously considered nanoplasmonic waveguides, or even waveguide-assisted Raman sensing, which are the topics of this thesis. This makes this project simultaneously enticing and challenging. Enticing because any reported results would constitute new research reported at NTNU, and challenging because large parts of the thesis had to be developed from scratch. The design and lithography process for the slot waveguides had to be developed, and the lab setup to measure Raman signal had to be built for the purposes of this thesis. The tuning of parameters necessary to achieve this, was done by the author. A large inspiration for this project was the work on waveguide-assisted SERS sensing reported by *Photonics Research Group* affiliated with Ghent University in Belgium.

About the Thesis

This thesis is organized in such a fashion that each chapter is should be a natural continuation of the last. The intention is to present relevant fundamental theory on photonics, then expand on this theory by reviewing recent developments in the field of SERS. Next, a simulation environment based on said theory is developed, and used to investigate a good waveguide design for SERS sensing. These simulation results are utilized to design a waveguide platform, which is subsequently fabricated at NTNU in the NorFab research facilities. Finally, the waveguides' performance are tested in an environment optimized for Raman measurement.

Some terms are used interchangeably. Silicon nitride, and the abbreviation SiN, both refer to the chemical compound Si_3N_4. Silicon dioxide, glass, and silica, all refer to the chemical compound SiO_2. The main waveguide structure considered in this thesis is a slot waveguide, which in other parts of the literature may be labelled a gap waveguide. For consistency, the term "gap waveguide" will not be utilized here. The slot, or gap, in the waveguide however, will be labelled "the gap".

This thesis was written in LaTeX, and the plots and figures were created with PGFPLOTS and TikZ. All figures and images were created by the author. On a sidenote, it should be mentioned that a *specral* colormap is commonly used in photonics plotting. This means that 2D surface plots are colored in such a fashion that the hue changes from blue at low values to red at high values, through the visible light spectrum. This thesis will employ another colormap, Matplotlib's *plasma* colormap. It is designed such that hue and brightness increment linearly with value, which is considered favorable since the human brain more easily interpret changes in intensity as changes in brightness, than hue. An explanation of the colormap and its purpose is found in the Matplotlib documentation [1].

Acknowledgements

A number of people deserve acknowledgement for contributing to the completion of this thesis. Firstly, Astrid Aksnes, my supervisor, who has been extremely helpful and supportive throughout this year, both during this master's thesis and the preliminary term project. I would also like to thank Jens Høvik, and the NTNU NanoLab staff in general, for their excellent help with the fabrication part of the thesis. In answering technical questions, and giving helpful advice, researchers Firehun T. Dullo and Karolina Milenko deserve a great thanks. Firehun kindly provided the wafers used for fabrication in this project. Next, I would like to thank all of my classmates, in particular Steffen Snippen Moe, Peter N. K. Sundgaard, and Petter Wiiker. Lastly, I would like to thank the people who helped me in correcting any errors in the final draft: Nikolai Stensø, Jan Ekern and Thale B. Mestvedt.

The fabrication part of this thesis was carried out in the NorFab cleanroom facilities at NTNU. NTNU NanoLab benefits from the support granted by the Norwegian Research Council (NRC).

Trondheim, June 2023

Joakim Ekern

Chapter 1

Introduction

Raman spectroscopy is a powerful tool in the analysis of any substance, since the fingerprint of specific molecules are revealed in the collected Raman spectrum. This fingerprint results from the scattering of light when molecules are illuminated, where the energy shift of an incident photon will correspond to a vibrational mode of the molecule. There is no need for labelling of the molecules, and since infrared light can be used to obtain a Raman signal, it is a non-destructive characterization technique. This is contrary to for instance X-ray spectroscopy, where high-energy photons will damage the sample. Although Raman spectroscopy is precise, it is limited by the small scattering cross-section of many biological molecules. Only one in a million photons is Raman scattered, making the Raman signal inherently weak, and hard to detect in the presence of noise and fluorescence. *Surface enhanced Raman spectroscopy* (SERS) produces an enhancement of the Raman signal, an effect which is attributed to the interaction of light with the conduction electrons of a metal. Two phenomena are responsible for this, namely surface plasmon-polariton (SPP) and localized surface-plasmon resonance (LSPR) [2]. A Raman signal enhancement of more than six orders of magnitude has been demonstrated, and the ultimate detection limit, capturing the Raman spectrum from a single molecule, was reached in 1997 [3].

In spite of the qualities of Raman spectroscopy, Raman microscopes are sparse, expensive, and large. Traditional microscopes operate in free-space, meaning that the laser excitation and the signal collection occur in air or vacuum. However, a new branch of SERS has emerged with the prospects of overcoming the disadvantages of free-space microscopes. This branch relies on waveguide based SERS platforms. Inspired by the enormous success of the electronics industry, in particular CMOS, a waveguide based SERS platform fully integrated on a chip is highly aspired. A waveguide has the ability to confine and direct the laser and Raman signal, and filtering and analysis of the signal can be done on a single chip. Additionally, the laser will be able to interact with molecules along the waveguide, increasing the interaction length between the laser and analytes. In complex samples, the signal can be multiplexed, such that multiple analyses can be performed in a single measurement.

However, such a platform imposes new challenges. The Raman signal from the waveguide itself, labelled the Raman background, introduces noise to the Raman signal from the analytes. Moreover, in order to confine light for use in Raman spectroscopy, the waveguides must have nanoscale dimensions. The same is true for the plasmonic structures, since LSPR is a nanoscale surface effect. This requires good control of nanoscale fabrication, if results are to be reliable and reproducible. Electron beam lithography (EBL) allows the creation of structures with a resolution of a few nanometers. However, EBL is not suited for mass-scale production, and EBL-free techniques are investigated. Recently, Turk *et al.* [4], were able to fabricate EBL-free nanoplasmonic slot waveguides approaching the performance of free-space SERS platforms.

The objective of this thesis is to develop, fabricate, and test a waveguide platform to be used in SERS. The next chapters will describe the work flow and creation of such a platform. Chapter 2 gives a short introduction of the most important concepts required to understand the operation of SERS in waveguides, such as basic electromagnetic theory, waveguide theory, and plasmon interactions. Chapter 3 expands on this theory, and considers recent developments in the field of SERS, which will influence the final waveguide design. In Chapter 4, a simulation environment employing the finite-difference eigenmode (FDE) method is designed. From this environment, a cross-sectional analysis of the sensing performance of a slot waveguide structure is presented. Chapter 5 takes advantage of these results, and a waveguide platform intended for fabrication and testing is designed. Chapter 6 realizes this design by means of EBL, and includes a justification of the most important fabrication parameters. The chapter also includes a description of the self-built Raman measurement setup which is utilized in performance testing. In Chapter 7 the fabricated waveguide design is tested by measuring the waveguide's Raman signal. Spectra were collected with and without isopropyl alcohol (IPA) deposited on the waveguides. Finally, in Chapter 8, the work presented in this thesis is shortly repeated, including the most important findings, and a possible continuation of the project is proposed.

Chapter 2

Fundamental Theory on Raman Scattering, Waveguides and Plasmons

Surface-enhanced Raman scattering (SERS) in waveguides is the topic of this thesis. It is an optical sensing technique, and relies on light in a waveguide interacting with molecules situated around its surface. Some of the light absorbed by the molecules is re-emitted as Raman Stokes scattered light of a longer wavelength. The energy difference between the incident and emitted light will correspond to the energy of one of the vibrational modes of the molecule. It is reliable and precise, because the light emitted from the analytes is specific to that molecule, like a molecular fingerprint. However, Raman scattering is inherently weak, only about one in a million photons incident on a molecule gets Raman scattered. This is a significant challenge, and some kind of enhancement of the scattered light is a prerequisite to enable SERS in waveguides as an analytical tool. A suitable enhancement effect is found in the *localized surface plasmon resonance* (LSPR) effect. In LSPR, electrons in metallic nanoscale sized structures oscillate in resonance with an electric field [2]. This resonance may produce a significant enhancement of the Raman signal, more than six orders of magnitude [5]. If the waveguide's surface is patterned with such plasmonic structures, a large enhancement of the Raman signal can be achieved.

It should be noted that SERS has been known since the 1970s, and it has previously been utilized in conventional free-space microscopes. The prospects of moving to a waveguide-based platform has the advantages of miniaturizing such a system, making it cheaper and simpler to use. When the technology is matured, an on-chip integrated photonic system for Raman spectroscopy can be realized. Previously, difficulty in controlling the morphology of nanostructures has limited the development of SERS, but advances in nanofabrication has increased control of structures on the nanoscale, paving the way for waveguide-assisted SERS.

This chapter presents the fundamental theory necessary to understand the topics discussed in this project. It should be considered an introduction, as lengthy derivations and complex explanations will be left out. More rigorous treatments can be found in the references and other sources. This chapter is intended to familiarize the reader with important concepts of photonics touched upon in this project. It should provide a basic understanding of scattering processes, the behavior of light in waveguides, and the behavior of plasmons in metal nanoparticles. Readers already familiar with these concepts may skip this chapter and go directly to Section 3, where SERS in waveguides is discussed in more detail.

In Section 2.1, some very basic principles of light propagation are stated, forming a platform for further theory. Section 2.2 briefly explains light scattering in general, and Raman scattering in particular. Section 2.3 continues with an introduction to light propagation in a guiding material, with focus on rectangular dielectric waveguides. This chapter closes with Section 2.4, an overview of the LSPR effect, which is the main effect responsible for SERS.

The basic theory presented in this chapter, as well as scattering and waveg-

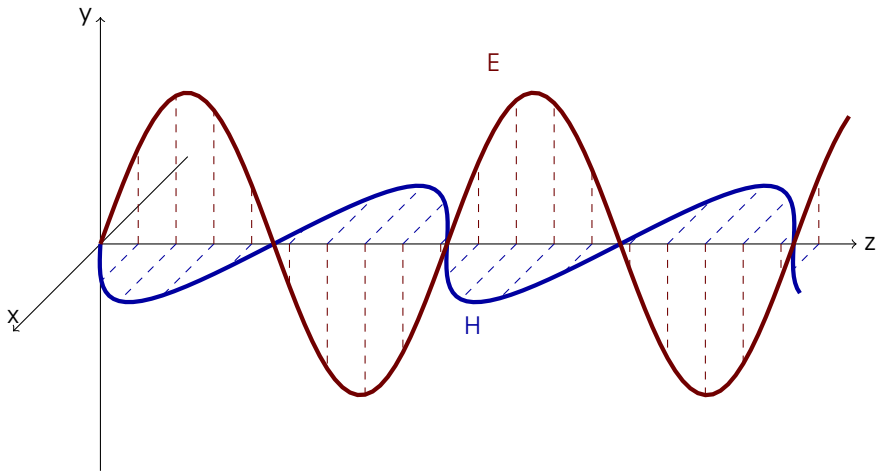


Figure 2.1: An electromagnetic light wave travelling along the z-direction. This wave is transverse electromagnetic (TEM), because the E and H fields oscillate linearly along y and x, respectively.

uides, is mainly adapted from Saleh & Teich [6]. The explanation of LSPR is mainly adapted from Stiles et al. [5]. The expressions for rectangular waveguides are taken from Yeh & Shimabukuro [7]. Pieter Wuytens' PhD thesis [8] on SERS and SiN waveguides has been used as a supplementary material to all topics. Please note that this chapter is reused from the term project leading up to this thesis, with only small corrections. The exception being Section 2.4, which has been extensively reworked.

2.1 An Electromagnetic Description of Light

That light possesses a wave-particle duality was first evident by Albert Einstein's discovery of the photoelectric effect. In most cases, a classical description of light as waves is sufficient, and Maxwell's equations describing the electromagnetic wave properties of light is the foundation of photonics. This means that a travelling light wave has an electric and a magnetic field which may interact with the surroundings, and it behaves like a wave, experiencing interference and diffraction. Such a wave is depicted in Figure 2.1. They are restated here as Equations (2.1)–(2.4) for reference and completeness, but the explanations of the terms and parameters are left out. The equations have the same form throughout the literature, so it should be straightforward to consult any good source on the topic to find a more detailed description of Maxwell's equations. The following expression of Maxwell's equations is for a source-free medium:

$$\nabla \times \mathcal{H} = \frac{\partial \mathcal{D}}{\partial t}, \tag{2.1}$$

$$\nabla \times \mathcal{E} = -\frac{\partial \mathcal{B}}{\partial t}, \tag{2.2}$$

$$\nabla \cdot \mathcal{D} = 0, \tag{2.3}$$

$$\nabla \cdot \mathcal{B} = 0, \tag{2.4}$$

where, all fields are real and time-dependent. $\mathcal{E}$ is the electric field, $\mathcal{D}$ is the electric displacement field, $\mathcal{H}$ is the magnetizing field, and $\mathcal{B}$ is the magnetic field.

Many problems may feature a small volume where the electromagnetic fields can be assumed to be constant. This is the case in the long wavelength limit, where the particle or volume of interest is at least one order of magnitude smaller than the wavelength of light. In these instances, Maxwell's equations can be replaced by the Laplace equation of electrostatics

$$\mathbf{E} = -\nabla V, \tag{2.5}$$

where V is the electric potential in volts.

Light is polarizable. A single photon will have a direction of propagation, and the E and H fields will oscillate normal to the direction of propagation, and to each other. In a Cartesian coordinate system, all the three parameters can be designated to a single axis, as depicted in Figure 2.1. This is labelled linear polarization. Let us only consider the E field in the following, since E and H are orthogonal anyway. Now, if a number of photons move along the same axis, their E fields may be oriented at any angle in the xy-plane. This will result in circular polarization. If the photons are out of phase as well, polarization is elliptical. This is depicted graphically in Figure 2.2.

The polarization of light is especially important in waveguides, which will be introduced in Section 2.3. In such a structure the picture is even more complicated, because waves move at an angle to the z-axis. However, either the E field or the H field will always be polarized in the xy-plane, making the wave either transverse electric (TE) or transverse magnetic (TM), respectively.

2.2 Raman Scattering

Scattering is one of the most important phenomenon that govern the interaction of light with particles. Here, different scattering processes, as well as Raman scattering, will be described conceptually. The goal is to briefly explain Raman scattering mechanisms, and why it is such a precise characterization technique. The processes are shown in Figure 2.3. In a scattering process, a photon incident on a molecule may scatter, so that it shifts its wavelength and/or direction of propagation. The wavelength of the scattered photon may be shorter, equal to, or longer than the incident photon. The latter being most likely — that some of the photon's energy is transferred to the molecule. The photon may change its direction of propagation as well. Depending on various parameters, all different directions of emission may have different probabilities.

As previously mentioned, Raman scattering is an inherently weak scattering process, and other types of scattering dominate. A scattering process works in the following way: Molecules will have an energy band gap between filled and empty electronic states. Incident photons may excite a particular electron to an energy level above the band gap. At some point, a photon with a shorter wavelength will be emitted when the electron relaxes from the excited energy state, back to the ground state. This scattering is fluorescence. In some cases, an incident photon may not have enough energy to excite an electron above the band gap. Instead, the electron may undergo an intraband transition, from one energy level below the band gap, to a higher energy level below the band gap. This scattering is infrared absorption, where no photon is emitted.

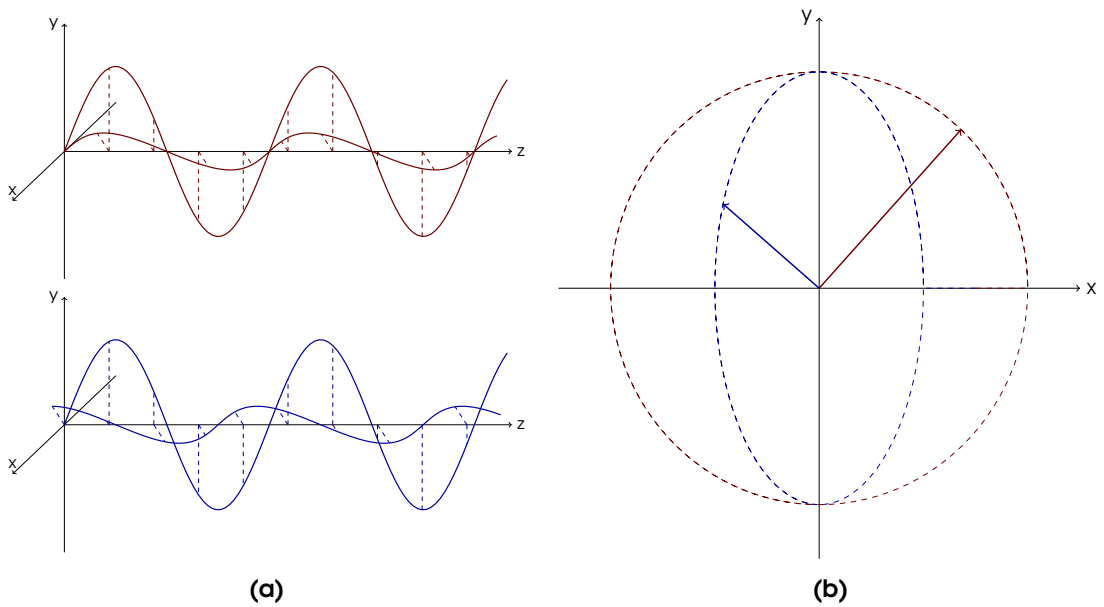


Figure 2.2: The principle of polarized light. Only the E field (red) or the H field (blue) of a light wave is shown at a time. (a) Shows two plots of two light waves oscillating together along the z-axis. In the top plot (red), the E field of one of the waves is at an angle 45° with respect to the other wave, but they oscillate in phase. In the bottom plot (blue), the H field of one wave is rotated as well, and the waves are out of phase by 90°. (b) Shows the resulting polarization. The arrows show the orientation of the E field (red) and the H field (blue) of the two plots at one particular point in time. The dashed lines show how the fields oscillate in the xy-plane. In-phase oscillation results in circular polarized light, while out-of-phase oscillation results in elliptical polarized light.

Other types of scattering do not necessarily involve electron transitions. One example is Rayleigh scattering, where a photon of the same wavelength is emitted, but in a random direction. This explains why the sky is blue, because particles in the atmosphere more easily scatters blue light than other wavelengths.

The final example involves an incident photon interacting with the vibrational states of the molecule, labelled Stokes or Raman scattering. When incident on a molecule, some of the photon's energy may be transferred to the molecule. This energy gain corresponds to one of the molecule's vibrational states. The photon is emitted with a longer wavelength. It has been Stokes shifted, or Raman scattered. The opposite can happen as well, that the emitted photon increases its energy with a value corresponding to a vibrational state, labelled anti-Stokes scattering.

Scattering cross-section

The scattering cross-section of a molecule is denoted σ and is a measure of how much of the incident light intensity is scattered by the molecule. If a molecule has a large σ, it will be easier to detect by means of optical sensing. Unfortunately, many biological molecules have a small Raman cross-section, which is one of the reasons why one would like to enhance the Raman signal. As demonstrated in [6], the scattering cross-section is proportional to λ^{-4}. Hence, shorter laser wavelengths is desirable in Raman spectroscopy. However, high-energy photons may damage biological samples. The therapeutic window, which is the wavelength range pre-

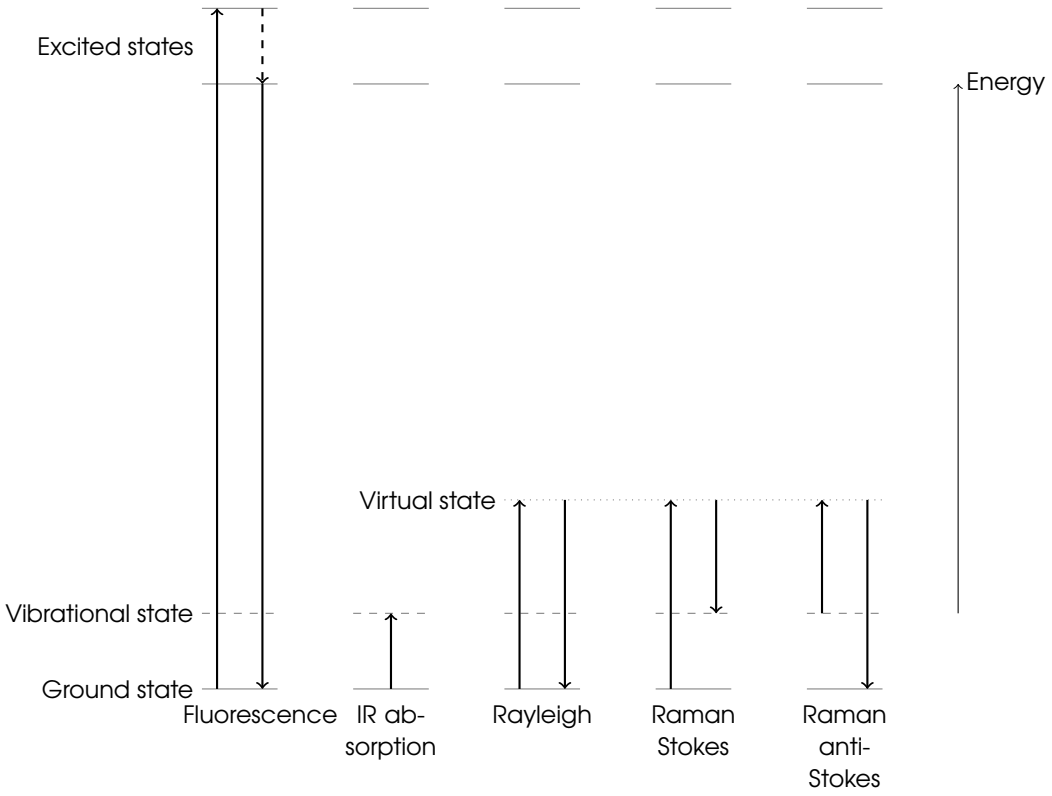


Figure 2.3: A simplified overview of selected scattering processes in a molecule. Some selected molecular energy states are shown. The vibrational states are much closer to the ground state than what is depicted. The excited states are electronic transitions. The virtual state is depicted in this fashion because, although it can be viewed as an excitation, energy is transferred instantly.

ferred for Raman spectroscopy, is between 600 and 900 nm.

2.3 Waveguides

Light is reflected at interfaces. This is evident in mirrors, which reflect nearly 100 % of incident light, and materials like water and glass which reflect light partially. Windows are a well known example of partial reflectors. If you stand in an illuminated room and look through a window in the dark of the night, you will only see your own reflection. Anyone standing outside, however, will see you clearly. This means that a window both transmits and reflects light. Materials that reflect and confine light can be used to guide waves of light from one point to another, if the material and design of the guiding structure is carefully selected. Any device which is capable of directing the propagation of light in such a fashion is called a *waveguide*. One simple example is a wave propagating between two perfectly reflecting mirrors. However, perfectly reflecting mirrors are practically nonexistent, so a more useful example is the *dielectric* waveguide. There, the wave is guided through the principle of *total internal reflection*. The discovery of waveguides has bent the propagation of light to our will, for example through optical fibers making the modern day internet possible.

This chapter will briefly go over the most important concepts of waveguides, with the goal of explaining the confining abilities of a rectangular dielectric waveguide, which is the structure used in Raman spectroscopy. The existence of an evanescent field will be presented as well. The theory presented here is mainly adapted from [6]. Of course, a proper treatment of waveguides requires Maxwell's equations to be solved with suitable boundary conditions, however, a more conceptual approach will be taken here.

Planar Mirror Waveguides

Recall the example mentioned above, with two perfectly reflecting mirrors guiding a wave. If they are lossless, any wave entering the waveguide will bounce back and forth between the mirrors, and emerge without loss of energy at the end of the waveguide. This is shown in Figure 2.4a. However, for a *mode* to be supported, the *roundtrip* condition needs to be satisfied. It states that for each roundtrip, the wave has to reproduce itself. This means that the distance travelled between two reflections needs to be an integer multiple of the wavelength. This ensures that a constant repeating pattern emerges in the waveguide. It is a geometrical condi-

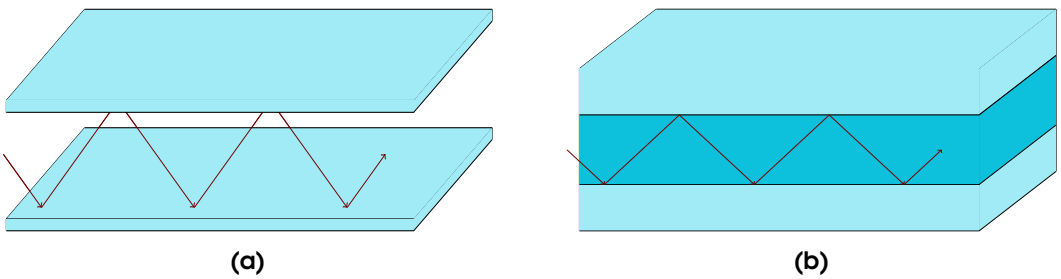


Figure 2.4: Two planar waveguides guiding a wave of light. (a) Planar Mirror Waveguide. The supported modes are those which between two reflections transverse a distance equal to an integer times the wavelength. (b) Planar Dielectric Waveguide.

tion which mathematically depends on the wave's entry angle θ to the waveguide and its wavelength λ, as well as the distance d between the mirrors,

$$\sin\theta_m = m\frac{\lambda}{2d}, \qquad m = 1, 2, \ldots . \tag{2.6}$$

In this expression, the phase shifts from the mirror reflections are hidden. When incident on a perfectly reflecting mirror, a wave will instantly experience a phase shift of π. The two reflections when travelling a roundtrip add up to 2π, such that the roundtrip condition is automatically satisfied. If the mirrors are not perfectly reflecting however, the phase shift is no longer 2π, but some other quantity 2φ. This is the case for dielectric waveguides.

Dielectric Waveguides

Dielectric waveguides are not perfect mirrors. Instead, light confinement in the waveguide depends on the ratio of the refractive index of the material in the waveguide, and outside of the waveguide. The refractive index of a material is a measure of the speed of light c in the material, $c = c_0/n$ where c_0 is the speed of light in vacuum, and n is the refractive index of the material. In vacuum $n = 1$, in water $n = 1.33$, in Si $n = 3.88$. Light is more easily confined to materials of high refractive index. If the refractive index in the waveguide is higher than in the surrounding material, the wave can be guided by *total internal reflection*. A wave incident on an interface between two materials will be refracted according to Snell's law. The angle of refraction depends on the angle of incidence and the refractive indices in the two materials. If the angle of refraction is 90°, the wave is not transmitted into the other material, but totally reflected. Solving Snell's law with this condition yields

$$\theta_c = \sin^{-1}\frac{n_2}{n_1}, \tag{2.7}$$

where n_1 is the refractive index of the guiding material, and n_2 the refractive index of its surroundings. θ_c is the critical angle, and an incident wave is totally reflected for all angles smaller than θ_c. Clearly, it is easier to confine a wave the smaller the ratio n_2/n_1.

In addition to this, the roundtrip condition needs to be satisfied. A correction in the expression for planar mirror waveguides is needed, because the phase shift introduced at the interfaces is 2φ for dielectric waveguides.

$$\frac{2\pi}{\lambda}2d\sin\theta - 2\varphi = 2\pi m, \qquad m = 0, 1, 2, \ldots, \tag{2.8}$$

where $\theta < \theta_c$. A wave propagating in a dielectric waveguide is depicted in Figure 2.4b. Some examples of dielectric waveguides are strip waveguides, rib waveguides, slot waveguides, and optical fibers. They are all depicted in Figure 2.5.

Rectangular Dielectric Waveguides

Up until now, dielectric waveguides of infinite width have been considered. However, in many cases the finite width of the waveguide cannot be ignored. Therefore, rectangular dielectric waveguides need to be considered. Including a finite width makes the waveguide confine light in both the x and y directions, making it possible to construct waveguides which guide only a single mode.

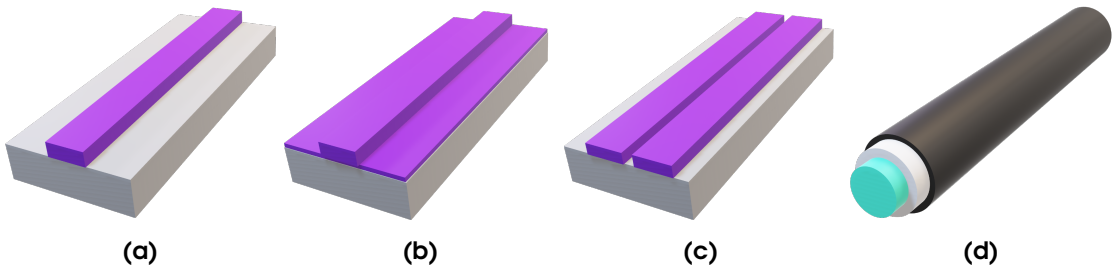


Figure 2.5: Different types of waveguides. (a) Strip (or ridge) waveguide, (b) rib waveguide, (c) slot waveguide, (d) optical fiber.

A rectangular waveguide may be surrounded by up to four different materials, one on each side. Including the waveguide material, a full treatment must take five wavenumbers (k_1, k_2, k_3, k_4, k_5) and five permittivity constants ($\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5$) into account. Most materials are nonmagnetic, so the permeability can be set to μ_0 everywhere. The height of the waveguide is b and the width is a. For TE modes, the dispersion relation of the propagation constants in y and x are derived in Yeh & Shimabukuro [7] as

$$k_y b = m\pi - \tan^{-1}\left(\frac{k_y}{\sqrt{k_1^2 - k_2^2 - k_y^2}}\right) - \tan^{-1}\left(\frac{k_y}{\sqrt{k_1^2 - k_4^2 - k_y^2}}\right), \tag{2.9}$$

$$k_x a = n\pi - \tan^{-1}\left(\frac{\epsilon_3}{\epsilon_1}\frac{k_x}{\sqrt{k_1^2 - k_3^2 - k_x^2}}\right) - \tan^{-1}\left(\frac{\epsilon_5}{\epsilon_1}\frac{k_x}{\sqrt{k_1^2 - k_5^2 - k_x^2}}\right), \tag{2.10}$$

where m and n are integers denoting the mode, k_y and k_x denote the wavenumbers in y (height) and x (width) respectively. The nearly identical expressions for TM modes can be found in [7] as well. As will be explained in Section 3.2, most waveguides used for SERS have the same structure: A SiN core on a silicon oxide bottom cladding, surrounded by air. This is the only waveguide type which will be considered in this project. Therefore, Equations (2.9) and (2.10) can be used to constrict the dimensions of the SiN waveguide, if it is to be single mode. The subscripts denote the material regions. 1 denotes Si_3N_4, 4 denotes SiO_2, while 2, 3 and 5 denote free-space. As will be mentioned in Section 3.2, the laser wavelength and Stokes shift will lie in the therapeutic window, between 600 and 900 nm. The waveguides need only support single modes within this range. Ignoring losses, this corresponds to $k_{max} = 2\pi/\lambda_{min} = 10.47 \times 10^6$ m^{-1} and $k_{min} = 2\pi/\lambda_{max} = 6.98 \times 10^6$ m^{-1} in free space.

Evanescent Field

Not all of the optical field distribution is contained within the waveguide. Some of the field escapes the waveguide, labelled the evanescent field. In many cases this is undesirable, because if the waveguide is surrounded by a lossy media, the evanescent field is associated with absorption. From an optical sensing point of view however, a large evanescent field is useful. This is because the field leaking out of the waveguide will interact with molecules around the waveguide. Thus, waveguide-based Raman spectroscopy would not be a reality without an evanescent field. Some waveguides have a larger optical field distribution into their sur-

roundings than others. Examples include slot waveguides and sub-wavelength grating waveguides (SWG). However, the waveguide's surface roughness will lead to scattering of the evanescent field. The trade-off between this loss factor and the increased interaction volume from a large evanescent field will be discussed in Section 3.2.

2.4 Surface Plasmon Interactions

Two phenomena are important for the electric field enhancement near rough metal surfaces, namely the surface plasmon-polariton (SPP) effect, and the localized surface-plasmon resonance (LSPR) effect. Both involve interaction between a phonon and a plasmon. The term *plasmon* refers to a quantum oscillation of the electron sea in metals, i.e., the conduction electrons. Charges moving in a magnetic field emit electromagnetic radiation. Thus, if incident light induces movement in a metal's conduction electrons, photons may be emitted to increase the total electromagnetic field.

SPP propagate at the interface between a metal and a dielectric [9]. The propagation length of a SPP on such an interface may be on the order of several hundred microns [10]. LSPR on the other hand, is an effect which is localized to the metal's edges, for instance at the tip of gold nano sized triangles [5]. In both cases, the electromagnetic fields may penetrate some distance into the dielectric surroundings. If the surroundings are molecules of interest, the interaction between these fields and the analytes may serve to enhance a Raman signal. This is the mechanism which is labelled surface-enhanced Raman scattering (SERS).

Although SPP and LSPR are closely related phenomena, there are some key differences. Firstly, SPP penetrate longer into the dielectric, about 150 nm, while LSPR penetrate about 25 nm. On the other hand, LSPR has a stronger field enhancement than SPP, typically 10^4 vs 10^2 enhancement, respectively [11].

In order to give a better understanding of the interaction between light and plasmons, the LSPR effect shall be described mathematically. The following explanation is adapted from P. L. Stiles *et al.* [5]. A condition for the LSPR effect to occur (in contrast to the SPP effect), is that the metal particle is significantly smaller than the wavelength of incident light. If this is the case, the electromagnetic field around the particle can be considered uniform, as shown in Figure 2.6, and plasmons oscillate and emit radiation in resonance with the incident light. In such a constant field, the Laplace equation of electrostatics can be used instead of Maxwell's equations. Consider a spherical nanoparticle on which z-polarized light is incident, the electric field E_{out} outside of the particle is given by

$$E_{\text{out}}(x, y, z) = E_0\hat{\mathbf{z}} - \alpha E_0\Big[\frac{\hat{\mathbf{z}}}{r^3} - \frac{3z}{r^5}(x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}})\Big], \tag{2.11}$$

where E_0 is the incident field and r is the radial distance. The polarizability is $\alpha = ga^3$, where a is the particle radius and g is defined as

$$g = \frac{\varepsilon_{\text{in}} - \varepsilon_{\text{out}}}{\varepsilon_{\text{in}} + 2\varepsilon_{\text{out}}}, \tag{2.12}$$

where ε_{in} is the dielectric constant of the metal nanoparticle, and ε_{out} is the dielectric constant of the environment. Note that in a plasmonic material ε_{in} will be

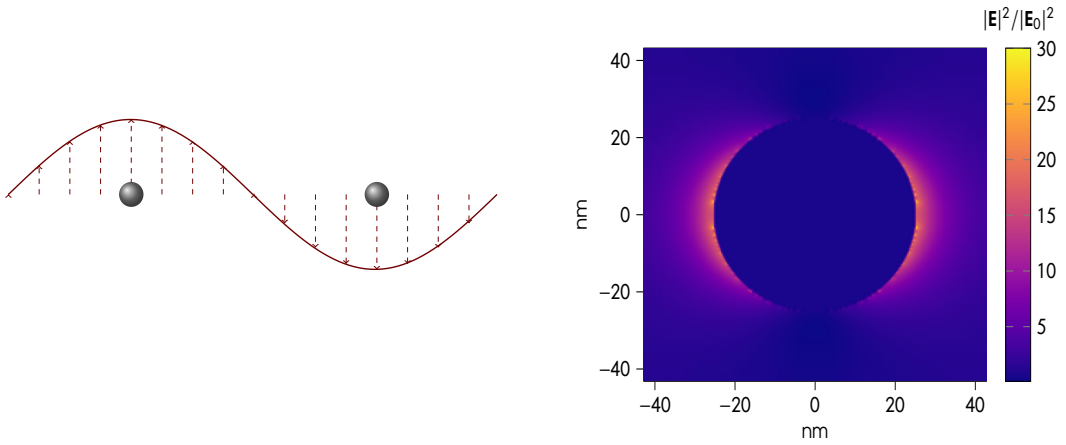


Figure 2.6: A small metal nanoparticle in the electric field of an incoming wave will be subject to a uniform field at any given point in time. Hence, all plasmons oscillate in resonance with the passing wave. This results in a field enhancement at the surface of the nanoparticle. Such an enhancement is shown in the field plot to the right, where $\lambda = 547$ nm light is incident on a spherical gold particle with a diameter of 50 nm in vacuum. The max enhancement is $|\mathbf{E}|^2/|\mathbf{E}_0|^2 = 30$. Visually, it can be seen that the enhancement is only large close to the particle. The electric field strenght was calculated from a 3D FDTD simulation in Lumerical.

negative, corresponding to enhancement. From the definition of g, it is clear that the maximal enhancement occurs when $\varepsilon_{\text{in}} \approx -2\varepsilon_{\text{out}}$.

Equation (2.11) shows that the field enhancement scales with r^{-3} (note that α will be negative), so the LSPR enhancement is a surface phenomenon. In the case of waveguide-based Raman spectroscopy, this means that analytes must be confined to a small sensing volume around the surface of the waveguide.

The magnitude of the Raman signal is linear with the electric field intensity, $|E|^2$. At the surface of the metal particle $r = a$, such that from Equation (2.11) we find that the radially averaged intensity $|\bar{\mathbf{E}}_{\text{out}}|^2$ is

$$|\bar{\mathbf{E}}_{\text{out}}|^2 = 2E_0^2|g|^2. \tag{2.13}$$

Here, the E-field enhancement is linear with $|g|^2$. However, when one considers the Raman scattering mechanism, it appears likely that this enhancement is applied twice. Light is first incident on the nanoparticle, at which point the field at the surface is given as $|\bar{\mathbf{E}}_{\text{out}}|^2$. This field then interacts with molecules close to the particle, and some of the resulting field is Stokes shifted. This scattered light is again incident on the nanoparticle, with a field given by $|\bar{\mathbf{E}}'_{\text{out}}|^2$, the prime referring to the Stokes shifted wavelengths. Hence, the enhancement factor can be written as

$$\text{EF} = \frac{|\bar{\mathbf{E}}_{\text{out}}|^2|\bar{\mathbf{E}}'_{\text{out}}|^2}{|\bar{\mathbf{E}}_0|^4} = 4|g|^2|g'|^2, \tag{2.14}$$

where the origin of the 10^4 enhancement of the Raman signal is evident, as for small Stokes shifts $g \approx g'$, such that EF scales with g^4. This is a major improvement, considering that only one in a million photons interacting with a molecule is Raman

scattered. EF can be calculated experimentally from

$$\mathrm{EF} = \frac{I_{\mathrm{SERS}}/N_{\mathrm{surf}}}{I_{\mathrm{NRS}}/N_{\mathrm{vol}}} = \frac{I_{\mathrm{SERS}}}{I_{\mathrm{NRS}}}\frac{N_{\mathrm{vol}}}{N_{\mathrm{surf}}}, \tag{2.15}$$

as given by [5]. I_{SERS} is the intensity from SERS, N_{surf} is the number of molecules bound to the metal nanoparticle, I_{NRS} is the normal Raman intensity, and N_{vol} is the number of molecules in the excitation volume. It will be easier to calculate the EF of a given structure, rather than predicting it analytically.

Chapter 3

The Physics of a Silicon Nitride Waveguide Platform for SERS Sensing

This chapter will build on the theory outlined in Chapter 2. In the following, the principles of Raman scattering, waveguides and plasmons will all be discussed simultaneously. The goal is to provide an understanding of what type of waveguide platform is best suited for Raman spectroscopy, which is the main issue of this project. Next, advantages of a waveguide based Raman sensing platform will be discussed. Then, in Section 3.2, follows an overview of the most important things to keep in mind when designing such a platform. Finally, in Section 3.3, a range of platforms reported in the literature are presented. The terms Raman spectroscopy, Raman scattering and Raman sensing will be used interchangeably. Also note that Si_3N_4, silicon nitride and SiN all refer to the same material. The theoretical considerations in this chapter are largely based on [5] and [11]. This chapter should leave the reader with a justification of the choices made in this project regarding design and analysis. Note that large parts of this chapter has been reused from the unpublished term project leading up to this thesis.

Raman spectroscopy was originally a free-space optical sensing technique. It was not until the discovery of SERS considerations were made to move Raman spectroscopy to a waveguide based platform. The strong enhancement of the Raman signal by surface plasmons on rough nanoscale metal surfaces made it possible to overcome previous issues with waveguides, such as the noisy Raman signal from the waveguide itself, labelled the Raman background. Additionally, advances in fabrication techniques at the nanoscale were a prerequisite to enable satisfactory control of the morphology of plasmonic structures and waveguides. The ultimate detection limit was reached when SERS was used to capture the Raman spectrum from a single molecule [3]. The advantages of a waveguide based platform are numerous, most evidently the ability to place an entire Raman sensing platform on a chip. An on-chip waveguide based SERS platform would enable multiplexing of samples, such that several molecules can be detected in parallel. Moreover, an on-chip system would be much cheaper and easier to handle than conventional free-space microscopes. *Photonic integrated circuits* (PIC's) are generally silicon compatible, which means that waveguide based SERS should be compatible with the huge CMOS industry. This has the two advantages that existing mass-scale fabrication techniques can be utilized, and that photonic and electric circuits can be integrated on the same chip. In total, significantly broadening the horizon of Raman spectroscopy's abilities.

3.1 Performance of a Raman Sensing Platform

The Raman signal from waveguide measurements arises when light from the laser pump interacts with analytes on the waveguide's surface. The performance of a Raman sensing system is thus characterized by the ratio of the Stokes signal power

P_s to the pump power P_s/P_p. This quantity must be captured in a suitable figure of merit for a Raman sensing platform. In general, the resulting Raman signal depends on the number of analytes on the surface and their Raman scattering cross-section, as well as the wavelengths of the pump, and the Stokes shift, the shape of the waveguide, and the waveguide and cladding materials. If the distribution of analytes is uniform along the waveguide, and the waveguide is uniform along its sensing region, then the ratio of Stokes to pump power can be written as

$$\frac{P_s}{P_p} = \eta\sigma\rho l, \tag{3.1}$$

where l is the length of the waveguide's sensing region, ρ is the density of analytes, σ is the scattering cross-section of the analytes, and η is labelled the *conversion efficiency*, which accounts for the waveguide's shape and material, as well as the pump and Stokes shifted light. It is assumed that the waveguide is lossless. Equation (3.1) is derived in [12]. Note that σ and ρ depends on the type of analyte present on the waveguide surface, not the waveguide itself. They may therefore be ignored when comparing different waveguide platforms. Additionally, since most waveguides are uniform (or periodic) in their sensing region, l only serves to increase the signal (in lossless waveguides). In the end, η is the quantity which depends on the waveguide's cross-section, and is a useful measure of the performance of a waveguide sensing platform. In Section 4.3 it shall be used as the sensing figure of merit (FOM) for the waveguides considered in this thesis. The general form of η, as given in [12] is

$$\eta = \frac{\lambda_s^2 n_g(\lambda_p) n_g(\lambda_s)}{n(\lambda_s)} \frac{\iint_{clad} |E_m(\mathbf{r},\lambda_s)|^2 |E_m(\mathbf{r},\lambda_p)|^2 dx^2}{\left(\iint_\infty \varepsilon(\mathbf{r},\lambda_s)|E_m(\lambda_s)|^2\right)\left(\iint_\infty \varepsilon(r,\lambda_p)|E_m(\lambda_p)|^2\right)}, \tag{3.2}$$

where λ_p and λ_s denote the pump and Stokes shift wavelengths, respectively, $\mathbf{r}$ is the position vector, E_m the modal electric field, $n_g(\lambda_i)$ denotes the wavelength-dependent group index, and n denotes the refractive index of the cladding. The integral in the numerator is over the waveguide's cladding.

The unit of the conversion efficiency η is in steradians, and it can be understood as a measure of the waveguides ability to excite, and subsequently collect light scattered from analytes at the surface. The excitation wavelength is λ_p, and the collection λ_s, which is why both the pump and Stokes shifted wavelengths are evaluated in the expression for η. The fact that the Stokes shift is small, allows for the approximation $\lambda_s \approx \lambda_p = \lambda$, which simplifies the above expression of η to

$$\eta = \frac{\lambda^2 n_g^2}{n_{clad}} \frac{\iint_{clad} |E_m(\mathbf{r},\lambda)|^4 |dx^2}{\left(\iint_\infty \varepsilon(\mathbf{r},\lambda)|E_m(\lambda)|^2\right)^2}. \tag{3.3}$$

The difference in results when including the Stokes shift, is about 10 % [13]. The modal fields E_m may be calculated with a mode solver.

3.2 Considerations for the Use of SiN Waveguides in Raman Sensing

In the following section, important considerations in design of a waveguide based Raman sensing platform are discussed. Raman spectroscopy is an extraordinarily

powerful analysis tool, because the Raman signal from an illuminated molecule is unique. It is like a molecular fingerprint. The optical sensing in a waveguide-enhanced platform occurs on analytes confined to a small sensing volume at the surface of the waveguide. For one particular mode of light, the performance of such a platform is determined by the mode's overlap with the surroundings, i.e., its evanescent field. It is a condition that the optical field in the waveguide extends into the surroundings, such that analytes can interact with light and emit a Stokes shifted signal. The performance is further characterized by the analytes' ability to bind to the waveguide surface. This follows from the fact that SERS is a surface phenomenon, and that in a waveguide platform the laser is injected at the same point as the signal is collected. Only analytes close to the surface will interact with light from the waveguide, and the Raman signal must be coupled back into the waveguide to be detected. In some cases bonding can be achieved by the use of an adhesion layer, however some biological molecules will bind to gold. As will be discussed later, gold is the most widespread plasmonic structure used in SERS. The performance by a Raman sensing platform is limited by scattering losses along the waveguide.

Material Choice

It is desirable to illuminate biological samples in the near infrared wavelength range, between about 600 and 900 nm. This is labelled the *therapeutic window*. Wavelengths within this range have low enough energy to not damage the sample, while simultaneously avoiding water's strong absorption of wavelengths longer than 1 micron. Additionally, it is desirable to use as short wavelengths as possible, because from Section 2.2, the scattering cross section scales with λ^{-4}. A suitable waveguide core must be optically transparent in this range. Moreover, an on-chip PIC requires waveguides which support tight bends, without leaking their field into the surroundings [11]. High index contrast waveguides (HIC), with a large refractive index difference between the core and cladding contain their fields even through tight bends. A waveguide intended for biological sensing will usually have an upper cladding of water or free-space, where the analyte interaction occurs, and a bottom cladding of an insulating layer like silicon oxide. As the refractive indices of air, water and SiO_2 are 1, 1.33 and 1.45, respectively, the HIC is measured relative to silicon oxide.

Silicon-on-insulator (SOI) is a HIC waveguide, as silicon has a refractive index of 3.5. SOI is dominant in the CMOS industry, and vastly used in PIC. Unfortunately, silicon is not optically transparent at the target wavelength range. Si_3N_4 waveguides on silicon oxide is the alternative, as SiN *is* transparent in the theurapeutic window. Since SiN has a refractive index of about 2, it is still considered a HIC waveguide. A minimal bend radius as low as 5 microns can be supported by SOI waveguides, but SiN waveguides support a minimal bend radius of 40 micron [11], which is acceptable. In the field of waveguide-enhanced Raman spectroscopy, SiN waveguides are dominant, especially when the injected laser wavelength is 785 nm. There are examples of alternative waveguide platforms. In blood analysis for instance, a Ta_2O_5 waveguide platform has been used [14], with a laser wavelength of 532 nm. Such platforms will not be considered in this thesis.

Waveguide Design

Continuing with SiN as the preferred waveguide, its specific design must be optimized for Raman sensing. Since analytes of interest will be close to waveguide's surface, some of the light's optical power must penetrate out of the waveguide in order for light to interact with the analytes. However, this evanescent field experiences scattering losses from surface roughness, which arise from imperfect fabrication techniques. Therefore, although an ideal waveguide for Raman sensing would have a large evanescent field, this would result in a short decay length due to scattering losses. Attempts have been made to quantify this trade-off, for instance by Kita et al. [15]. A dimensionless figure of merit was derived to distinguish between the performance of different waveguide designs as

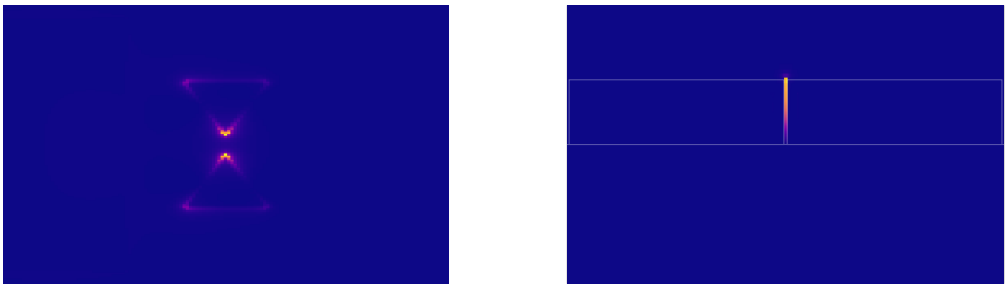
$$\text{FOM} = \frac{\beta}{\alpha_s}, \tag{3.4}$$

where α_s is the scattering loss per unit length and β is the Raman gain coefficient. β is an alternative way to capture the waveguide's coupling efficiency, given as η in Equation (3.3). In [15], a classical strip waveguide, a slot waveguide and a sub-wavelength grating (SWG) waveguide were analysed in terms of their performance in Raman sensing, using the above figure of merit. Analysis was restricted to the telecommunications wavelength at 1550 nm. Nevertheless, it was concluded that the slot waveguide is best for Raman sensing, which coincides with recent SERS research preferring the slot waveguide over the strip waveguide. These results will be discussed further in Section 3.3. The TE slot waveguide was identified as the best waveguide for Raman sensing, outperforming the TM strip waveguide by a factor of 5, which in turn outperformed the SWG waveguide by a factor of 2. The reason why waveguide performance differs under TE and TM operation, is due to the spatial confinement of the evanescent electric fields. For TM modes, the modal fields overlap with the top of the waveguide, while the TE modes overlap with the sidewalls of the waveguide. If analyte interaction occurs at the waveguide's sidewalls, such as in a slot waveguide, TE modes would naturally perform better.

The type of plasmon interaction (SPP or LSPR) will depend on the waveguide design. In waveguides such as the slot waveguide, SPP propagates in the gap. This results in a large enhancement of the electric field in the gap. LSPR dominated platforms rely on nanostructures with sharp edges, such as nanotriangles or bowties. In such configurations, the Raman enhancement occurs on analyte molecules on the edges of the metal nanoparticles. This difference is illustrated in Figure 3.1.

3.3 Relevant Results in Publications on SERS

There are numerous publications in the field of SERS in general, but also waveguide-assisted SERS in particular. The research carried out by *Photonics Research Group* affiliated with Ghent University in Belgium has been given special attention in this project. Work on SiN waveguides for Raman sensing with gold bowties on SiN waveguides (LSPR interaction) [16], [17], [13], [11], gold nanotriangles on SiN waveguides (LSPR interaction) [18] and nanoplasmonic slot waveguides (SPP interaction) [19], [20] has been reported. Nanoplasmonic slot waveguides were reported to have the best performance [4], approaching that of free-space gold nanodomes [21].



(a) (b)

Figure 3.1: Illustration of the difference between the LSPR and SPP interactions in waveguides. The intensity denotes the magnitude of the electric field intensity $|E|^2$. (a) Shows a bowtie structure placed on top of a waveguide. The electric field is strongest at the triangles' edges. Image is reused from the term project. (b) Shows the cross-section of a slot waveguide, with the waveguide structure overlaid as white lines. Clearly, the electric field is confined to the gap.

Similar work on nanoplasmonic slot waveguides has been produced outside of this research group [22]. Other waveguide structures, such as suspended silicon nitride waveguides have been investigated as well [23].

Some general remarks on the findings in these papers are described in the following. Regarding the dimensions of the waveguide, the cross-sectional height is usually set to 220 nm, while different widths have been investigated numerically and experimentally. This height correspond to the transmission height for telecommunications optics, while the widths are tuned to support single mode operation in the therapeutic window. The excitation laser is 785 nm, and the observed Stokes shift is commonly 1339 cm^{-1}, which results in a Raman signal at 877 nm. This corresponds to NO_2 stretching, which is commonly found in biological molecules. nitrothiophenol (NTP) is one such molecule, with the advantage that it binds to gold as a monolayer [24]. This makes NTP useful in assessing the biological enhancement of Raman signal. Gold is used as the plasmonic material, since gold's surface plasmons has a desirable response to the laser wavelength and Stokes shift. Silver would be an alternative, but its toxic properties poses an unnecessary obstacle. In nanoplasmonic slot waveguides, Al_2O_3 has been deposited as a *spacer* layer to narrow the slot, before the deposition of gold. It was shown that the enhancement of the Raman signal was dependent on the combined thickness of the Al_2O_3 and Au layer [19]. This is useful, because aluminum oxide can in theory be deposited with a precision of a single atomic layer, by means of atomic-layer deposition (ALD). Because of low adhesion between SiN and gold, a thin ($\sim$ 2 nm) layer of Ti is deposited to improve adhesion. These results have significantly influenced the waveguides designed and simulated in this project. As will be mentioned in Section 4, the waveguide dimensions are inspired from these findings, and gold is the chosen plasmonic material.

Chapter 4

Numerical Computer Simulations in Lumerical

In designing a photonic structure, it is of critical importance to predict its performance prior to fabrication, to ensure the system meets the requirements of the designer. The knowledge of how a system's performance depends on its components and dimensions, may be obtained either experimentally or theoretically. Experiments performed on previous platforms provide insight into the performance of future platforms. Photonic theory may be used to predict the performance of completely new platforms. Some properties of particular importance in the field of waveguides are the effective index n_{eff}, loss α, the propagation constant β, and the confinement factor Γ. For any photonic system, these values are revealed by solving Maxwell's equations, which were introduced in Section 2.1. However, analytical solutions of these equations exist only for very simple systems. In predicting the performance of slot waveguides such as those described in this thesis, one must turn to numerical solutions of Maxwell's equations. This usually involves discretizing Maxwell's equations, and solving them numerically in a 2D or 3D grid. There exist a family of such finite-difference method (FDM)s designed to solve Maxwell's equations numerically. The method employed in this thesis is the finite-difference eigenmode (FDE) method. The mathematical framework needed to understand these approaches will be shortly described in Section 4.1. The program used to carry out the numerical simulations in this project is the *Lumerical* software provided by Ansys. The details of how this software was employed is described in Section 4.2. The most important results of these simulations and how they impacted the final design of the waveguide platform is described in Section 4.3.

4.1 Numerical Computation Method: FDE

The finite-difference eigenmode (FDE) method was first proposed by Zhu & Brown [26], where vectorial finite-differences are used to compute Maxwell's equations in a time-static 2D grid. It is of particular use in waveguide analysis, since they often have a uniform cross-section. The results computed with the FDE method, are therefore valid throughout the structure, as long as the cross-section stays the same. In broad terms, it works by solving an eigenvalue problem of one particular electromagnetic frequency ω in a 2D cross-section with known refractive indices. The modal field profile, as well as the effective refractive index $n_{eff} = c\beta/\omega$ and the absorption loss α_s are calculated. The starting points of the method, are finite-differences, and the 2D Yee cell.

Finite-differences are a way of approximating the derivative of a function $f(x)$ at a point x_0 when the function itself is known. The central difference of a function $f(x)$ around the point x_0 is

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0+\Delta) - f(x_0-\Delta)}{2\Delta}, \tag{4.1}$$

where Δ is a small increment in x. This is useful for solving differential equations, such as Maxwell's equations. In such a case, f would be the wave-equation of

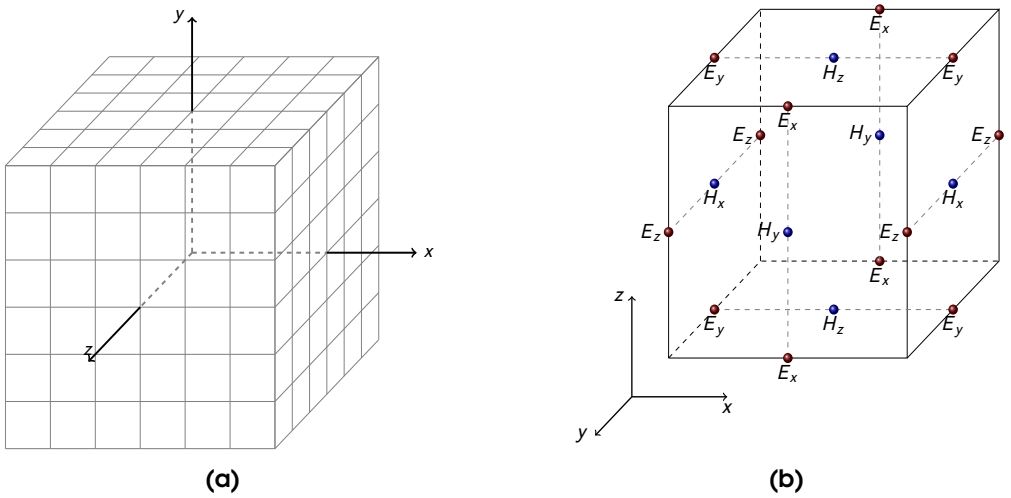


Figure 4.1: Illustration of the division of space in the FDE and FDTD methods. (a) Some kind of domain split into $6 \times 6 \times 6$ distinct cells, like a mesh. This particular mesh is uniform, however in FDTD simulations the mesh need not be uniform nor isotropic. (b) One single cell (the Yee cell) showing the positions of the E (red) and H (blue) components. The thin dotted lines show how the next H_i component is calculated through two neighbouring E_j components, details are found in [25]. This figure is reused from the term project.

either the electric E, or the magnetic H field. The first step in a FDE computation, would be to divide the analysis domain into special cells, labelled Yee cells. The Yee cell was first proposed by Kane Yee [25] in 1966, as a way to compute Maxwell's equations in the time-domain, the so-called finite-difference time domain (FDTD) method. The cell consists of several decomposed electromagnetic vector fields E_i and H_i, distributed in such a fashion that the values of adjacent vector fields are calculated from numerically solving Maxwell's equations as finite-differences. Through this iterative algorithm, which will be labelled a *simulation*, the entire field in the domain may be calculated. The details of the position of the fields in the Yee cell, or the equations used, are not given here. A 3D Yee cell is shown in Figure 4.1. A 2D Yee mesh actually consists of 3D Yee cells, but with extension in only two of the three spatial dimensions.

One of the advantages of the FDE method proposed by Zhu & Brown [26], is index-averaging. At curved interfaces, a single Yee cell may extend through two materials of different refractive indices. In such a cell, an average refractive index is assigned. The final step before starting a simulation, is to define boundary conditions. One common boundary condition, which will be employed in this thesis, is perfectly matched layer (PML). Such a boundary is designed to absorb all incident fields, which is good to simulate *open borders*, where the surroundings of the domain extends indefinitely. In the calculations carried out in Zhu & Brown, accurate results were obtained when the grid size was approximately $\lambda/15$ [26].

In summary, the FDE method is excellent when calculating guided modes of a uniform waveguide structure. However, the fields in only a single cross-section is calculated. In a structure where the cross-section changes, other methods should be considered, such as the eigenmode expansion (EME) method [27]. In such methods, the electromagnetic fields of several cross-sections are solved by FDE, and then the fields are propagated between the cross-sections. Another case where

FDE may not be as useful, is for scattering, or time-dependent fields. FDTD is more suited to consider such problems, since Maxwell's equations are solved in the time-domain [27]. In the waveguide structure considered in this thesis, described in detail in Section 4.2, the *active* part of the waveguide — where Raman interaction takes place — *is* uniform, and the fields of interest are the guided mode. In such an analysis, the FDE method is well suited. However, in the structures away from the active region, such as taper regions, EME could have been employed. Additionally, accurate analysis of scattering and plasmon interactions may require FDTD. This thesis is restricted to FDE analysis, but in future analysis of this waveguide platform, EME and FDTD may be considered to provide a complete analysis of the structure.

FDE In Lumerical

Lumerical's FDE solver is available in *MODE*, one of Lumerical's products. Other FDE solvers are available, such as *COMSOL Multiphysics*, but Lumerical MODE is employed in this thesis. Information on how the FDE method is implemented, including guides and examples, are available on Lumerical's web page [28]. Some adjustments are made to the FDE method described above. The most important difference is that although the vector E_i and H_i fields are situated at different points in the Yee cell, they are all interpolated to a single grid point, once a simulation has stopped running. This is user-friendly, as every mesh cell (corresponding to a point in space) will contain all six components of the E and H vector fields.

One advantage of Lumerical, is the possibility of defining a non-uniform mesh. This is useful when considering the slot waveguides in this thesis, since some index layers are thin, and the gap in the structure is narrow. A mesh refinement was employed in and around the gap, to obtain accurate results. This mesh is described in Section 4.2-B, and shown in Figure 4.4.

4.2 Numerical Simulation Setup in Lumerical

After selecting FDE as the computation method of choice, the next task is to implement a complete simulation environment in Lumerical's MODE software. The simulation should be an as accurate representation of the waveguide's cross-section as possible, while simultaneously being computationally feasible. In order to achieve this, the following elements need to be carefully defined:

(i) *A spatial restriction of the simulation domain*, i.e., defining the cross-sectional area in which the calculations are carried out. This domain must contain the waveguide structure, possible external fields, and objects which may cause reflections back into the structure.

(ii) *The size of the Yee cell*, labelled the *mesh size*. Recall that Zhu & Brown obtained good results for a mesh size of $\sim \lambda/15$ [26]. Lumerical allows for mesh cells to be rectangular, and the mesh need not be uniform. This makes it possible to define a refined mesh in regions where the refractive index changes rapidly, such as in the waveguide gap.

(iii) *The boundary conditions*, which define what should be done with fields incident on the domain boundaries. Examples of boundary conditions are reflecting, absorbing, or conducting.

(iv) *The refractive indices* n in the domain. This involves inserting the waveguide structure, and defining the refractive index of the surroundings, which in this case will be air, $n_{air} \approx 1$. Initially, this is defined independently of the domain and mesh size. A meshing step before the simulation starts, assigns each mesh cell a refractive index according to the index-averaging FDE algorithm.

(v) *The frequency* ω (or wavelength λ) for which the calculations are carried out. Lumerical allows for dispersion sweeps, meaning that several frequencies are calculated in a single simulation.

The rest of Section 4.2 concerns how these five elements were implemented in the simulation platform, which was developed for this thesis. Section 4.2-A describes the waveguide design in detail, solving point (iv) above. Section 4.2-B defines the properties of the simulation domain, solving points (i)—(iii) above. Finally, for point (v), the wavelengths of interest are the laser pump $\lambda_p = 785$ nm, and possible Stoke shifts. As mentioned in Section 3.3, the Stoke shift of 1339 cm^2, which is $\lambda_s = 877$ nm at this pump λ_p, is commonly used. This shift corresponds to N-H stretching. This is but one of the large number of vibrational modes present in biological molecules. However, analysis will be restricted to these two wavelengths, λ_p and λ_s. The pump wavelength is of course critical for the performance, while λ_s will be considered to predict the effect of a Stokes shift in the waveguide. More details of the simulation setup, in particular how to replicate these simulations in Lumerical, and the computational scripts used, are given in Appendix B.

A: Waveguide Design

The waveguide structures considered in the simulations in this thesis were Si_3N_4 slot waveguides, equivalent to two SiN ridge waveguides separated by a gap. The waveguides were placed on top of a 2.4 µm thick insulating SiO_2 slab. As mentioned in Section 3.2, the fact that SiN on SiO_2 are HIC waveguides, and the large thickness of the SiO_2 layer, should ensure that light is confined to the Si_3N_4 core, and prevent any light from coupling into the Si wafer beneeth the oxide. For this reason, the Si wafer was omitted in the simulations. Figure 4.2 shows a sketch of the waveguide design. The refractive indices of Si_3N_4 and SiO_2were taken from experimental measurements by H. R. Philipp [29] and E. D. Palik [30], respectively, provided in Lumerical's material database. From now on this waveguide structure will be labelled the *slot waveguide*, the two Si_3N_4 ridges the *rails*, the distance between the rails the *gap* g_0, the width of each rail the *rail width* r or simply the *width*, the height of each rail the *rail height* h or simply the *height*, and the total width of the waveguide structure (2 × rail width + gap) the *waveguide width*, as labelled in Figure 4.2. A variety of slot widths and gaps were simulated, and the waveguide height was set to either 220 nm or 140 nm. 220 nm is a standard height in integrated SOI photonics, since they typically operate in a single TE mode at the 1550 nm transmission wavelength. As mentioned in Section 3, other studies have employed SiN waveguides for Raman sensing at this transmission height. This has the advantage of being compatible with existing CMOS lines, however shorter wavelengths, such as 785 nm or even visible light, are used in Raman spectroscopy. Shorter wavelengths require smaller waveguide dimensions as well, to achieve single TE mode operation. For this reason, this thesis focuses on slot waveguides with a height reduced from the transmission height, to 140 nm.

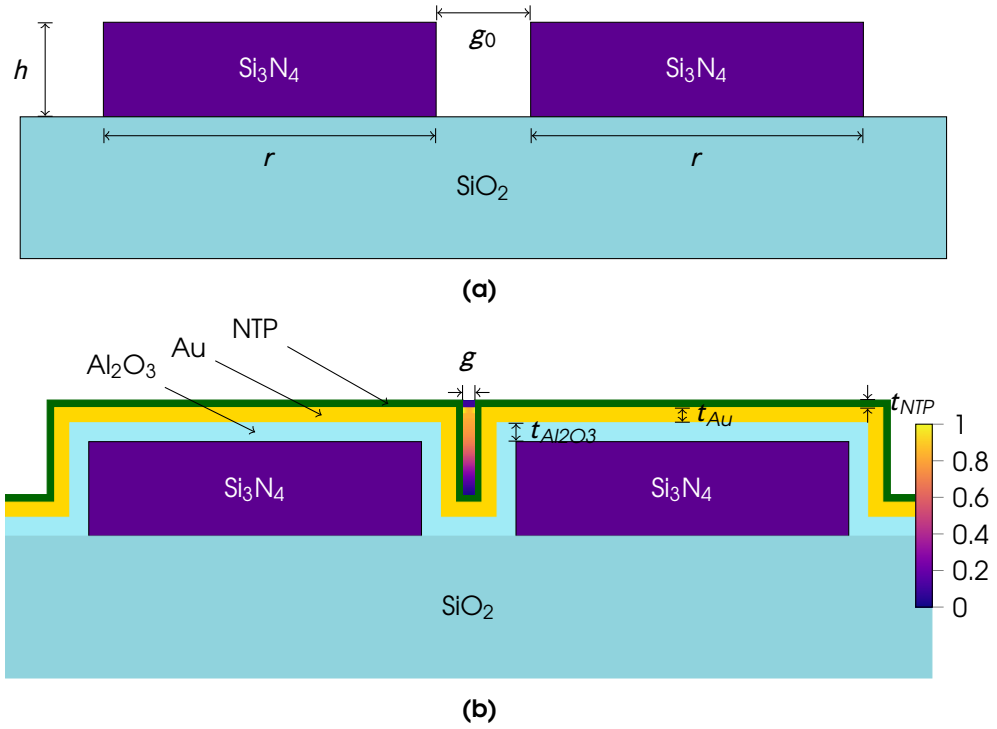


Figure 4.2: Cross-sectional sketch of the waveguide design used in the simulations, with the relevant lengths indicated. (a) shows the plain SiN waveguide, while in (b) all three additional layers of alumina, gold and NTP are present. The inset in the gap is a plot of $|E|^2$, normalized to a maximum value of 1. In the (b) configuration, nearly all the electric field power is confined to the gap. A 3D image of the waveguide is displayed in Figure 4.3.

On top of the waveguide, layers of Al_2O_3, Au, and NTP were placed. The refractive indices of Al_2O_3 and Au were taken from E. D. Palik [30] and Johnson & Christy [31], respectively. The refractive index of NTP was set to 1.538 [24]. In the simulations, zero, one, two, or all three of the layers were present, but their respective order was always the same. The ordering was NTP on gold on alumina, due to their respective purpose. The NTP was intended to model a monolayer of an organic substance susceptible to Raman scattering, which is why it must be placed on top, and its thickness is set to only 1 nm. The gold is the plasmonic structure responsible for enhancement of the electric field due to surface plasmon-polariton (SPP), and the alumina is a spacer layer, serving as a way to increase the rail width r and narrow the gap g. A summary of the materials, their ordering and respective thicknesses is found in Table 4.1.

This amounts to a large amount of simulation variables. There is a total of 8 possible layer configurations, two waveguide heights (220 nm and 140 nm), and two wavelengths (λ_p, λ_s). This brings the number of simulation configurations up to $2^5 = 32$. In addition, the rail widths r, the gap g_0, and the alumina t_{Al2O3} and gold t_{Au} thicknesses may vary continuously. This vast number of simulation configurations, makes it necessary to restrict analysis to certain values of interest. Next, such restrictions will be imposed on the above variables. The simulation results presented in Section 4.3 follow these restrictions. However in preliminary simulations, and for verification, alternative simulations were carried out. These simulations showed that the imposed restrictions were reasonable, and are displayed in Appendix C.

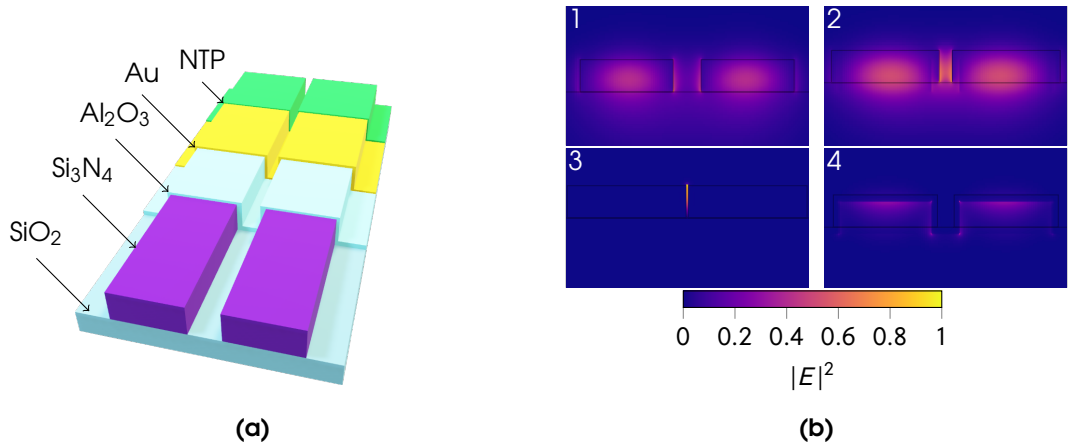


Figure 4.3: (a) Model of the slot waveguide considered in all simulations, with all possible layers included. (b) Fieldplots of normalized $|E|^2$ in the cross-section of the waveguide for different configurations. (1) Plain SiN waveguide, showing limited modal overlap in the gap. (2) SiN coated with alumina, showing an increased modal overlap in the gap. (c) SiN coated with alumina and gold, resulting in a strong electric field localized in the gap. (4) SiN coated with gold. Here, the strongest electric field interaction is localized to the gap edges, which is typical of the LSPR interaction. The interfaces between the waveguide structures and air are marked with black lines.

Table 4.1: Overview of the materials utilized in the Lumerical simulations. Their order was fixed, and they are tabulated here in the same fashion. Consequently, the material at the bottom of the table (SiO_2) is the bottom structure in the waveguide. NTP is always on top when present. The refractive index of NTP is assumed the same for all wavelengths, and is not taken from a model.

Material	Always Present	Model	Thickness [nm]	Purpose
NTP	✗	None, $n = 1.538$	1	Model monolayer
Au	✗	Johnson & Christy [31]	5 — 30	Plasmon layer
Al_2O_3	✗	E. D. Palik [30]	5 – 70	Spacer layer
Si_3N_4	✓	H. R. Philipp [29]	140 or 220	Waveguide
SiO_2	✓	E. D. Palik [30]	2400	Insulating layer

Table 4.2: Summary of the relevant variables used in the simulations. The Boolean parameters indicate that some layers may either be present or absent. The parameters NTP, Au, and Al_2O_3 refer to the additional layers, and t_{layer} to their respective thicknesses. The remaining parameters are: h, the waveguide height; g, the gap; r, the rail widths; λ the wavelength at which the waveguide platform is evaluated. Considering the Boolean and discrete parameters, results in $2^5 = 32$ possible simulation configurations. However, the NTP layer was included whenever alumina or gold was present, the height was mostly set to 140 nm, and the wavelength was set to 785 nm.

Parameter	Type	Value
NTP	Boolean	True/False
Au	Boolean	True/False
Al_2O_3	Boolean	True/False
t_{NTP}	Fixed	1 nm
t_{Au}	Continuous	$t_{Al2O3} + t_{Au} < g_0/2$
t_{Al2O3}	Continuous	$t_{Al2O3} + t_{Au} < g_0/2$
g_0	Continuous	> 0
r	Continous	> 0
h	Discrete	140 nm or 220 nm
λ	Discrete	785 nm or 877 nm

The physical restriction on r and g_0 is that they be larger than zero, however, wide rails result in a multimode waveguide, and a large gap will not give the a strong modal gap confinement, necessary for Raman enhancement. This puts an upper bound on r and g_0. The thicknesses t_{Al2O3} and t_{Au} may not be so large that they cover the gap, such that $t_{Al2O3} + t_{Au} < g/2$, where t_{NTP} is subtracted from the right hand side whenever NTP is present. In fact, whenever the alumina layer or the gold layer is included, the NTP layer is included as well. The effect of the NTP layer is to confine more light near the waveguide surface. This resembles the effects of a biological monolayer, such that including it makes the simulations a more realistic representation of experimental measurements. Lastly, considering the waveguide height and wavelength of interest, the height shall mostly be set to 140 nm, and the Stokes shift will mostly be ignored. Previously, the performance of SiN waveguides at the 220 nm transmission height, has been reported, both depending on the gap and rail widths [13], and depending on the alumina and gold thicknesses [19]. Such considerations will in general not be repeated here. Instead, the height of the waveguides simulated in this thesis will be 140 nm, unless explicitly stated otherwise. Dhakal *et al.* [13] also reported that including (Equation (3.2)) or excluding (Equation (3.3)) the Stokes shift resulted in a 10 % discrepancy. In Appendix C similar results are shown. This relatively small error, and the fact that simulation time is doubled when accounting for the Stokes shift (since the modal fields have to be calculated at two wavelengths λ_p and λ_s), it was omitted. A summary of all variables, their values, and type, is found in Table 4.2.

B: Simulation Domain

The simulation domain had to be large enough to contain the evanescent field from the waveguide. The waveguides were designed to be single TE mode, so the evanescent fields would mostly be horizontal. Additionally, the width was larger than the height. Therefore, a rectangular domain was used. The width was set

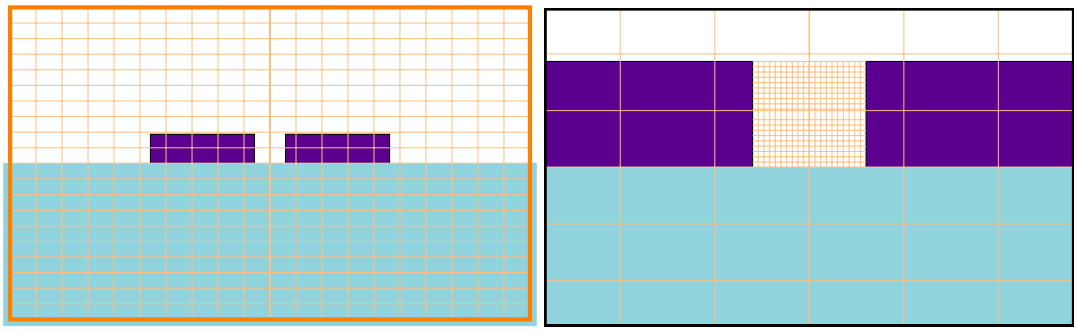


Figure 4.4: Sketch of the simulation domain and the mesh (left), and a magnified sketch of the mesh refinement region in the gap (right).

to 2,500 nm, and the height was set to 1,500 nm, with the center of the gap at the origin, as displayed in Figure 4.4. This means that for 140 nm high and 500 nm wide rails, the domain boundaries were at least 500 nm away from each rail both horizontally and vertically. This is a reasonable buffer, as evanescent fields typically do not penetrate more than some hundreds of nanometers into the surroundings. Rigorous simulations confirming this were not carried out, but the modal fields did not appear to change significantly with increased domain size.

The mesh size was set to 10 nm in each direction, which is $\sim \lambda/80$ for $\lambda_p = 785$ nm. This is a finer mesh than the $\lambda/15$ mesh, reported by Zhu & Brown to yield good results. A finer mesh was selected, simply because even at a 10 nm mesh, simulations completed relatively quickly. Again, rigorous simulations were not carried out, but decreasing the mesh below 10 nm appeared to significantly increase computation time, while still producing near identical results.

However, a refined mesh region was used in the gap, where the mesh size along the waveguide width was set to 0.5 nm, and 2 nm along the height. The reason for this choice is that the minimal layer thickness is NTP, with 1 nm. The mesh should be smaller than the minimal feature. In the direction of the waveguide height, however, this requirement was relaxed. This is because since the guided modes should be TE, they oscillate horizontally in the gap. Therefore, it is not as critical to have such a fine mesh along the height. An illustration of the mesh refinement is shown in Figure 4.4.

The boundary conditions of the simulation domain were set to PML on all sides, to simulate reflection-free absorption of light from the waveguide structure into the surroundings. There should be no reflections from the surroundings, and the insulating silica layer should prevent conduction into the Si wafer.

These domain properties were valid for nearly all simulations, only in a few cases were adjustments made to the domain described above. Some simulations included very large rails, in which case the domain was extended horizontally. Additionally, in simulations involving very large gaps, with only a small modal confinement in the gap, the mesh refinement in the gap was removed.

4.3 Simulation Results

The results presented in this section were computed using the finite-difference eigenmode (FDE) solver in Lumerical MODE. The FDE method was described in Section 4.1. The simulation setup and waveguide design was described in Section 4.2, includ-

ing justifications for the choice and value range of variables. The Lumerical script used to calculate the simulation results can be found in Appendix B. The structure considered is a Si_3N_4 slot waveguide on a SiO_2 insulating slab. Additional layers of a dielectric (NTP), gold, and alumina may be present on the waveguide. The parameters of interest are the dimensions of the waveguide, i.e., the gap g_0 and rail widths r, and the thickness of the alumina t_{Al2O3} and gold t_{Au} layers. In all calculations, the fundamental TE mode is considered where relevant, at the Raman excitation laser wavelength λ_p = 785 nm. The waveguide height is 140 nm, unless stated otherwise.

Two properties are considered, the *conversion efficiency* η, and the absorption loss α_a. This loss is calculated directly in MODE, while the computation of η requires an additional step, since Equation (3.3) must be calculated. Details of this calculation are available in Appendix B. The conversion efficiency η (units in steradians), given in Equation (3.3), is proportional to the Stokes power to the pump power in the waveguide, $P_s/P_p \propto \eta$, as stated in Equation (3.1). Thus, increasing η will increase the Raman signal obtained from a waveguide. This makes tie conversion efficency a reasonable choice of figure of merit (FOM) for waveguide-assisted SERS. Note that only the absorption loss α_a is considered in this section. However, in high index contrast (HIC) waveguides such as SiN on SiO_2, scattering loss is the dominant loss contribution. Considerations of the scattering loss require additional steps, which are not computed here. Kita *et al.* [15] calculated scattering losses of certain SOI sensing platforms, with the use of perturbation theory and the finite-difference time domain (FDTD) method, which was deemed to be out of scope for this thesis.

Waveguide Dimension Analysis

To predict the performance of the waveguide design, the *conversion efficiency* η, as given in Equation (3.3), was chosen as the figure of merit. The dependence of η relative to the gap and width of the waveguide slot is plotted in Figure 4.5. Two waveguide heights were considered. 220 nm, which is the transmission height commonly used in integrated photonics, and 140 nm, the height used in this thesis. η was evaluated at the fundamental TE mode (if multiple modes were present), within a few nanometers of the waveguide surface. The simplified expression for η was used, meaning that the Stokes shift was omitted. Figure 4.5 shows that η increases significantly for a smaller gap. This is due to an increased modal confinement in the gap [13]. The conversion efficiency is nearly constant for the 140 nm high waveguide, while it increases for the 220 nm waveguide.

An expanded dimension analysis is displayed in Figure 4.6, where both the gap g and the rail widths r are varied simultaneously. It is shown that the conversion efficiency η increases with a narrow gap and thin rails, while the loss appears to increase as the rail width r decreases. In the waveguide design, the number of modes is best controlled with the rail width r, while the conversion efficiency η is best controlled with the gap g. Note that g_0 will be relatively large, such that the waveguide platform will be compatible with deep UV lithography (DUVL), which will be discussed in Section 6.1. In the waveguide design, a rail width of 500 nm and a gap of 150 nm are targeted. This is indicated by the red lines in Figure 4.6. The gap will then be narrowed with a spacer layer, which is considered next.

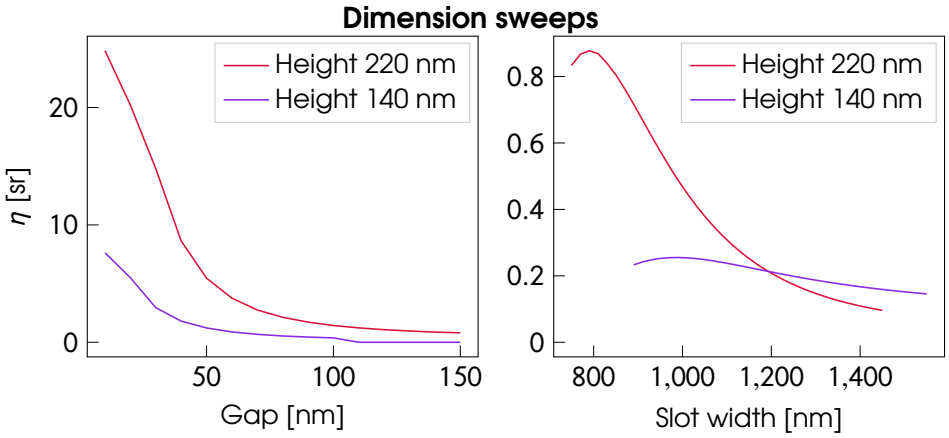


Figure 4.5: Conversion efficiency η depending on the gap and width dimensions of the slot, at two heights (220 nm and 140 nm). η was evaluated at the fundamental TE mode and at the pump wavelength λ_p = 785 nm. For the gap calculations, the rail width was 350 nm, and for the width calculations the gap was 150 nm. Clearly, the dominant contribution to η is from a small gap. The relation between slot width and rail width is: slot width = 2 × rail width + gap.

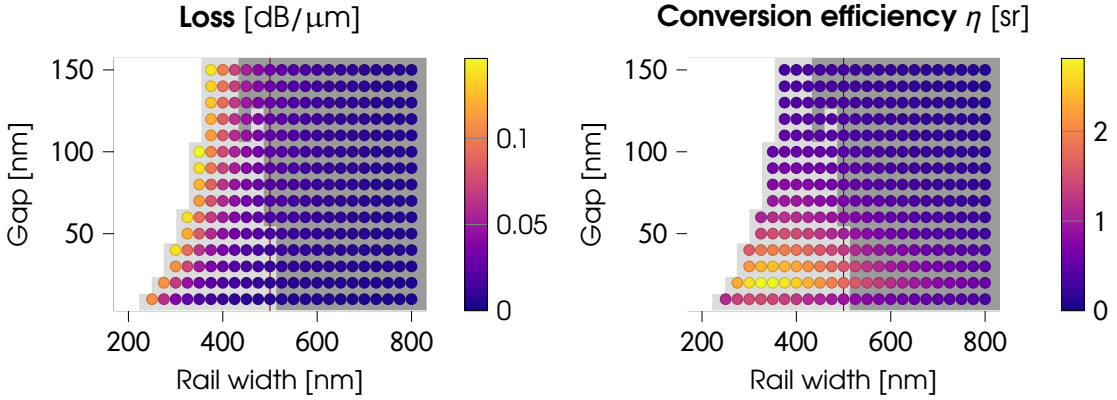


Figure 4.6: 2D plots of the absorption loss in dB/μm (left) and the conversion efficiency η (right) calculated for a given rail width r and gap g of a slot Si_3N_4 waveguide platform. Each circle represents a simulation point. The pump wavelength λ_p = 785 nm was evaluated, and the waveguide height was set to 140 nm. The number of modes present in the waveguide is indicated by the shade of grey. Light grey = 1 mode, and dark grey = 2 modes. In the white regions, no modes are present. The number of modes increases with a narrow gap and wide rails, however no modes are guided when the rail width approaches 200 nm. Similarly, both the loss and conversion efficiency η increases when the slot's dimensions are reduced. The loss at the lines r = 450 nm, r = 500 nm, and r = 550 nm are shown in Figure 5.2. The r = 500 nm lines are highlighted in red.

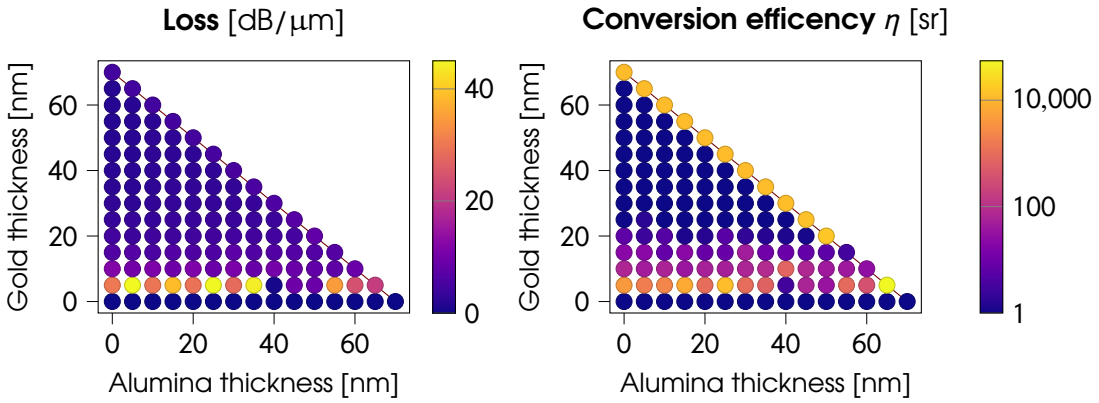


Figure 4.7: 2D plots of the absorption loss in dB/μm (left) and the conversion efficiency η (right) for varying thicknesses of the Au and Al_2O_3 layers. Note that η is a logarithmic plot. The slot gap, the rail widths, and the height was fixed at $g = 150$ nm, $r = 500$ nm, and $h = 140$ nm, respectively. The fundamental TE mode was evaluated at the pump wavelength $\lambda_p = 785$ nm. Clearly, η is the largest when the final gap is narrow, $g \sim 10$ nm. This occurs along the marked diagonal. A higher resolution plot around this gap g is shown in Figure 4.8. η is only larger than unity when the gold layer is present. The small η for gold layers thicker than 20 nm, may be explained by an increased absorption in the plasmonic material [19].

Plasmon and Spacer Thickness

To reduce the gap size of waveguides fabricated using photolithograpghy, A. Raza *et al.* [19] proposed using Al_2O_3 deposited on the waveguide with ALD as a spacer layer to narrow the gap. In Figure 4.7 it is shown how the conversion efficiency η would behave with the inclusion of both alumina and gold on top of the waveguide. The inclusion of gold shows an significant increase in η compared to the plasmon free waveguide considered in Figure 4.5. The magnitude of η reaches its maximum $\sim 10^4$ when $t_{Au} + t_{Al2O3} \approx 70$ nm, reducing the final slot gap to only 10 nm. These alumina and gold thicknesses around this narrow gap are shown in Figure 4.8. There, it is shown that a narrow gap 6 nm $\leq g \leq$ 12 nm is required to get a large η, and that η is nearly constant for a given gap g. Of course, this does not hold anymore when the plasmon layer is too thin to give a significant enhancement, which occurs when $t_{Al2O3} \approx 55$ nm, such that $t_{Au} \leq 9$ nm. Up until this alumina thickness, a large η is obtained. Therefore, as stated by A. Raza *et al.* [19], it is sensible to target a relatively thick alumina layer, since its thickness is well-controlled with ALD.

Slot Width Analysis

Up until this point, only the *active region* of the waveguide has been considered, i.e., the section of the waveguide where Raman interaction should take place. However, large parts of the waveguide will consist of section of different dimensions, such as tapers, and input and output regions. The final design will be described in detail in Section 5, and the total length of a single waveguide may be on the order of several millimeters. The active region, on the other hand, will only be some tens of microns. Therefore, it is useful to consider the losses for other waveguide dimensions as well. The losses of a taper were calculated, its width w ranging from $w_{min} = 2r + g = 2 \times 500$ nm + 150 nm = 1150 nm to $w_{max} = 5$ μm. This means that

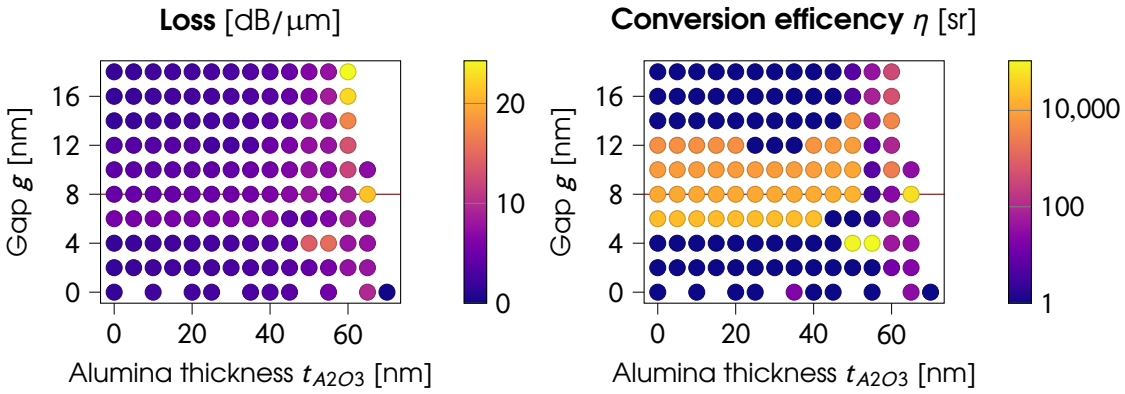


Figure 4.8: 2D plots of the absorption loss α (left) and conversion efficiency η (right) for different layer thicknesses, close to the $t_{Au} + t_{Al2O3} \sim 70$ nm line in Figure 4.7. The y axes denote the final gap $g = 150 - 2(t_{Au} + t_{Al2O3} + t_{NTP})$, where $t_{NTP} = 1$ nm. The horizontal lines in this figure, correspond to the diagonal lines in Figure 4.7, and the red line in Figure 4.7 is marked in red here as well, as $g = 10$ nm.

the absorption losses are calculated for all taper dimensions, from the width at the input coupling, to the dimensions of the active region. The loss turns out to be highly dependent on the slot width, as shown in Figure 4.9. As the slot width decreases, the waveguide becomes single mode, and the losses increase exponentially. The Raman signal will be measured in transmission, as described in Section 6.5. Thus, away from the active region, losses should be as small as possible to preserve the signal. Hence, to preserve the signal in the waveguide, the active region should not be longer than necessary, and the waveguide width should be increased outside of the active region.

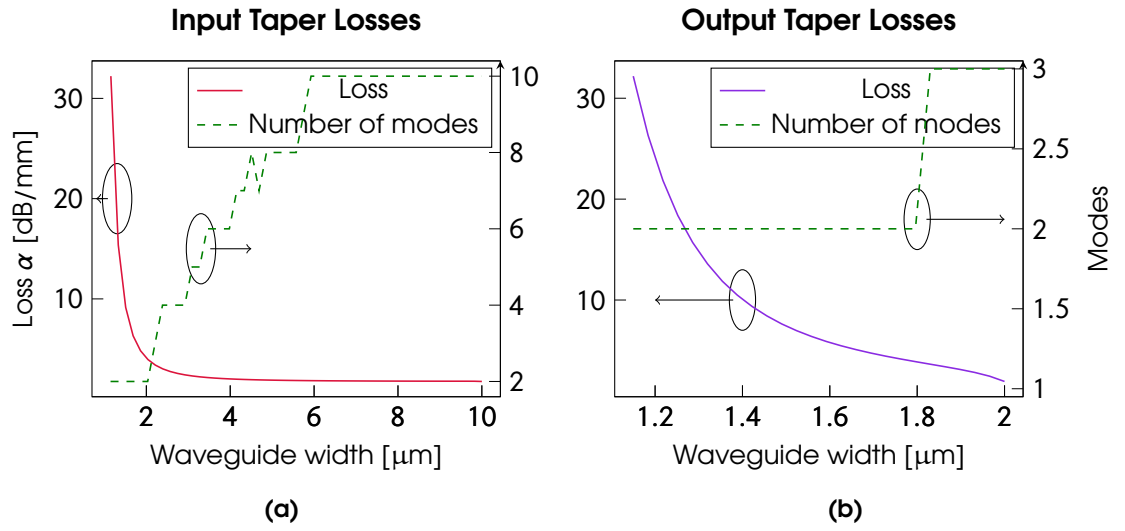


Figure 4.9: Calculated absorption losses in the taper regions. Both the input taper (a) and the output taper (b) are shown, as well as the number of modes. The loss of the first TE mode is given at the pump wavelength $\lambda_p = 785$ nm. The waveguide height was 140 nm. For both tapers, the smallest slot width refer to a gap $g_0 = 150$ nm, and rail widths $r = 500$ nm. The gap is narrowed down until it disappears at the largest slot width.

This concludes the prediction of the waveguide performance and losses, depending on its dimensions (rail width r and gap g), and the thicknesses of the plasmon layer t_{Au} and the spacer layer t_{Al2O3}. This will make it possible to design a waveguide platform with maximum conversion efficiency of $\eta \sim 10^4$ in the active region, and minimal loss away from it, within the fabrication limits imposed by for instance lithography. The final waveguide design is described in detail in Section 5.

Chapter 5

Waveguide Design

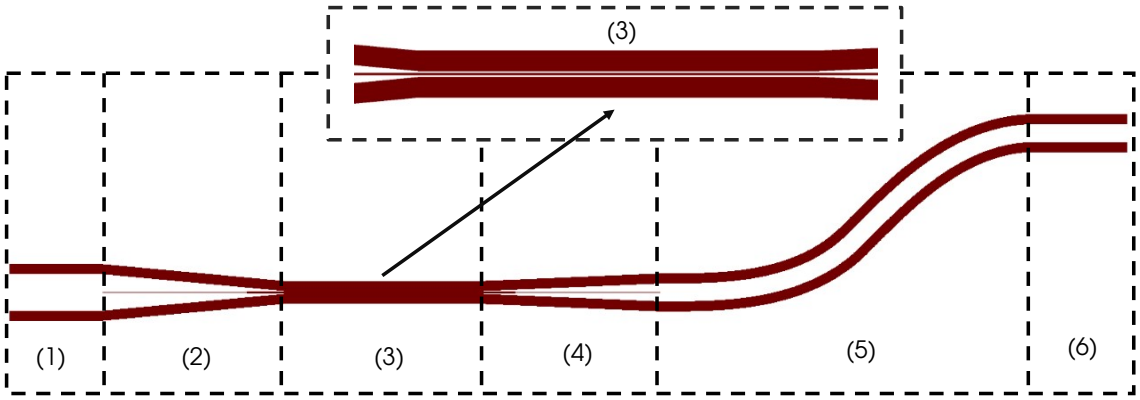


Figure 5.1: Layout of the final waveguide design. It consists of six different regions, as listed in Table 5.1. The dimensions are **not** to scale, as the waveguide length is compressed to display all six regions simultaneously. (1) Input, (2) input taper, (3) active region, (4) output taper, (5) Euler bend, (6) output. Inset: magnified sketch of the active region. This is an etch mask, meaning that the red areas correspond to the parts of the SiN thin film which is etched away. Thus, light can be guided in the white areas. The waveguide structure itself is the white areas between the red lines.

Based on the theory presented in Sections 2 and 3, and the simulation results in Section 4, a suitable waveguide design was chosen for fabrication. Note that this design is intended for transmission measurements, where the pump light is coupled in from free space at one end of the waveguide, and coupled out to free space at the other end, before being collected at the Raman spectrometer. The layout of a single waveguide consists of six distinct regions. Their names and purpose listed from input to output are as follows:

- (1) *Input*, the entrance to the waveguide where light is coupled in from free space. The waveguide is wide here, and there is no gap, in order to effectively couple an incident laser beam. This section must be scribed and broken to allow direct coupling from free space. Its lengths should be as short as possible to minimize absorption losses.

- (2) *Input taper*, where the dimensions of the waveguide are tapered down from wide at the input, to dimensions suitable for Raman interaction. The waveguide must be tapered over a distance sufficient to ensure a stable beam and reduce scattering losses at the surface.

- (3) *Active region*, here, Raman interaction should take place. The plasmonic layer and the alumina spacer layer will be deposited here. The length of this region is limited by losses, as the waveguide is designed to work in transmission.

Table 5.1: Overview of the waveguide regions and their dimensions in the longitudinal plane. The height is determined by the etch, and described in Section 6.3. Note that the length and width of the active region were varied from waveguide to waveguide. The maximum and minimum values are included here. The lengths of the input (1) and output (6) regions are labelled as zero, since they should be made as small as possible to reduce absorption losses. In reality, their lengths are on the order of millimeters, to make it easier to scribe and break each waveguide chip during fabrication.

	Region	Length [μm]	Width [μm]
(1)	Input	0	10
(2)	Input taper	600	varies
(3)	Active region	10 — 100	0.850 — 1.150
(4)	Output taper	300	varies
(5)	Euler bend	100	2
(6)	Output	0	2

- (4) *Output taper*, here the waveguide dimensions are increased again to reduce losses and increase the numerical aperture of the waveguide when it is coupled out of the waveguide.
- (5) *Euler bend*, a bend which shifts the input and output beams laterally in order to avoid interference, increasing the signal-to-noise ratio (SNR).
- (6) *Output*, the output from the waveguide. Light is coupled back into free-space. Similarly to the output region, this section must be scribed and broken, and should be as short as possible.

In summary, the only region important for Raman interaction is the active region, while the tapers contribute to reduced losses and simplified coupling. The Euler bend prevents scattered light at the input from contaminating the output light. The dimensions of the aforementioned regions are listed in Table 5.1.

First, the dimensions of the active region were evaluated. To make this design compatible with standard photolithography processes suitable for industrial production, the gap g_0 was set to 150 nm. The narrowing of the gap would then be achieved with an alumina spacer layer. A suitable rail width of $r = 500$ nm was determined from Figure 4.6. At this r, the waveguide is double mode for a 150 nm gap, and becomes single mode when the gap is 50 nm or less. Assuming that the narrowing of the gap $g_0 = 150$ nm can be achieved with a spacer layer of alumina, a gap size of 50 nm would be achieved with a 50 nm thick Al_2O_3 layer. In that case the waveguide is single-mode with a 28 dB/mm absorption loss. This is shown in Figure 5.2. The loss at $r = 500$ nm increases from 6.5 dB/mm when $g = 150$ nm, to 32 dB/mm when $g = 10$ nm. At $g = 50$ nm, the loss is 28 dB/mm.

The next task to consider was to ensure that light is successfully guided into and out of the active region, by the use of tapers. Common NTNU NorFab practice is to make tapers about 600 microns long, when used to connect free space input couplings, and single mode waveguides, similar to this design. Therefore, the length of the input taper was set to 600 μm, while the output taper was set to a length of 300 microns. It was made shorter since the final width of the output taper is 1/5 of the width of the input taper. The absorption losses associated with each taper can be estimated from Figure 4.9 by integrating the graphs for the lengths of the tapers

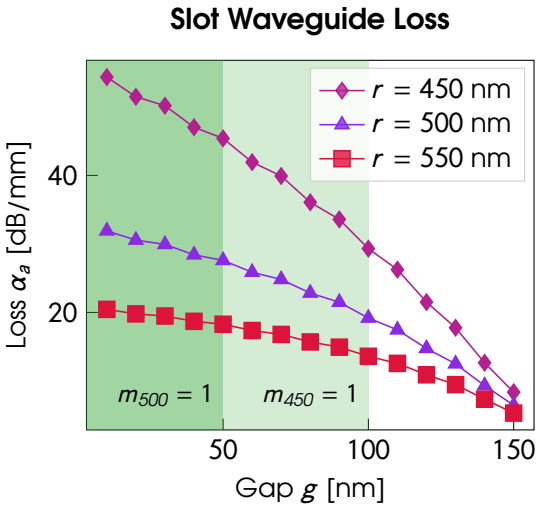


Figure 5.2: Absorption loss α_a of a plain SiN slot waveguide plotted against the gap size g at selected rail widths r. The data points are marked, and the lines serves as a guide to the eye. The data is the same as in Figure 4.6. The number of modes m_r for a given rail width r is indicated by the shade of green. When $r = 550$ nm, the waveguide guides at least two modes for all gap sizes. In the light green region $r = 500$ nm becomes single mode when $g \leq 50$ nm, and in the dark green region $r = 450$ nm becomes single mode when $g \leq 100$ nm.

(600 microns for the input taper, and 300 for the output). This results in absorption losses of $\alpha_{t,in} = 1.70$ dB and $\alpha_{t,out} = 2.57$ dB, for the input taper and the output taper, respectively.

The Euler bend was included to shift the input and output beams, and a lateral shift of 100 microns was deemed sufficient. This corresponds to the approximate width of the pump beam in our experimental setup, see Figure 6.8. To achieve this shift, two Euler bends in series with bend angles of 30° and -30° and bend radii of 50 µm was included in the design. The optical length of this region is 100 microns, and its losses may be estimated from Figure 4.9, in the same fashion as for the output taper. The absorption loss of the Euler bend is $\alpha_{Euler} = 0.15$ dB.

Finally, the input and output regions were included for fabrication purposes. As mentioned, their inclusion is not necessary for waveguiding purposes. This is because light can be coupled directly into the input taper, and exit directly from the Euler bend. Instead, these regions are included for ease of fabrication and treatment in the lab. When a waveguide chip is fabricated, it is not feasible to pattern the chip edges. Consequently, the chip must be cut (or scribed) to expose the coupling regions. A high cutting accuracy enables short input and output regions. In this thesis however, the scribing was done by hand, and total chip length would be about 1 cm, meaning that the input and output regions constitute about 90 % of the chip length. An image of the scribing process is shown in Figure 5.3. Once again using Figure 4.9, setting the average loss of the input and output to $\alpha_{in,out} = 1.85$ dB/mm, and assuming the length to be 9 mm, estimates the loss to 16.6 dB.

A summary of the estimate of each region's absorption loss contribution given in this section is found in Table 5.2. Clearly, the largest contribution originates from the input and output. The loss from the active region is sensitive to the inclusion of the alumina and gold layers, and since the measurements are in transmission, it is

Table 5.2: Estimated absorption loss contribution from each waveguide region. The tapers and the Euler bend contribute to 4.42 dB of the loss in all configurations. The contributions from the input, output, and the active region may vary depending on the cutting of the chip, and the length of the active region. If the total chip length is 1 cm, their loss from the input and output is 16.6 dB. The loss from the active region strongly depends on the gap size, as shown in Figure 5.2. The loss listed here is an average value. If the active region is 100 µm, the loss is 2.15 dB. For such a chip, the total absorption loss amounts to 23.2 dB. Scattering losses and plasmonic effects are ignored.

	Region	Absorption loss
(1)	Input	1.70 dB/mm
(2)	Input taper	1.70 dB
(3)	Active region	~ 21.5 dB/mm
(4)	Output taper	2.57 dB
(5)	Euler bend	0.15 dB
(6)	Output	1.93 dB/mm

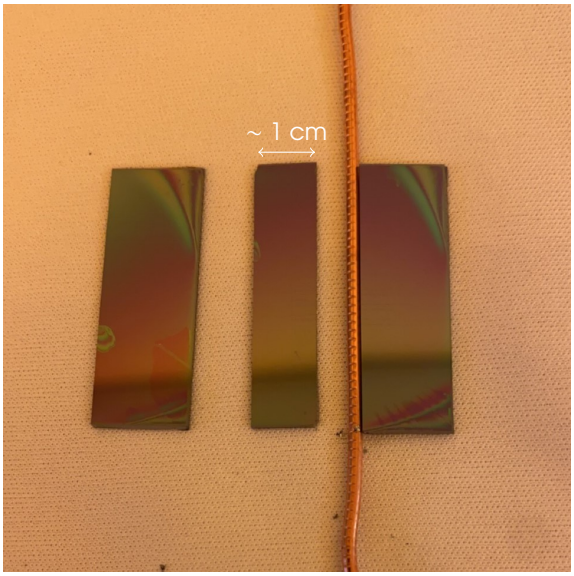


Figure 5.3: Image after manual scribing of one of the chips fabricated in this thesis. The chip containing the waveguides is the part in the middle. The waveguide structures are not visible in this image, but their lengths lie in the horizontal direction. The width of the chip is about 1 cm. Inputs and output regions are present on the left and right parts of the sample, respectively. Prior to scribing, photoresist is applied to protect the pattern. The resist is visible as the green shading at the sample edges.

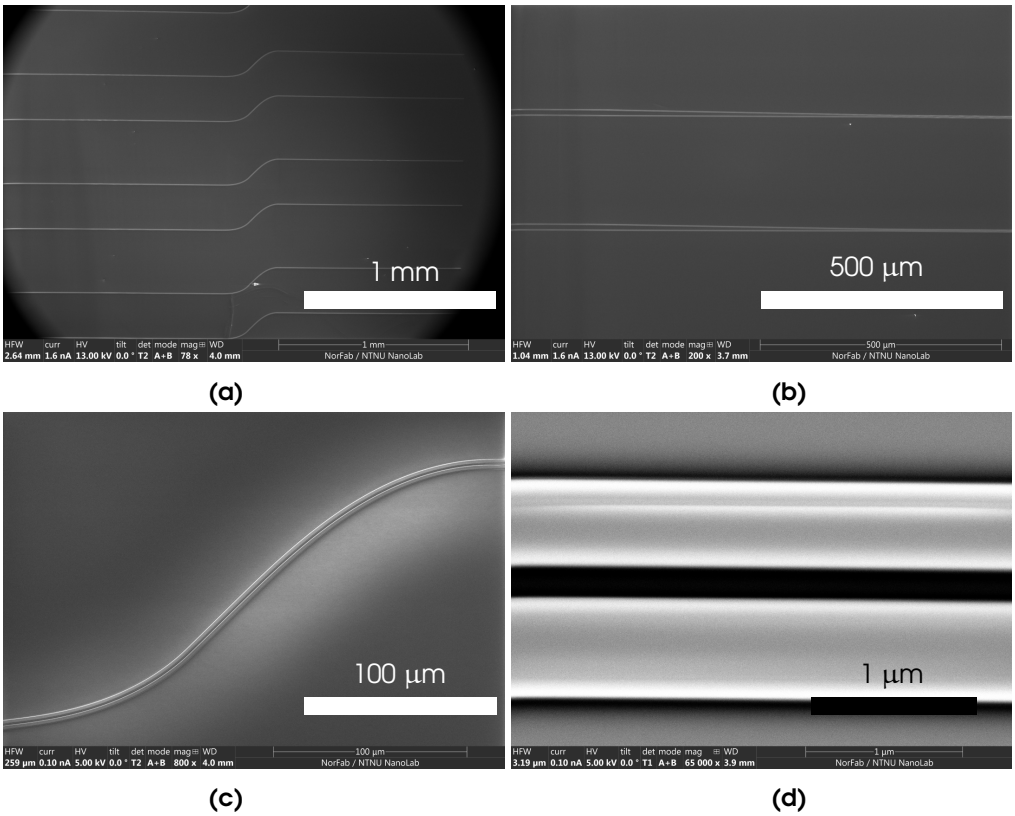


Figure 5.4: Top-view SEM images taken at the NTNU NorFab facilities, showing the waveguide structure resulting from the design described in this section. (a) An overview of several waveguides, with the active region and Euler bends. (b) The end of two input taper regions. (c) A single Euler bend. (d) The slot waveguide in the active region, with rail widths of 500 nm and a gap of 150 nm.

important that a sufficient amount of light survives passing through the waveguide. Although the pump laser is unwanted at the detector, the Raman signal generated in the active region is subject to the same attenuation as the laser beam. Hence, it is reasonable to assume that if the laser is extinct, so is the Raman signal. SEM images of the platform fabricated from this design are displayed in Figure 5.4.

In the above loss considerations, the scattering loss was ignored. However, in high index contrast (HIC) waveguides, such as this SiN platform, the dominant loss does indeed arise from scattering losses at the surface [11], [15]. This is due to surface or sidewall roughness, and limited through good fabrication techniques. The sidewall roughness is of particular importance in this design, since the slot waveguides are designed to be single TE mode. As mentioned in Section 3.2, TE modes overlap with the sidewalls.

Consequently, the above loss analysis is somewhat limited, and dependent on fabrication. The loss calculations only serve as a guideline, and the main point to take away is that the longitudinal direction should be reduced where possible, in particular at the input and output. Additionally, losses from the plasmonic layer are omitted. The Raman signal will be enhanced up to a certain plasmonic layer length, when attenuation in the gold exceeds the enhancement [19]. The optimal length of the plasmonic layer should be verified experimentally.

The Python code used to create the waveguide masks is available in Appendix

D, and on GitHub [32]. Images of the waveguides resulting from this design are shown in Figure 5.4. Additional images are shown in Appendix A.

Chapter 6

Fabrication of a Silicon Nitride Waveguide Platform for SERS Measurements

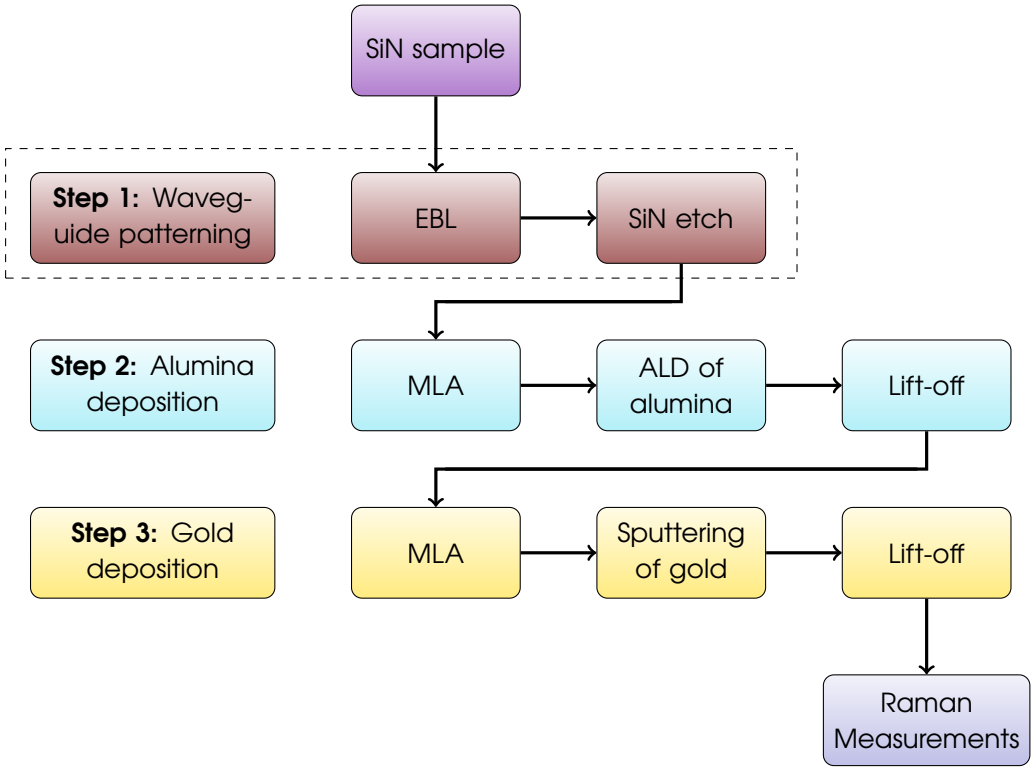


Figure 6.1: The process flow chart consisting of three consecutive fabrication steps labelled steps 1 to 3, and one measurement step. The abbreviations stand for: electron beam lithography (EBL), maskless aligner (MLA), and atomic layer deposition (ALD). In the processing, only the first step was completed, which is why it is highlighted by the dashed box.

In this thesis, the fabrication of a waveguide platform for surface-enhanced Raman scattering (SERS) sensing is described, involving the patterning of silicon nitride (SiN) waveguides and subsequently coating them with Al_2O_3 and Au. The fabrication process can be viewed as a three step process, involving one waveguide patterning step and two metallization steps. The process flow is outlined in Figure 6.1, and will be discussed in detail in the following sections.

Firstly, in Section 6.1, follows an overview and short explanation of the processes used in the fabrication of the waveguide platform, including lists of the instruments and chemicals used. Next, in Sections 6.2 to 6.4, the different processing steps with the relevant parameters are presented. Any reader already familiar with the processing techniques presented here, should feel free to skip the introductory section 6.1, and go directly to Sections 6.2 to 6.4 for the experimental details. This chapter closes with Section 6.5, an overview of the measurement setup.

6.1 Overview of Fabrication Methods

As mentioned in Section 3.3, deep UV lithography (DUVL) has been successfully employed to create slot waveguides with dimensions similar to those in this thesis. However, e-beam lithography (EBL) was employed here instead, as it is the sole lithography tool available in NTNU, NorFab capable of producing features smaller than 1 micron. The main distinction between DUVL — and photolithography in general — and EBL lies in their respective speed and resolution. EBL is slower but can pattern smaller structures, whereas the minimum dimensions in photolithography are limited by the wavelength of the light source. The typical resolution of DUVL is 50-100 nm [19]. In contrast, the diameter of an electron beam can be reduced to a few nanometers, allowing for smaller resolution. Due to the small number of samples produced in this research project, the slow writing speed of EBL is not an issue. However, for industrial applications, DUVL would be preferred. Although EBL was used in this thesis, the smallest features patterned in this design are 150 nm, making this platform compatible with DUVL as well.

After the lithography step, parts of the sample are left resist-free. These exposed areas of the SiN thin film can be etched away to create the waveguide pattern. Etching involves introducing reactive molecules to the sample surface, which degrade it. It is important for the waveguide cross-sections to be as rectangular as possible, which means that the etch should be anisotropic, and directed parallel to the waveguides' sidewalls. This is usually achieved with a *dry etch*, where the surface is degraded from reaction with gas molecules. When the gas is an ion plasma, a voltage bias at the sample surface can increase the directionality of the plasma, making the etch more anisotropic. Such an etch was employed in this thesis, using an inductively coupled plasma-reactive ion etch (ICP-RIE), with CF_4 and O_2 as the reactive gases. The etch rate of our system was found to be about 120 nm/min for Si_3N_4.

Deposition of the alumina spacer layer and the plasmonic gold layer was not performed in this thesis. Had it been performed, the metal would be deposited on parts of the waveguides using lift-off. The dimensions of the metal layers are on the order of microns, so the use of EBL is not necessary. Instead, a maskless aligner (MLA) may be used. An MLA is a lithography tool which writes a pattern with a laser source, but without the use of masks. The deposition of alumina may be achieved with atomic layer deposition (ALD), and the deposition of gold with sputtering or e-beam evaporation. The necessary instruments are readily available in the NTNU NorFab facilities.

The following instruments were used, all of which are located at NTNU in the NorFab cleanroom facilities:

- *ELS-G100* from Elionix. Electron beam lithography system
- *Plasmalab System 100 ICP-RIE 180* from Oxford Instruments. Inductively coupled plasma-reactive ion etch system
- *Femto* from Diener Electric. Plasma cleaner
- *MLA 150* from Heidelberg. Maskless aligner, laser write tool. Not employed in this thesis as lift off was not performed.

- *Savannah S200* from Veeco. Atomic layer deposition system. Not employed in this thesis, as alumina deposition was not performed
- *Custom ATC-2200V* from AJA International Inc. E-beam evaporator and sputtering system. Not employed in this thesis, as gold deposition was not performed.

Although some of these instruments were not employed in this thesis, they are included to highlight the fact that they are necessary to complete the development of the waveguide platform designed in this thesis. In combination with these instruments, the following chemicals were used:

- *AR 600-549* from Allresist. Positive e-beam resist,
- *AR 600-71* from Allresist. E-beam resist remover,
- acetone, isopropyl alcohol, and deionized water for cleaning, and
- various process gases in instruments, most notably CF_4, O_2 and N_2.

Finally, the instruments used to characterize the waveguide platform during and after fabrication were:

- *F 20 Reflectometer* from Filmetrics. Used to measure the thicknesses of the Si_3N_4 and resist thin films.
- *Dektak 150 Profilometer* from Veeco. Used to measure the etch depth of the SiN thin film.
- *SEM APREO* from FEI. Used to image the waveguide structures, both in top-view and in the cross-section.
- An optical microscope to monitor the quality of the fabrication *in-situ*, such as the resist quality, successful development, successful resist removal, etc.

6.2 Fabrication of Silicon Nitride Samples

The wafers required in this thesis, $Si_3N_4/SiO_2/Si$, were not available at the NTNU NorFab facilities, and would have to be made in-house or ordered from a manufacturer. Silicon nitride thin films may be deposited using plasma enhcanced chemical vapour deposition (PECVD) or low pressure chemical vapour deposition (LPCVD). Originally, attempts were made to deposit SiN thin films using the in-house PlasmaLab System 100-PECVD from Oxford Instruments. The optical performance of these thin films is unknown. Although Si_3N_4 thin films for optical purposes have been deposited with PECVD [33], [34], LPCVD are known to have excellent optical properties [35]. No LPCVD system was available at the NTNU NorFab facilities. However, SINTEF (MinaLab at Gaustadalléen 23C, 0373 Oslo) were able to provide the desired SiN wafers. The wafers were fabricated in the following fashion: Starting from silicon wafers, they underwent a thermal oxidation process to grow the oxide layer, followed by the deposition of Si_3N_4 using low pressure chemical vapour deposition (LPCVD). According to SINTEF, the thickness of the oxide layer was 2.4 μm, and the thickness of the silicon nitride layer was 140 nm. All experimental results displayed in

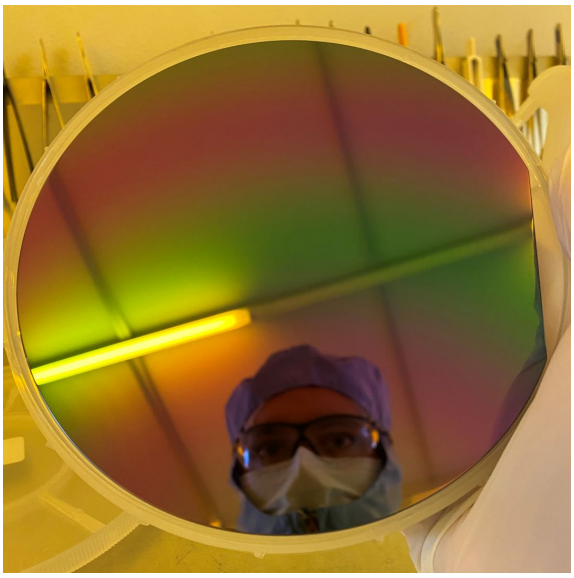


Figure 6.2: One of the 6" SiN wafers provided by SINTEF. The apparent color of the wafer is highly sensitive to the angle of the incident light, due to the very thin silicon nitride layer.

this thesis were obtained using these wafers. One such wafer is displayed in Figure 6.2. The wafers arrived on the 12th of May, meaning that all fabrication described in this section was performed within the span of a few weeks, up until the thesis deadline on the 12th of June.

6.3 Waveguide Patterning by EBL and Dry Etch

The waveguides are patterned in two steps, starting from a cleaned SiN sample. Firstly, the etch pattern is created by e-beam lithography, subsequently the exposed parts of the SiN thin film are etched away by an inductively coupled plasma-reactive ion etch. The lithography process consists of six steps, the etching is done in a single step, and finally the resist should be removed and the sample cleaned, resulting in a total of nine steps, as displayed in Figure 6.3. A summary of the most important process parameters is found in Table 6.1. The details of each step are described in the following.

Electron Beam Lithography (EBL)

A six step lithography process was employed in this thesis, as shown in steps 1 to 6 in Figure 6.3. Steps 1 to 3 involves applying resist on the samples. First, the samples were cleaned, and then dehydration baked on a hot plate at 150 °C for 5 minutes. This is to remove water molecules that may be present on the sample, in order to improve resist adhesion. Next, a positive e-beam resist (AR-P 6200) was spin coated onto the sample, with a spin time of 60 seconds at 4000 rpm and 1000 rpm/s acceleration. Reflectometer (F20, Filmetrics) measurements confirmed a resist thickness of about 440 nm for our chip samples, however the resist datasheet [36] should be consulted before determining the spin parameters. The samples were subsequently soft baked at 150 °C for 1 minute. This decreases the viscosity of the resist by removing excess water, and after this step the sample is ready for exposure.

Step 4 was the exposure, performed in the "Elionix ELS-G100", a 100 kV EBL system.

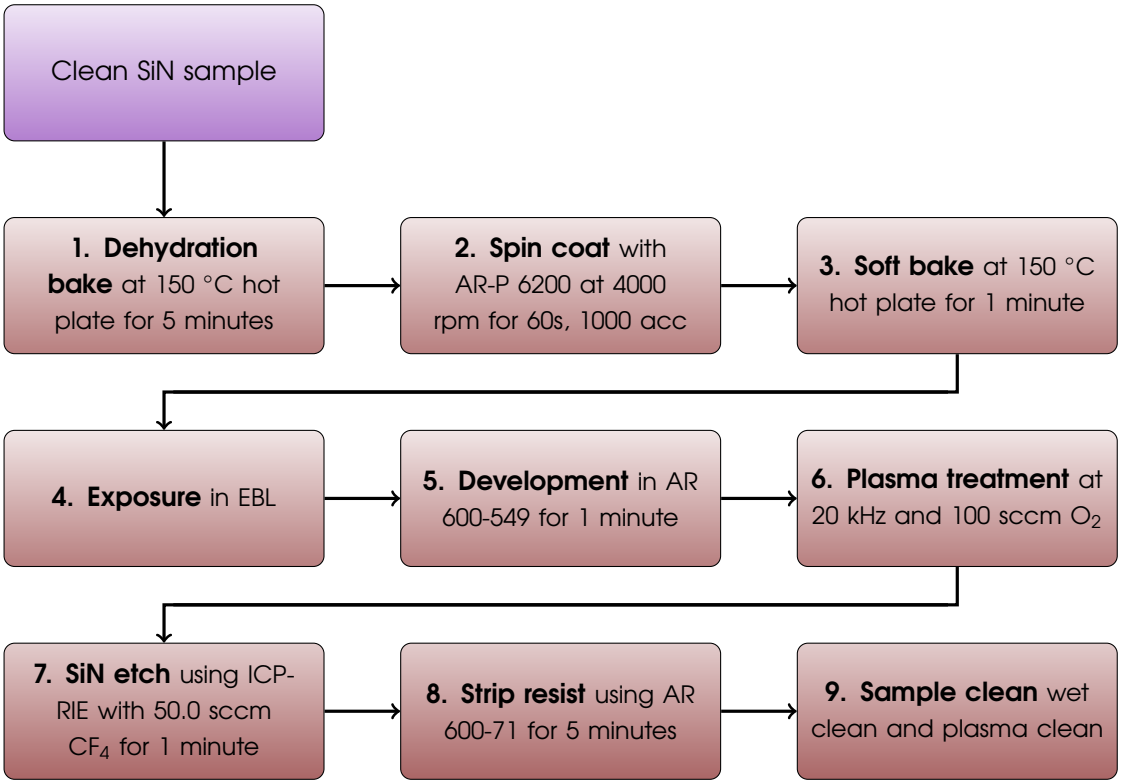


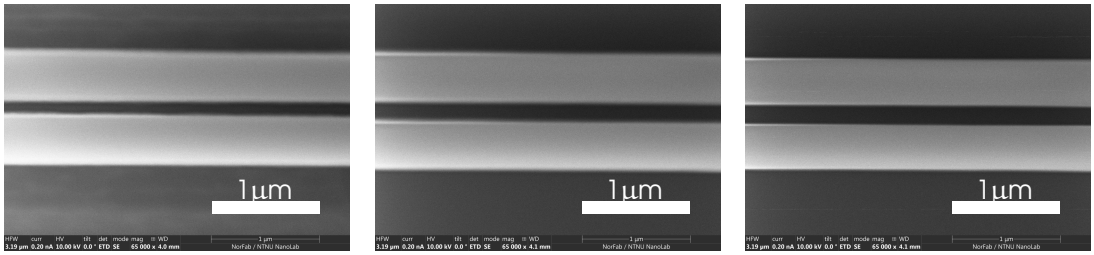
Figure 6.3: Flow chart of the nine step waveguide patterning process, from a clean $Si_3N_4/SiO_2/Si$ sample to SiN waveguides. The first six steps are lithography steps, while the final three steps are the etching of the thin film, removal of the resist, and cleaning of the sample. The parameters stated in the chart are intended for the materials and instruments employed in this thesis, and may be adjusted if desirable.

The beam current was set to 2 nA, and the objective aperture to 120 μm, resulting in a beam diameter of 3 nm. The write field size was set to 500 μm, with 1,000,000 dots. This results in a minimum pixel resolution of 0.5 nm, but the pitch size was set to 7, resulting in an actual pixel diameter of 3.5 nm. The electron area dose was set to 170 $\mu C/cm^2$, resulting in a dose time of 0.01041 μs pixel. These parameters are summarized in Table 6.1. A brief rundown of the parameters is given in the next paragraph. A detailed explanation of the parameters and recommendations are found in the Elionix manual. The manual is not available to the public, but is provided at the NTNU NorFab facilities.

The minimum beam diameter is determined by assigning the beam current. A small beam can write smaller features, but at an increased scan time. A beam current of 2 nA, resulting in a 3 nm beam diameter, is relatively high, but since the waveguides designed in this project are mostly straight, and the smallest feature (the gap) is 150 nm, speed was preferred over accuracy. Assigning the write field divides the exposure pattern into distinct grids. The sample stage is moved between write fields, which is a possible source of *stitching errors*, visible as misalignment of features at write field borders. If the write field is large there will be less stitching, but an increased write field area has two other unwanted effects. The first is lower magnification, decreasing the beam focus, and the second is that the beam will expose the write field edges at a slight angle. In general, a write field side length of 500 μm is a common choice, and it was the only field size selected in this thesis. Next, the write field's pixel size is determined by the number of dots and the pitch. The minimum pixel diameter is determined by the number of dots, as 500 μm / 1,000,000 = 0.5 nm. This is however, much smaller than the minimum beam diameter (3 nm in our case). Such a small pixel size only serves to increase scan time. A pitch of 7 means that only every seventh pixel is exposed, resulting in an effective pixel size of 0.5 nm × 7 = 3.5 nm. This is slightly larger than the minimum beam diameter. Ideally, the pixel diameter should be 80 % of the beam diameter, to ensure as smooth edges and as fast scan time as possible. However, the fact the accuracy requirements in this design is relaxed, allows for a larger pixel size. Finally, the dose electron area dose is set to 170 $\mu C/cm^2$. This results in a dose time of $(3.5\ nm)^2 \times 170\ \mu C/cm^2$ / 2 nA = 0.01041 μs per pixel (or dot). The minimum dose time of this system is 0.01000 μs/dot, so this dose is acceptable.

The ideal area dose was determined from experimental dose testing, which means exposing the same pattern with different doses, and visually observing how the resolution of the resulting features changes. A plot of the dosetest is shown in Figure 6.5, and SEM images of three different doses are displayed in Figure 6.4. From these result, a dose of 170 $\mu C/cm^2$ was deemed ideal. More SEM images at this dose are shown in Figure 5.4 and Figure 6.6. None of the images display stitching errors or low resolution edges. This shows that the EBL parameters selected in this thesis were acceptable to create well defined SiN waveguides.

After exposure, the samples were developed in AR 600-549 developer for 1 minute (step 5). The final lithography step (step 6) was a plasma clean with the "Femto Plasma Treatment Machine" from Diener Electronic, with O_2 gas at 50 % gas flow and 50 % generator power for 1 minute. This corresponds to a plasma frequency of 20 kHZ and a flow rate of 100 sccm. This step is aimed at improving the vertical lines in the resist. Following the completion of this step, the sample is ready for dry etch, to create the waveguide pattern in the SiN thin film.



(a) $d = 120$, $g = 130$, $r = 390$ **(b)** $d = 170$, $g = 150$, $r = 380$ **(c)** $d = 200$, $g = 150$, $r = 340$

Figure 6.4: Top-view SEM slot waveguide images taken at NTNU NorFab after etching. Different electron doses are shown, but the same lithography mask was used in all three instances. The images are labelled according to electron dose d [µC/cm^2], gap g [nm] and rail width r [nm]. (b) shows the dose selected for further fabrication, 170 µC/cm^2. Figure 6.5 plots the relation between dose and dimensions.

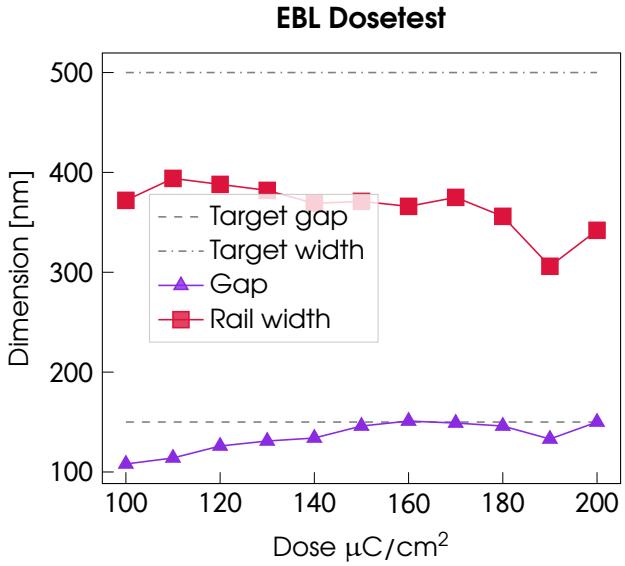


Figure 6.5: Relation between dose and slot dimensions when using a lithography mask as described in Section 5, where the mask gap is 150 nm and the rail width 500 nm. Measurements are made after etching, from inspection of SEM images. The dose number denotes the number of electrons incident on a single dot (grid point). Low doses result in a narrow gap, while higher doses slightly reduce the rail width. Images of selected doses are displayed in Figure 6.4. Other relevant EBL parameters are listed in Table 6.1.

Table 6.1: Overview of the process parameters in the lithography step to pattern the SiN waveguides. Note that a dehydration step should precede the spin coating step, including cleaning if necessary. Between exposure and etch, it is necessary to develop the sample in AR 600-549 for 1 minute, and the sample should be subject to a short plasma clean. The parameters listed here may be sensitive to for instance sample size, and should always be verified. Remember to consult the relevant datasheets.

Spin coat		EBL Elionix		ICP-RIE Chiller	
Parameter	Value	Parameter	Value	Parameter	Value
Resist	AR-P 6200	Voltage	100 kV	Temperature	20°C
Velocity	4000 rpm	Beam Current	2 nA	Pressure	5×10^{-5} Torr
Acceleration	1000	Aperture	120 µm	CF_4 flow	50.0 sccm
Time	60 s	Beam Diameter	3 nm	O_2 flow	5.0 sccm
Soft bake	150 °C, 1 min	Write Field	500 µm	Etch time	70 s
		Dot Number	1,000,000	Etch rate	~ 120 nm/min
		Dose	170 µC/cm^2		
		Pitch Size	7		
		Dose time	0.01041 µs/dot		

Dry Etch Using Inductively Coupled Plasma-Reactive Ion Etch (ICP-RIE)

Following exposure and development, a dry etch method was used to pattern the SiN thin film. The "Plasmalab System 100 ICP-RIE 180" from Oxford Instruments was used. Inductively coupled plasma-reactive ion etch (ICP-RIE) is an anisotropic dry-etching process using a chemically reactive ion plasma, which is directed towards the sample surface with a DC bias.

The SiN etch (step 7) was performed at 20 °C and vacuum (5×10^{-5} Torr). The process gases were CF_4 and O_2 at 50.0 sccm and 5.0 sccm, respectively. From profilometer (Dektak 150, Veeco) measurements, the etch rate of the silicon nitride samples in this thesis was found to be about 120 nm/min. To etch through the silicon nitride layer, the samples were etched for 70 seconds. The etch parameters are listed in Table 6.1. Figure 6.6 shows the cross-section of a waveguide after etching, where the desired etch depth is reached. After etching, the resist was removed (step 8) with "AR 600-71", and cleaned. Finally, any residue resist was removed with a plasma clean (step 9), this time with 100 % generator power and 100 % O_2 gas flow for 10 minutes, i.e., 40 kHz plasma frequency and 200 sccm O_2 gas. The oxygen plasma degrades organic material, such as resist.

This completed the lithography step, labelled *Step 1* in Figure 6.1, and resulted in the creation of SiN waveguides. An image of a chip after patterning (i.e., after etching and cleaning) is shown in Figure 6.7.

6.4 Metallization of Alumina and Gold

The completion of the waveguide platform is achieved by depositing alumina and gold in the active region. This is necessary to get the desired plasmon enhancement of the Raman scattering. This was not done in this project, but should be done in future projects. This would require two lithography steps and two metal deposition steps, labelled *Step 2* and *Step 3* in the process flow chart in Figure 6.1.

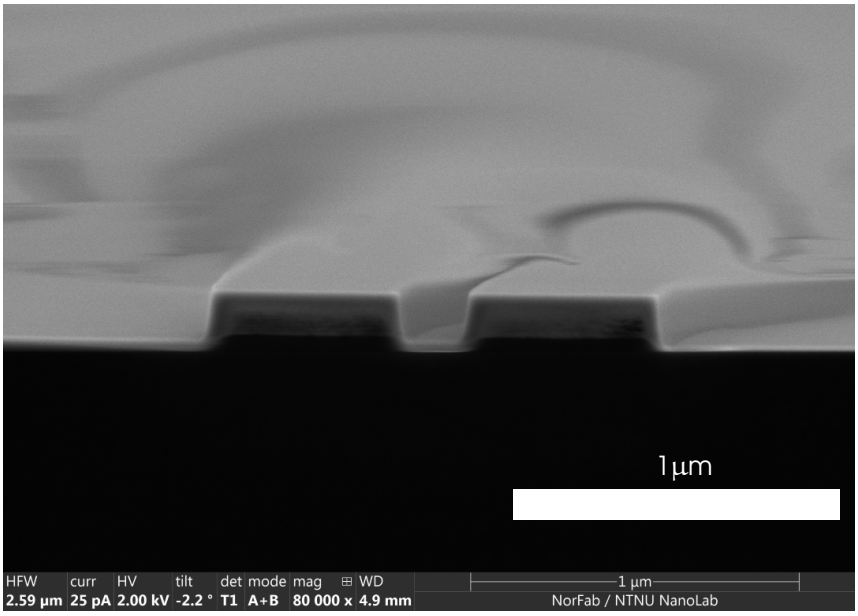


Figure 6.6: SEM image taken at NTNU NorFab facilities showing the cross-section of one of the waveguides fabricated in this thesis. The etch depth is ~ 145 nm, meaning that the entire SiN thin film not part of the waveguide was successfully etched away.

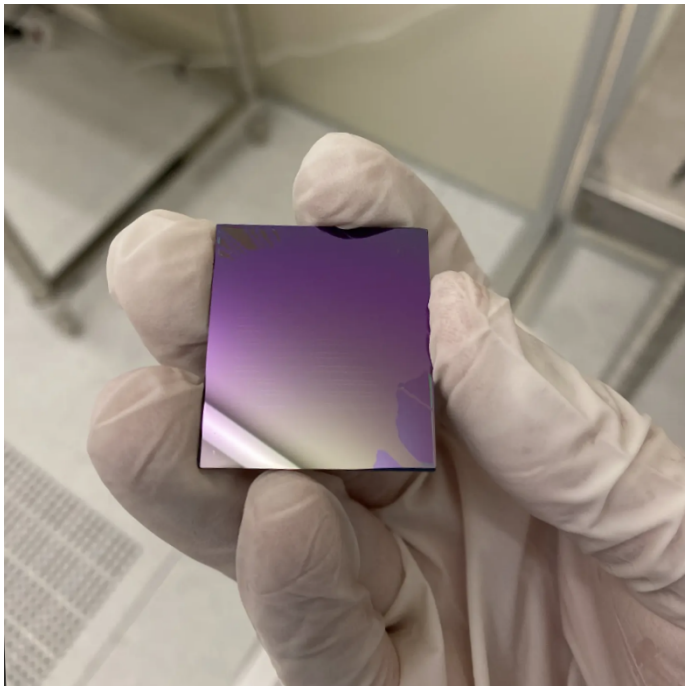


Figure 6.7: Image of a SiN waveguide chip post-patterning, i.e., after lithography, etching, and cleaning. The matte areas of the chip were not covered with resist (or only covered with a thin layer) during processing, and were etched. This does not affect performance, as these areas are free of any waveguide structures.

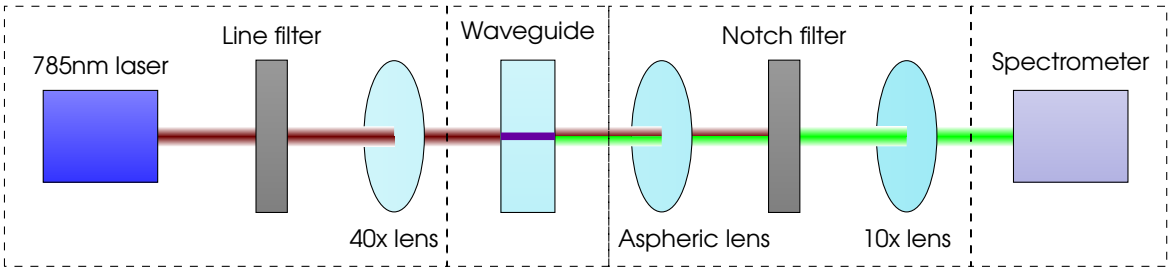


Figure 6.8: Sketch of the experimental Raman setup. Three lenses are used to prevent the beam from diverging. The laser line filter removes light not at 785 nm. The notch filter removes the pump laser, to produce a Raman signal. Details of each component is found in Table 6.2. The dashed boxes indicate the components placed on the same stage, as shown in Figure 6.9.

6.5 Raman Measurement Setup

The measurement setup is displayed in Figures 6.8 and 6.9. They show how the waveguides' Raman signal were measured in transmission. A 785 nm single mode fiber-coupled laser diode (LP785-SAV50, Thorlabs) was injected into the SiN waveguide sample through a laser line filter (FL780-10, Thorlabs) and a 40x focusing lens (UPLFNL 40X, Olympus). The transmitted beam containing both the remaining pump beam and the Raman signal was then focused by a aspheric lens (C220TMD-B, Thorlabs) into a notch filter (NF785-33, Thorlabs) to remove excess pump light. 10x lens (1-U2B223, Olympus). The Raman signal was then fiber coupled into the spectrometer (QE Pro Raman Spectrometer, Ocean Optics) through a 10x focusing lens (UPLNFL 10X, Olympus). A summary of the components in the setup is listed in Table 6.2.

The setup is self-built and aligned for the purposes of this thesis. It was loosely based on a previous setup in the same lab [37], although that setup was designed to work in reflection. The setup constructed here is designed to work in transmission. It is made as compact as possible to minimize beam divergence, and includes as few optical components as possible to minimize losses. Components following a lens have been positioned at the lens' focal length. The components were placed on four stages, allowing for accurate and independent control of all three spatial dimensions. The waveguide stage had to be independent to measure all the waveguides on the same chip. Additionally, the lateral positions of the input and output beams differ by 100 µm, due to the design described in Section 5. This made it necessary to place the components on three stages, and the fourth stage was included because the spectrometer input fiber did not fit on the third stage.

The spectrometer had a resolution of $\Delta\lambda = 0.2$ nm, and a range of 780—1013 nm. In units of wavenumber, this correspond to $\Delta k \approx 3.5$ cm^{-2} and a range of 0—2873 cm^{-2}, when the Stokes shift results from a 785 nm excitation laser. The spectrometer was operated at an integration time of either 10 seconds or 30 seconds.

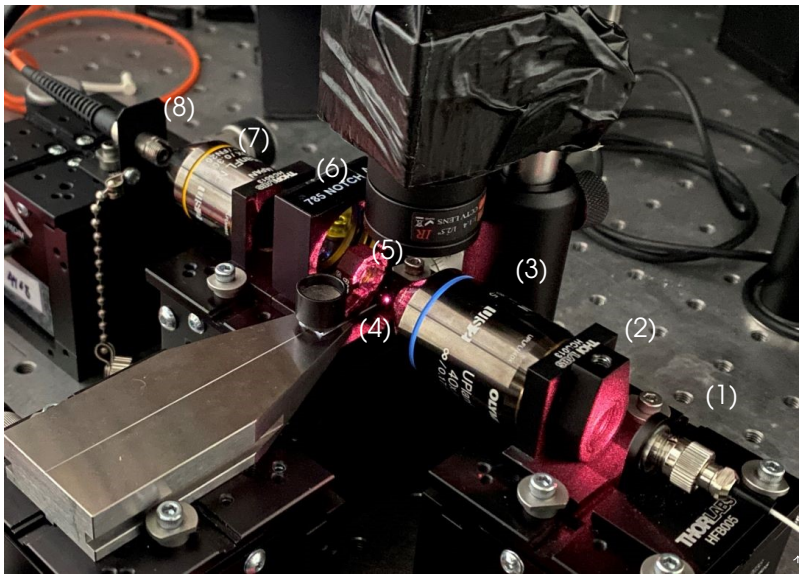


Figure 6.9: Image of the self-built experimental setup, under operation. The red/pink light is the 785 nm near-infrared pump light, and the beam can be seen injected into the chip. The component numbering is: (1) Fiber pump laser, (2) laser line filter, (3) 40x focusing lens, (4) waveguide chip on a metal sample holder, (5) aspheric lens, (6) notch filter, (7) 10x focusing lens, (8) fiber to spectrometer. An infrared camera is situated above the chip. The optical components following a lens, are positioned at the lens' focal points. Four stages with independent control of all three spatial directions were used, which allows for accurate alignment. The first stage holds the components (1)—(3), the second (4), the third (5)—(7), and the fourth (8).

Table 6.2: Components used in the measurement setup, including their most important specifications, such as the numerical aperture (NA), the focal length (f), the magnification (m), and the wavelength resolution ($\Delta\lambda$), where relevant.

Type	Specs	Model	Manufacturer
785 nm laser	Singlemode	LP785-SAV50	Thorlabs
Laser line filter		FL780-10	ThorLabs
40x lens	NA=0.75, f=4.5 mm, m=40x	1-U2B5272	Olympus
Aspheric lens	NA=0.25, f=11 mm	C220TMD-B	Thorlabs
Notch filter		NF785-33	ThorLabs
10x lens	NA=0.25, f=10.6 mm, m=10x	1-U2B223	Olympus
Raman spectrometer	$\Delta\lambda$ = 0.2 nm	QE Pro	Ocean Optics

Chapter 7

Experimental Results and Discussion

The final step in this thesis involved testing of the fabricated waveguide platform, by measuring the waveguides' Raman spectra. The results of these measurements are presented here. Unfortunately, all Raman spectra were nearly identical, displaying a few distinct peaks in an otherwise noisy signal. The measurements were expected to contain some Raman background signal from the SiN waveguides, but this was not observed. Possible explanations for this will be discussed in this section, but it is believed that large losses in the waveguide, as well as a low power laser, might have contributed to this. The fact that the experimental setup was designed and built for this thesis specifically, made it of critical importance to test the setup, and ensure alignment. The entire Section 7.2 is devoted to an explanation of the sample alignment. This was considered necessary to explain why there was no detectable Raman signal, and to suggest improvments to the setup, in case the work presented in this thesis should be continued.

Whenever *setup* is mentioned in this section, it refers to the self-built experimental setup described in Section 6.5, and shown in Figures 6.8 and 6.9. The Raman measurements were performed on this setup. The maximum output of the excitation laser (LP785-SAV50, Thorlabs) was measured to be about 13 mW with a power detector, before injection into the optical setup. In all measurements described in this section, the laser was operated at this maximum power. The spectrometer (QE Pro, Ocean Optics) detected wavelengths from 780 nm to 1014 nm, with a resolution of 0.22 nm. This corresponds to a Stokes shift of 0 to 2900 cm^{-2} at a 785 nm pump laser. The integration time was set to either 10 seconds, or 30 seconds.

7.1 The Measured Waveguide Platform

One of the fabricated chips was tested. 16 waveguides were patterned on the chip, according to the fabrication process described in Section 6.3. Eight different designs were produced, resulting in two samples for each design. The different waveguides are listed in Table 7.1. The waveguide design was previously described in Section 5, and the details of the different waveguide regions will not be repeated here. The reasons for testing different waveguide designs were twofold, namely (a) test the effects of the input, output, taper, and Euler bend regions (designs 1—3), and (b) to test the effects of the different dimensions in the active region (designs 4—8).

7.2 Alignment Tests

To ensure successful coupling of pump light into the waveguide chip, the alignment of the setup was tested. Since the setup was self-built for this thesis, alignment testing was deemed necessary, and relevant to, both for reproducibility and documentation purposes. Consider the setup in Figure 6.9. In this setup, the sample (4) must be correctly aligned in all three spatial directions (x, y and z) relative to both

Table 7.1: The waveguides tested in this thesis. The design details and the properties of each waveguide region are described in Chapter 5. The index *i,j* labels each waveguide according to design *i*, and sample number *j*. In the designs 1—3, the Euler bend is absent. In design 1, there is no active region, and the waveguide width is the same at the input and output. Design 2 has a narrower width at the output than design 3, since design 2 lacks an output taper. In the designs 4—8, the only difference is the rail width r in the active region. The gap g = 150 nm, and the input, output, taper, and Euler bend regions remain unchanged.

Index		Regions Present	Purpose
1.1	1.2	Only input	Test coupling without tapering or bend
2.1	2.2	Input taper	Test effect of input taper
3.1	3.2	Input & output tapers	Test effect of output taper
4.1	4.2	All. Rails r = 500 nm	Test dimensions in active region
5.1	5.2	All. Rails r = 550 nm	Test dimensions in active region
6.1	6.2	All. Rails r = 600 nm	Test dimensions in active region
7.1	7.2	All. Rails r = 650 nm	Test dimensions in active region
8.1	8.2	All. Rails r = 700 nm	Test dimensions in active region

the objective lens (3) and the aspheric lens (5). The numbers in parenthesis refer to the labelling in Figure 6.9, and the direction x is parallel to the pump beam, y is transversal to the beam, and z is the height. In total, six independent positions must be aligned correctly.

The alignment to the aspheric lens (5) was simplified through the following assumptions. Upon exiting the chip, the transmitted beam was assumed to have a very large divergence. Therefore, the sample was placed as close to the lens as possible, solving the alignment in x. Next, alignment in y and z was solved by centering the output beam of the objective lens (3) at the center of the aspheric lens. Thus, the alignment to the aspheric lens was assumed to be achieved whenever the sample was aligned to the objective lens. The remaining challenge was to ensure the alignment of the sample and the objective lens. Ideal alignment of the sample and objective is assumed to be achieved in each spatial direction when the following conditions wareere met:

(i) In x when the input edge of the sample is positioned at the lens' focal point, i.e., at the smallest beam width,

(ii) in y when the width of the incident beam enters the waveguide of interest,

(iii) in z when the height of the incident beam was centered at the waveguide structure itself, i.e, just below the sample surface.

The two following methods were employed to assess the quality of the alignment:

(a) Visual inspection of images, meaning that the laser beam's shape and the illumination of the sample was observed, and

(b) observation of fluctuations in the experimental data, as the sample was measured at different alignment positions.

The following will describe how each method was employed, and their challenges. The visual method (a) was limited by the camera's resolution. In the images shown

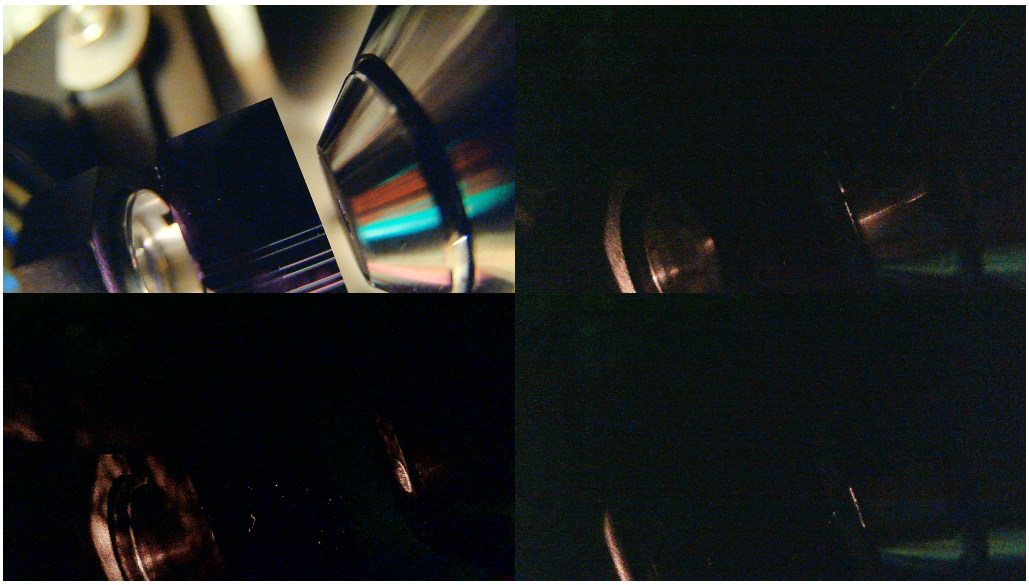


Figure 7.1: Images showing how the sample stage was aligned in z. The camera captured near infrared light. Top left: Image in the light, for reference. Top right: The stage is too high, the laser beam does not hit the sample. Bottom left: The stage is a little too low, as the laser beam illuminates the surface of the sample. The bright spots are dust particles on the sample surface scattering the laser light. Bottom right: The sample stage is too low. No coupling is observed, and (almost) no light is reflected off the aspheric lens, as it lies in the shadow of the sample.

in Figure 7.2, the input and output beams, and the waveguides are distinguishable, but their widths do not occupy more than a few pixels in the image, owing to the low resolution camera. This made it difficult to determine the exact position of the beam, relative to the waveguide. One major improvement would have been to include an optical microscope in the setup. However, it was not possible to obtain and integrate an optical microscope into this setup within the thesis' time constraints.

Condition (i) above, was solved by adjusting the x position of the objective lens, such that the beam was as narrow as possible at the sample edge. Figure 7.1 shows how the height of the sample was aligned, solving the z condition (iii) above. Successful z alignment was determined from observing the illumination of the sample surface, the reflection off the sample edge, and the amount of light transmitted through to the aspheric lens. In general, if the dust particles at the surface are illuminated, the stage was too low. If the sample was too high, the laser beam was blocked. This was observed as reflection on the input side of the waveguide, and that the aspheric lens was covered in shadow.

Results of the visual alignment in y are shown in Figure 7.2. The input beam appears to be successfully coupled into waveguide 2.1 (Figure 7.2b), and into the SiN thin film at both sides (Figures 7.2a and 7.2c). This is observed from the significant changes in the position and intensity of the output beam, resulting from only small shifts of the stage in y. Additionally, if the laser is coupled into the SiN thin film, it would be wide at the output, which is the case here. The output beams can be compared to Figure 7.4b, which shows a narrower output beam.

The other method to confirm alignment was the use of experimental measure-

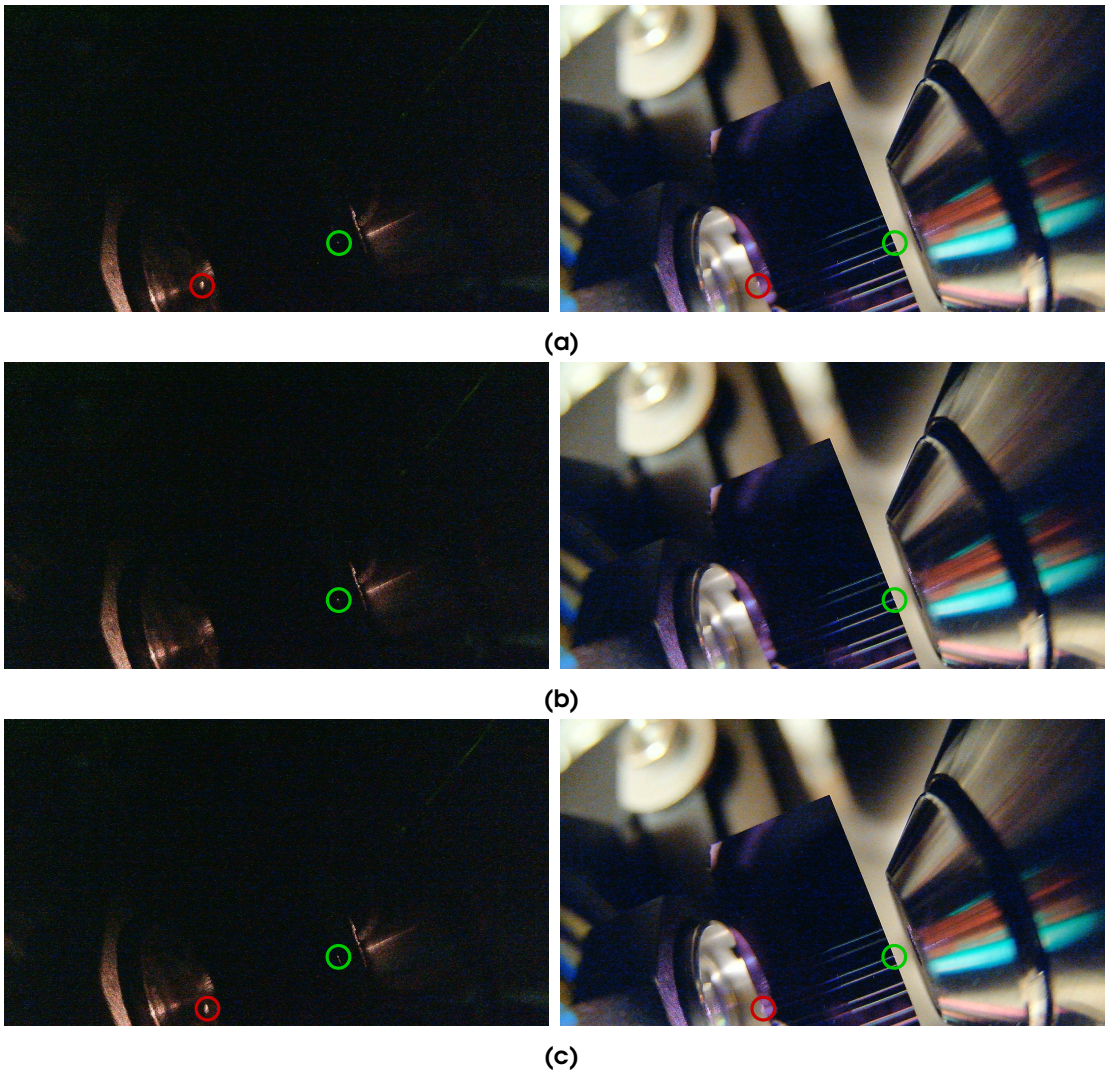


Figure 7.2: Images of the alignment testing, showing attempts to couple light into the waveguide labelled 2.1 in Table 7.1. The camera captured near infrared light. Each position is shown both in the dark (only laser light) and with ceiling lights. Between each image, the sample stage is shifted a small distance (on the order of microns) to the right relative to the beam. This resulted in the output beam appearing at different positions. The points where the beam is incident on the chip are marked with green circles, and the transmitted beam is marked with red circles. (a) Light is likely coupled into the SiN thin film to the right of the waveguide. (b) Light is possibly coupled into the waveguide, since the output beam is not detected in the image, indicating that the wave was attenuated in the waveguide. (c) Light is likely coupled into the SiN thin film to the left of the waveguide. The output beams of (a) and (c) are shifted a distance relative to each much larger than the respective input beams.

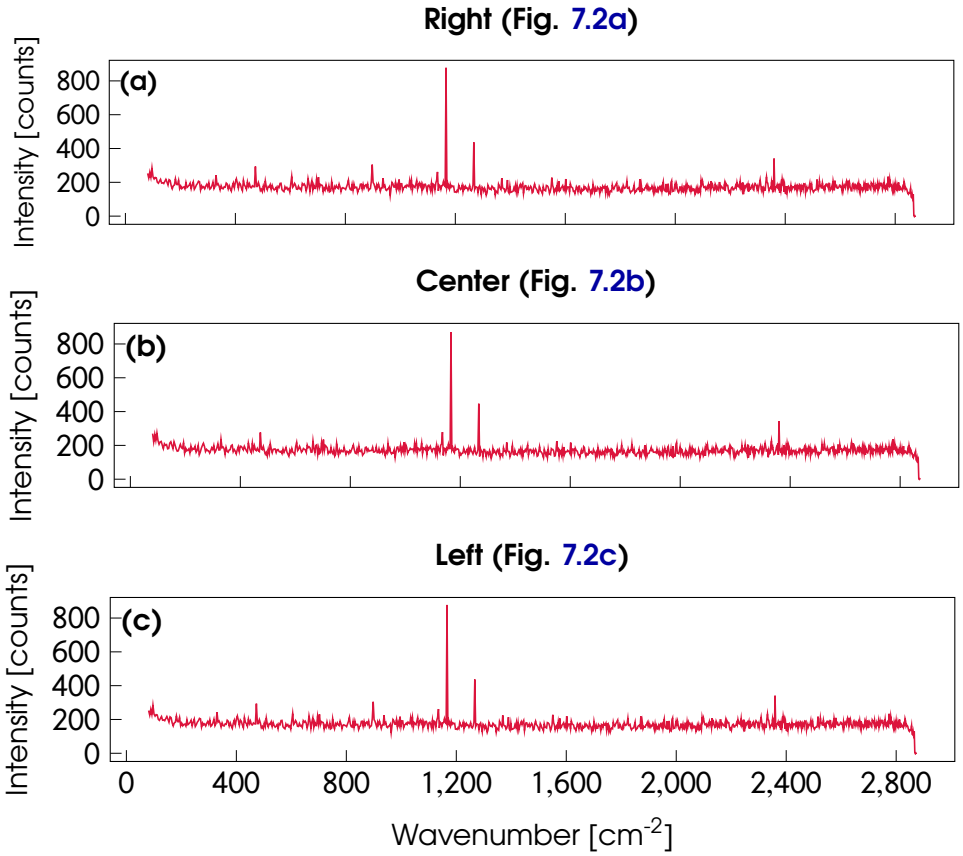


Figure 7.3: Nearly identical Raman measurements from the three different sample positions shown in Figure 7.2. There is no trace of the expected Raman background signal from the SiN waveguide. The integration time was 10 seconds. The two most prominent peaks are from the Stokes shifts at 1166 cm^{-2} and 1267 cm^{-2}.

ments. The Raman signal was assume to change depending on wether or not coupling into the sample was achieved, as coupling should result in a Raman background contribution from silicon nitride. In [13] and [11], for instance, this background is shown as a very broad intensity peak starting below $k \sim 1000$ cm^{-2}. When the beam was coupled into the sample, the signal collected at the detector should be reminiscent of such a background spectrum. In these Raman measurements, that was not the case. This is shown in Figure 7.3, where the Raman spectra are taken at the alignment positions shown in Figure 7.2. Instead of broad background peaks, the intensity signal remains nearly flat at all Stokes shifts. Although the input beam was coupled into different positions at the sample, the Raman signals remained invariant. Attempts of achieving good alignment are shown in Figure 7.4, including the collected Raman spectra. Even thought the laser beam appears to both enter and exit the sample, the Raman signals show no sign of any SiN background contribution. As before, the signal is noisy and flat.

Of course, one major source of error is whether or not the input beam was acually coupled into the waveguides. In this setup, the exact position of the input beam was not possible to determine. As mentioned above, the inclusion of an optical microscope, might have solved this issue. On the other hand, if coupling was indeed achieved, the fact that the background spectrum is missing might be explained by two possible causes. Either that (i) the transmitted beam did not reach

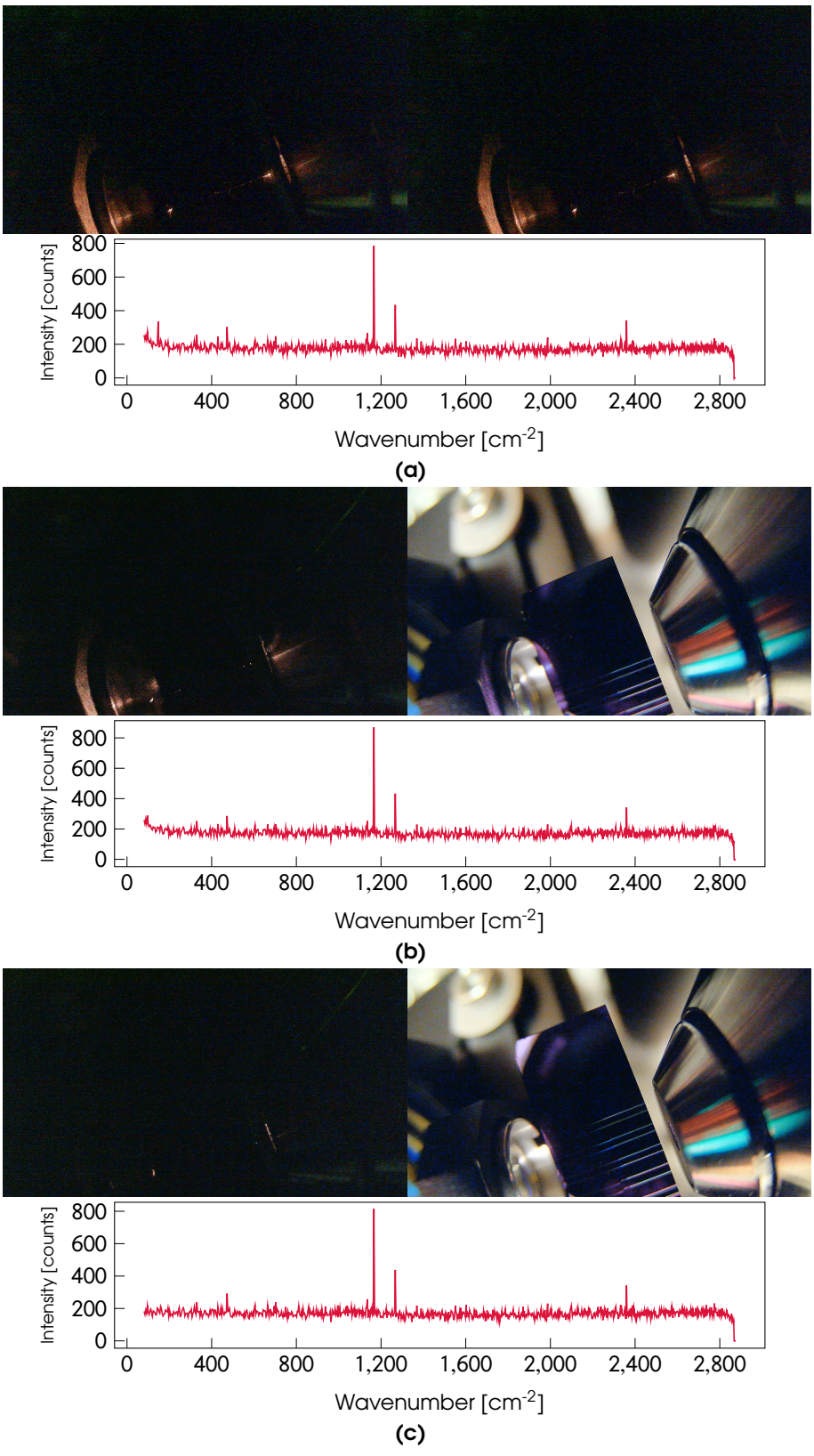


Figure 7.4: Raman measurements when coupling was possibly achieved, as both the input and output beams were visible. Nonetheless, the spectra remained similar as before. The waveguides tested were (a) 1.1, (b) 1.2, and (c) 3.1 (see Table 7.1 for the indices). The integration time was 10 seconds.

the detector, or that (ii) the transmitted beam was indistinguishable from noise, due to a low signal-to-noise ratio (SNR). (i) might be a result of beam attenuation through the optical components leading to the detector. If the transmitted beam power was small, it might be completely attenuated on its way to the detector. One solution of (i) would be to increase the laser power, to make sure the transmitted beam is strong enough to not be attenuated. The laser used in these experiments however, operated at its maximum power. Increasing the laser power would mean the acquisition of a higher power laser. Additionally, increasing the power would not improve the signal if (ii) is correct and the issue is a low SNR. Increasing the SNR would involve reducing losses in the waveguides, or increasing the Raman enhancement. Loss considerations were made in Section 5, and one solution may be a more compact waveguide. Increasing the Raman enhancement was considered in Section 4.3, and involves narrowing the waveguide gap and depositing a plasmonic layer on the waveguide. This was not done, due to the time restrictions of this thesis. In the expansion of this work, increasing the Raman conversion by including the plasmon layer would be a natural continuation.

7.3 Raman Measurements with IPA

Although the SiN Raman background spectrum was not identified, measurements were performed with isopropyl alcohol (IPA) deposited on the waveguides. The purpose was to collect the Raman spectra from different waveguides, and compare the intensity peaks of the wavenumbers associated with IPA. The intensities would be a measure of the performance of the different waveguide structures. Prior to IPA deposition, the background spectrum of three different waveguides (1.1, 4.2 and 5.2) were measured, and the mean of these spectra were used to compare with the IPA spectra. The resulting Raman spectra from waveguides 4.2 and 5.2 with IPA, are shown in Figure 7.5. These plots turned out very similar to the other plots shown in this section. To highlight any IPA peaks, the Raman spectra are shown in percentage deviation from the mean. The IPA peaks are located at 819, 953, 1134, and 1450 cm^{-2} [13], none of which are visible in Figure 7.5. The reasons for this are likely the same as for the measurements without IPA, which was discussed previously in Section 7.2.

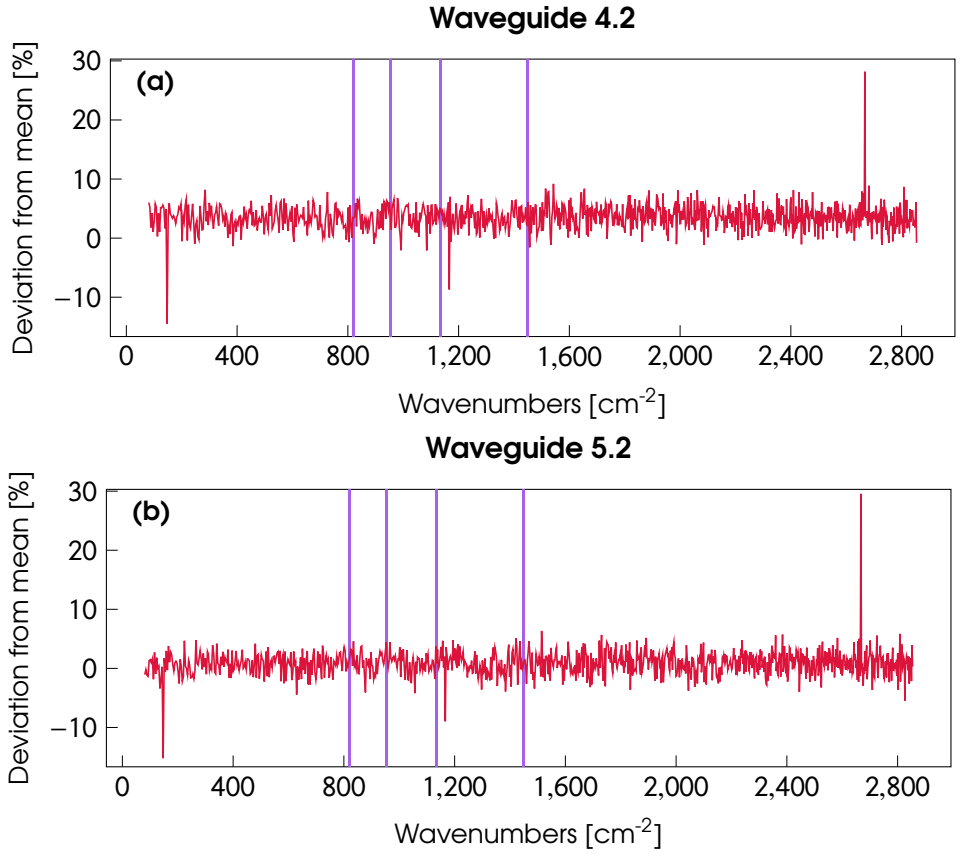


Figure 7.5: Raman measurements with IPA, displayed as percentage deviation from the mean. The mean, or the background spectrum, was calculated by taking the average of three Raman measurements without IPA. The indices of the measured waveguides were (a) 4.2, and (b) 5.2 (see Table 7.1 for the indices). Neither spectrum show a significant deviation from the background spectrum. The IPA peaks are included as verticle, purple lines.

Chapter 8

Conclusion

This thesis focused on developing a slot waveguide sensor for use in surface-enhanced Raman scattering (SERS). One major advantage of a waveguide-based SERS platform, is the prospects of placing an entire integrated photonic circuit on a chip. This would allow for cheap mass-scale production, and the ability to create sensory chips tailored to highly specific sensing tasks. Based on the theory presented in Chapters 2 and 3, the slot waveguide was selected for future investigations. This was in part due to the high performance of nanoplasmonic slot waveguides recently reported by Turk *et al.* [4], with a sensing performance approaching that of free-space systems. Additionally, work on nanoplasmonic waveguides reported by the *Photonics Research Group* at Ghent University, served as an inspiration for this thesis. In particular, the EBL-free nanoplasmonic slot waveguide platform proposed by A. Raza *et al.* [19].

Continuing with the slot waveguide as the sensing platform of choice, numerical simulations were carried out in Chapter 2 with the goal of determining the optimal dimensions for SERS sensing. A finite-difference eigenmode (FDE) simulation environment was designed in Ansys' *Lumerical* software. In the simulations, several sensing performance benchmarks were accounted for. These involved how the waveguide's dimensions influenced its single mode operation, conversion efficiency η, and absorption loss α_a. A suitable SiN slot waveguide in regards to these benchmarks was identified. This slot waveguide was 140 nm high, with a 150 nm gap and 500 nm wide rails, and an alumina spacer layer and a plasmonic gold layer with a combined thickness of 68—71 nm. This waveguide would operate at a single mode, with a maximal calculated conversion efficiency η on the order of 10^4. The Raman signal to pump ratio is proportional to the conversion efficiency, i.e., $P_s/P_p \propto \eta$, meaning that η is directly related to the enhancement of the Raman signal.

In Chapter 5 this optimized cross-sectional waveguide design was subsequently extended upon, in order to place the design on a chip. This involved the addition of taper regions and an Euler bend, in order to successfully couple an incident laser beam, guide the beam to the active sensing region, and finally couple it back out to free-space for detection at a spectrometer. The widths and shapes of these taper regions were designed to have minimal losses, based on previous loss considerations of the waveguide platform in Chapter 4.

Next, in Chapter 6, the waveguide design was realized by fabrication in the NorFab cleanroom facilities at NTNU. This processing step involved waveguide patterning on SiN wafers supplied by SINTEF, by means of electron beam lithography (EBL) and inductively coupled plasma-reactive ion etch (ICP-RIE). The characterization these structures was performed using various structures, most notably a scanning electron microscope (SEM). The process parameters of the EBL were tuned to suit this project, and the etch rate of the SiN sample during ICP-RIE was identified. The characterization showed that the fabrication of slot SiN waveguides with the desired dimensions was successfully achieved. Due to the time restrictions of

this project, the alumina spacer layer and the plasmonic gold layer considered in simulations, were not deposited on the waveguides. This step is left as a natural continuation of the work presented here.

Finally, one of the fabricated SiN slot waveguide chips was tested in a Raman measurement setup. The setup was self-built for the purposes of this thesis, and involved identifying the necessary optical components, and correct alignment to ensure coupling of the laser beam into the spectrometer. The Raman spectra of different waveguides were collected, both with and without isopropyl alcohol (IPA). This resulted in the collection of the SiN waveguide's background spectrum, and the IPA Raman spectrum in the SiN background spectrum, respectively. The Raman spectrum of IPA is distinguished from four characteristic peaks [13]. However, the resulting measurements showed no notable distinction between the two types of spectrum, and the measured spectra were thought to mostly consist of noise.

Several factors were mentioned in an attempt to explain the absence of the expected measurement results. One possibility was a high signal-to-noise ratio (SNR) in the waveguides, which may attenuate the incident laser beam completely, including any Stokes shifted light. Another possibility was misalignment of the sample or optical components in the setup. The last possibility which was discussed was a low excitation laser power resulting in attenuation of any Stokes shifted light prior to reaching the detector. Possible solutions to these three issues were proposed. Either, by including the plasmonic layer in the waveguide structure to enhance the Raman signal, expanding the setup with an optical microscope, or acquiring a higher power laser.

The lack of experimental results that characterize the performance of the fabricated SiN slot waveguides, does not render the work in this thesis futile. On the contrary, useful contributions were made in both theory and practice. The theoretical performance of SiN slot waveguides were calculated, and a fabrication process to realize such waveguides in the NTNU NorFab facilities was established. These contributions may be taken advantage of, should this project be repeated or expanded upon. A reasonable continuation of this thesis, would be to make the slot waveguides plasmonic, goal which has been repeated throughout this thesis. Its inclusion might result in the creation of a high-performing Raman sensor, which with maturing technology may be placed on a single chip.

References

[2] M. Moskovits. "Surface roughness and the enhanced intensity of Raman scattering by molecules adsorbed on metals". In: *The Journal of Chemical Physics* 69.9 (1978). Cross-Link, pp. 4159–4161.

[3] Katrin Kneipp et al. "Single Molecule Detection Using Surface-Enhanced Raman Scattering (SERS)". In: *Phys. Rev. Lett.* 78 (9 Mar. 1997), pp. 1667–1670. DOI: 10.1103/PhysRevLett.78.1667.

[4] Nina Turk et al. "Comparison of Free-Space and Waveguide-Based SERS Platforms". In: *Nanomaterials* 9 (2019). DOI: 10.3390/nano9101401.

[5] Paul L. Stiles et al. "Surface-Enhanced Raman Spectroscopy". In: *Annual Review of Analytical Chemistry* 1 (2008), pp. 601–626. DOI: 10.1146/annurev.anchem.1.031207.112814.

[6] Bahaa E.A. Saleh and Malvin Carl Teich. *Fundamentals of Photonics*. 3rd ed. 111 River Streeet, Hoboken, NJ 07030, USA: John Wiley & Sons, Inc., 2019. ISBN: 9781119506874.

[7] C. Yeh and F. I. Shimabukuro. *The Essence of Dielectric Waveguides*. 223 Spring Street, New York, NY 10013, USA: Springer, 2008. ISBN: 978-0-387-30929-3.

[8] Pieter Wuytens. "Surface-Enhanced Raman Spectroscopy for Intracellular Sensing and Protease Activity Detection: From Chip Technology to Applications". PhD thesis. Ghent University, 2017.

[9] Pierre Berini. "Figures of merit for surface plasmon waveguides". In: *Opt. Express* 14.26 (Dec. 2006), pp. 13030–13042. DOI: 10.1364/OE.14.013030.

[10] William L Barnes. "Surface plasmon–polariton length scales: a route to sub-wavelength optics". In: *Journal of Optics A: Pure and Applied Optics* 8.4 (Mar. 2006), S87. DOI: 10.1088/1464-4258/8/4/S06. URL: https://dx.doi.org/10.1088/1464-4258/8/4/S06.

[11] Ananth Z. Subramanian et al. "Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip". In: *Photon. Res.* 3.5 (Oct. 2015), B47–B59. DOI: 10.1364/PRJ.3.000B47.

[12] Ashim Dhakal et al. "Efficiency of evanescent excitation and collection of spontaneous Raman scattering near high index contrast channel waveguides". In: *Opt. Express* 23.21 (Oct. 2015), pp. 27391–27404. DOI: 10.1364/OE.23.027391. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-23-21-27391.

[13] Ashim Dhakal et al. "Single mode waveguide platform for spontaneous and surface-enhanced on-chip Raman spectroscopy". In: *Interface Focus* 6 (2016). DOI: 10.1098/rsfs.2016.0015.

[14] David A. Coucheron et al. "Chip-Based Resonance Raman Spectroscopy Using Tantalum Pentoxide Waveguides". In: *IEEE Photonics Technology Letters* 31.14 (2019), pp. 1127–1130. DOI: 10.1109/LPT.2019.2915671.

[15] Derek M. Kita et al. "Are slot and sub-wavelength grating waveguides better than strip waveguides for sensing?" In: *Optica* 5.9 (Sept. 2018), pp. 1046–1054. DOI: 10.1364/OPTICA.5.001046.

[16] Frédéric Peyskens et al. "Waveguide excitation and collection of surface-enhanced Raman scattering from a single plasmonic antenna". In: *Nanophotonics* 7 (May 2018). DOI: 10.1515/nanoph-2018-0003.

[17] Frédéric Peyskens et al. "Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform". In: *ACS Photonics* 3 (2016), pp. 102–108. DOI: 10.1021/acsphotonics.5b00487.

[18] Pieter C. Wuytens, Andre G. Skirtach, and Roel Baets. "On-chip surface-enhanced Raman spectroscopy using nanosphere-lithography patterned antennas on silicon nitride waveguides". In: *Optics Express* 25.11 (2017), pp. 12926–12934. DOI: 10.1021/acsphotonics.5b00487.

[19] Ali Raza et al. "ALD assisted nanoplasmonic slot waveguide for on-chip enhanced Raman spectroscopy". In: *APL Photonics* 3.116105 (2018). DOI: https://doi.org/10.1063/1.5048266.

[20] A. Raza et al. "E-beam-lithography free plasmonic slot waveguides for on-chip Raman spectroscopy". In: (2018), pp. 1–2. URL: https://ieeexplore.ieee.org/document/8428070.

[21] Pieter C. Wuytens et al. "Gold nanodome SERS platform for label-free detection of protease activity". In: *Faraday Discuss.* 205 (0 2017), pp. 345–361. DOI: 10.1039/C7FD00124J.

[22] Shiyu Li et al. "Surface-enhanced Raman scattering sensor based on hybrid deep slot waveguide on an integrated photonic platform". In: *J. Opt. Soc. Am. B* 36.9 (Sept. 2019), pp. 2423–2428. DOI: 10.1364/JOSAB.36.002423.

[23] Qipu Cao et al. "Surface-enhanced Raman scattering using nanoporous gold on suspended silicon nitride waveguides". In: *Opt. Express* 26.19 (Sept. 2018), pp. 24614–24620. DOI: 10.1364/OE.26.024614.

[24] Pan Du et al. "In situ surface-enhanced Raman scattering monitoring of reduction of 4-nitrothiophenol on bifunctional metallic nanostructure". In: *Japanese Journal of Applied Physics* 57.3 (Feb. 2018), p. 030308. DOI: 10.7567/JJAP.57.030308.

[25] Kane Yee. "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media". In: *IEEE Transactions on Antennas and Propagation* 14.3 (1966), pp. 302–307. DOI: 10.1109/TAP.1966.1138693.

[26] Zhaoming Zhu and Thomas G. Brown. "Full-vectorial finite-difference analysis of microstructured optical fibers". In: *Opt. Express* 10.17 (Aug. 2002), pp. 853–864. DOI: 10.1364/OE.10.000853. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-10-17-853.

[27] Dominic Gallagher. "Industry Research Highlights, Photonic CAD Meatures". In: *LEOS Newsletter* 22.1 (2008), pp. 2–3.

[29] Herbert R. Philipp. "Optical Properties of Silicon Nitride". In: *Journal of The Electrochemical Society* 120.2 (Feb. 1973), p. 295. DOI: 10.1149/1.2403440. URL: https://dx.doi.org/10.1149/1.2403440.

[30] H.R. Philipp. "Silicon Dioxide (SiO2) (Glass)". In: (1997). Ed. by Edward D. Palik, pp. 749–763. DOI: https://doi.org/10.1016/B978-012544415-6.50038-8.

[31] P. B. Johnson and R. W. Christy. "Optical Constants of the Noble Metals". In: *Phys. Rev. B* 6 (12 Dec. 1972), pp. 4370–4379. DOI: 10.1103/PhysRevB.6.4370.

[33] A. Gorin et al. "Fabrication of silicon nitride waveguides for visible-light using PECVD: a study of the effect of plasma frequency on optical properties". In: *Opt. Express* 16.18 (Sept. 2008), pp. 13509–13516. DOI: 10.1364/OE.16.013509.

[34] I. Guler. "Optical and structural characterization of silicon nitride thin films deposited by PECVD". In: *Materials Science and Engineering: B* 246 (2019), pp. 21–26. ISSN: 0921-5107. DOI: https://doi.org/10.1016/j.mseb.2019.05.024.

[35] N. Daldosso et al. "Fabrication and optical characterization of thin two-dimensional Si3N4 waveguides". In: *Materials Science in Semiconductor Processing* 7.4 (2004). Papers presented at the E-MRS 2004 Spring Meeting Symposium C: New Materials in Future Silicon Technology, pp. 453–458. ISSN: 1369-8001. DOI: https://doi.org/10.1016/j.mssp.2004.09.023. URL: https://www.sciencedirect.com/science/article/pii/S1369800104000757.

[37] Karolina Milenko et al. "Optimization of SERS Sensing With Micro-Lensed Optical Fibers and Au Nano-Film". In: *Journal of Lightwave Technology* 38.7 (2020), pp. 2081–2085. DOI: 10.1109/JLT.2019.2958128.

LINKS

[1] *Choosing Colormaps in Matplotlib*. Accessed June 2023. URL: https://matplotlib.org/stable/tutorials/colors/colormaps.html.

[28] Lumerical's Webpage. URL: https://www.lumerical.com/.

[32] Joakim Ekern's GitHub profile. URL: https://github.com/joakimekern.

[36] Allresist. *Positive E-Beam Resists AR-P 6200* (*CSAR 62*). Accessed June 2023. Cross-Link.

Appendix

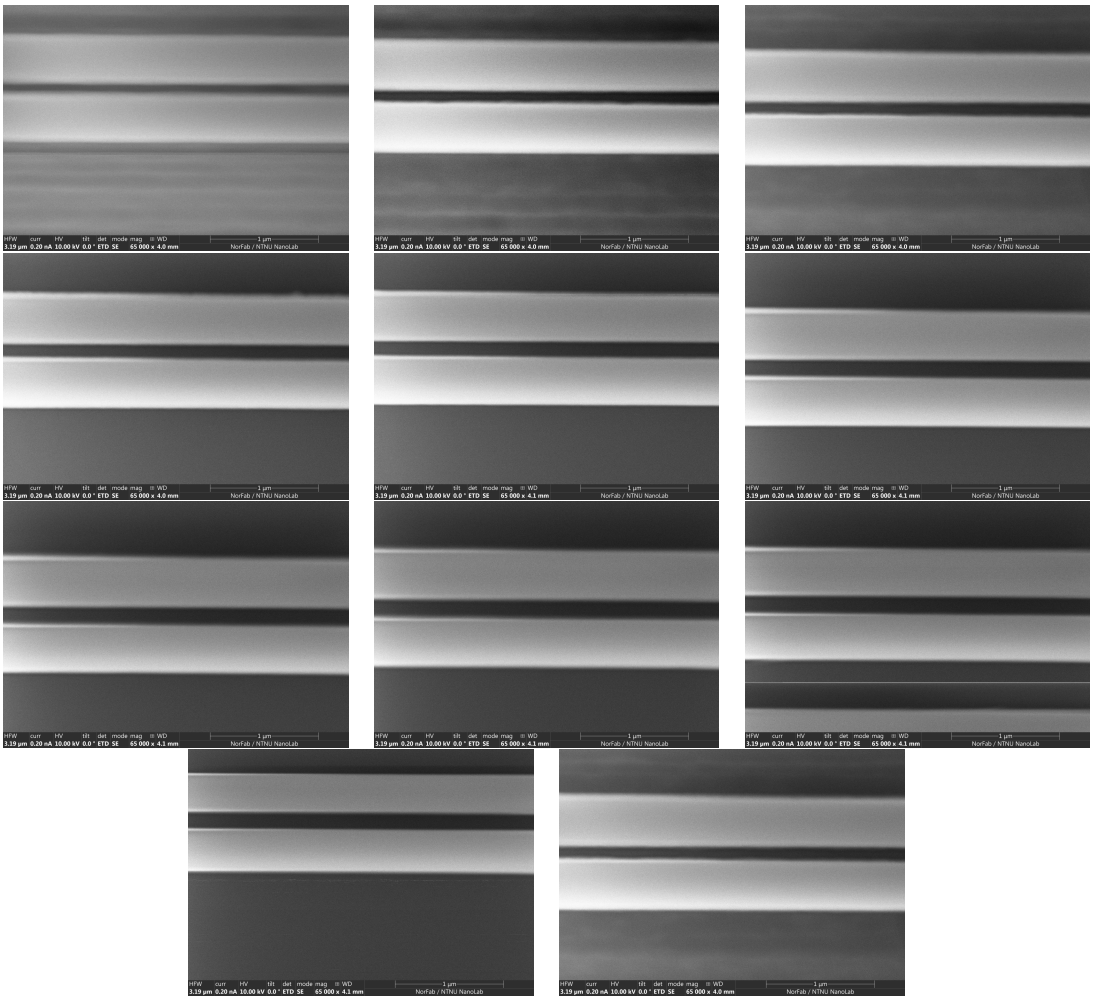


Figure A.1: Results of the EBL dosetest. The area doses were varied from 100 to 200 $\mu C/cm^2$. The doses are incremented by 10 from left to right, and top to bottom. Hence, the top left dose is 100 $\mu C/cm^2$, the top right dose is 120 $\mu C/cm^2$, etc.

Appendix A

Fabrication Details

This part of the appendix documents the testing of the process parameters which resulted in the fabrication method presented in Section 6.

Dosetesting

A dose test was performed, the results of which are shown in Figure 6.4. The SEM images resulting in the dosetest are shown in Figure A.1.

Cross-Sections

The cross-sections of the six different waveguide designs were imaged with SEM. These waveguides are identical to the waveguides fabricated in Chapter 6, and their dimensions are listed in Table 7.1. Here, the waveguide designs 2 and 3 are not imaged, since their cross-sections are identical to that of design 1, except for in

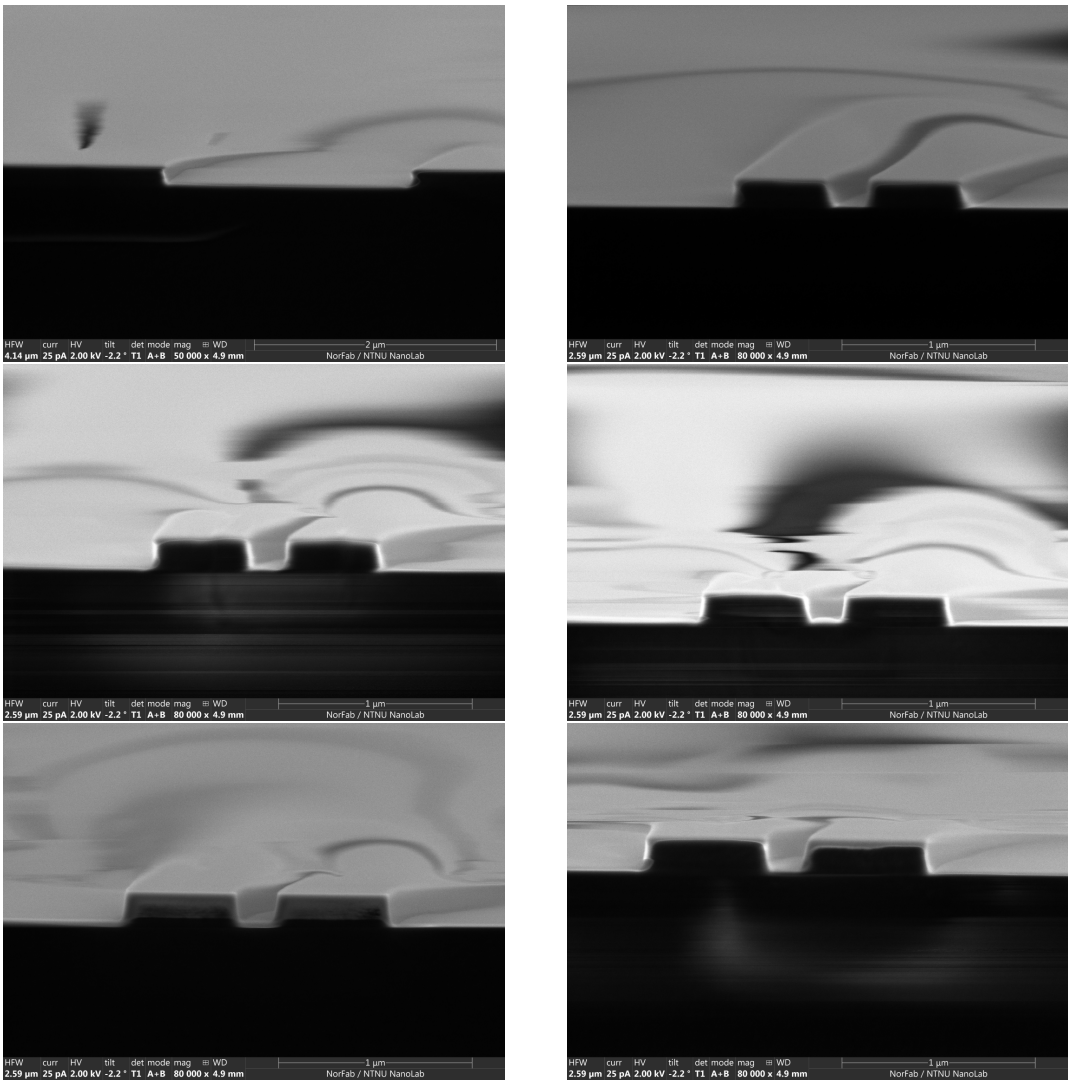


Figure A.2: Cross-sections of the fabricated designs. The rail widths r are given in Table 7.1, and correspond to the rail width of the etch mask described in Chapter 5, which not necessarily is their widths as depicted here (the fabricated with is slightly narrower, about 100 nm). From top to bottom, and left to right, the designs were: 1, 4, 5, 6, 7 and 8.

their taper regions.

Mask Test

Different rail widths of the mask was tested. This test is the same as the one showed in Figure A.2. Here, top view images are shown instead of the cross-section.

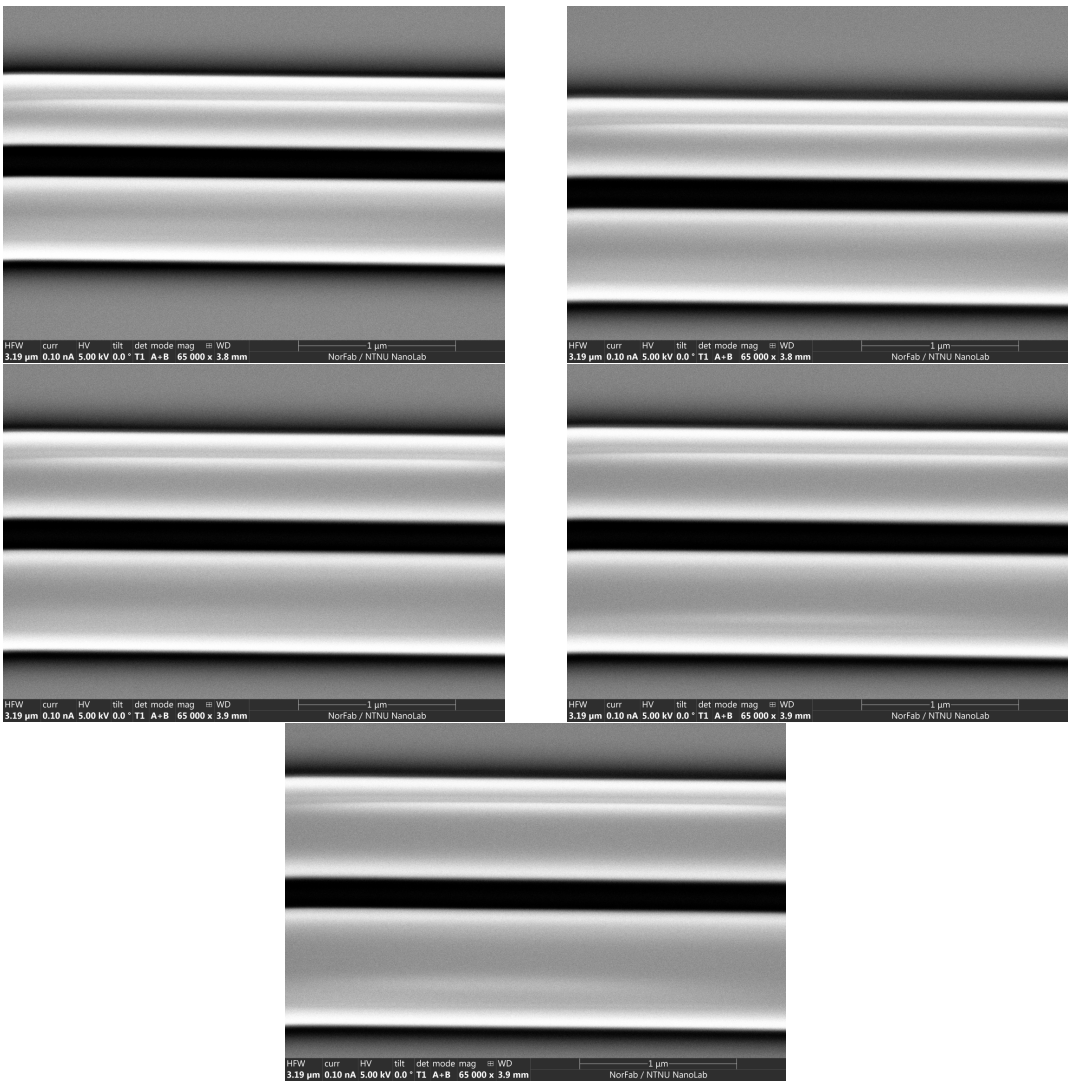


Figure A.3: Top view of the same structured as imaged in Figure A.2 of the fabricated waveguide structures listed in Table 7.1. In this figure however, only designs 4-8 are shown, where the design is varied in the active region. The rail width r increases from left to right, and top to bottom in the figure. The top left rail width mask was r = 500 nm.

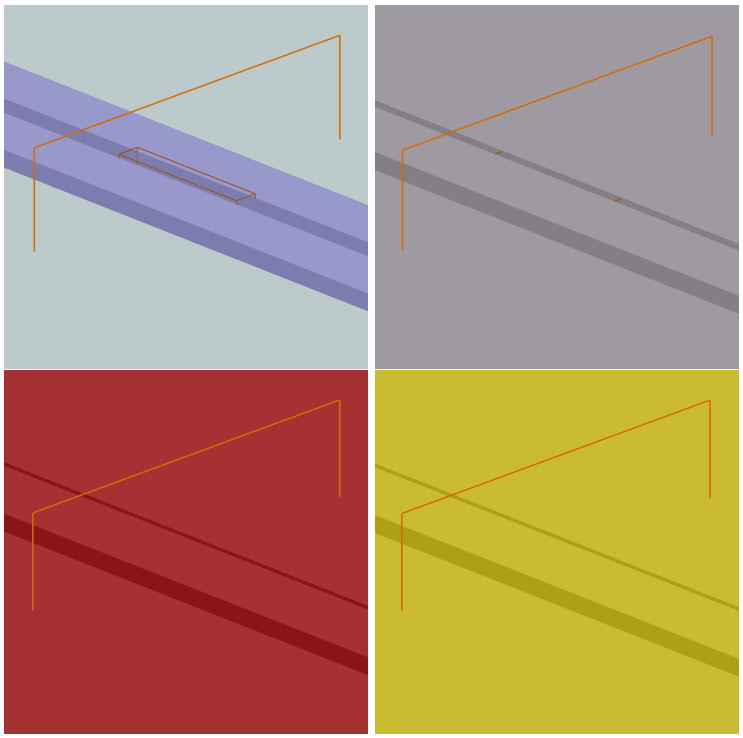


Figure B.4: Screenshots of the *Perspective view* of the simulation setup with various layers. Clockwise from top left: plain SiN slot waveguide without additional layers; with Al_2O_3 layer; with Au layer; with NTP layer. All image rights belong to Ansys.

Appendix B

Details of the Simulation Setup in Lumerical

An overview of the simulation setup employed in this thesis using the Lumerical software from Ansys is found in Section 4.2. Here follows a complete description of the setup, including any code necessary to replicate the simulations carried out in this thesis. This implementation may serve as a guide for creating similar simulations in Lumerical. Note that this may not be the best way to implement the simulations, and any users should feel free to make any changes and improvements when necessary. The explanation assumes that the user is already familiar with the objects and the nomenclature in Lumerical.

The Lumerical version used in this thesis was 2023 R1.

Setup

A setup script was utilized to simplify the setup and modification of different Lumerical simulations. The two screenshots in Figures B.5 and B.6 from Lumerical display what variables and objects were defined in order for the setup script to work. It is by no means necessary to replicate this setup method in order to create simulations in Lumerical, but this was the technique I found to be the most useful. Note that the "mesh order" property of the materials in the database was changed. This value governs which material should be included in the mesh, if two materials overlap. It is 2 for all materials by default, and favors the materials with the lowest number in

#	Name	Type	Value	Unit
1	wg_width	Length	400	nm
2	wg_height	Length	140	nm
3	wg_gap	Length	150	nm
4	wg_length	Length	100000	nm
5	t_siox	Length	2400	nm
6	t_alox	Length	40	nm
7	t_au	Length	15	nm
8	t_ntp	Length	1	nm
9	y_meshsize	Length	10	nm
10	z_meshsize	Length	10	nm
11	y_span	Length	2500	nm
12	z_span	Length	1500	nm

Figure B.5: The setup variables defined within the *::model* scope. These variables are employed by the setup script.

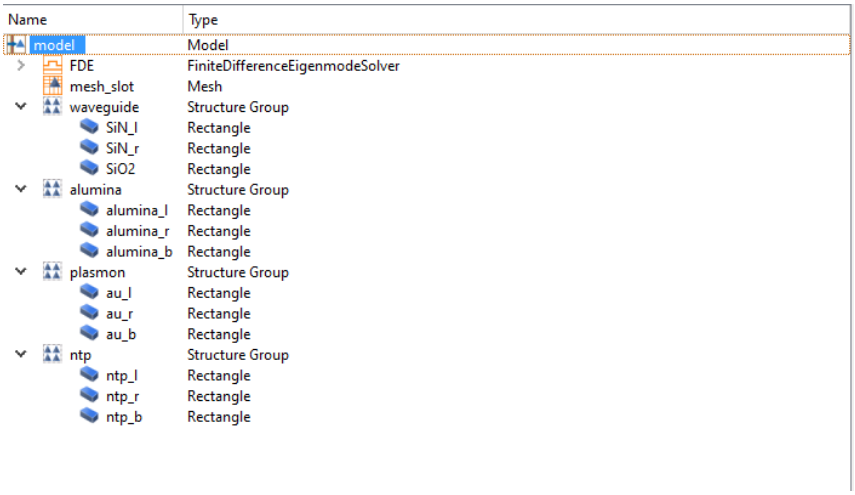


Figure B.6: The objects which was defined to create the SiN slot waveguides.

case of overlap. To get the desired structure the mesh order was changed, such that Si_3N_4 was preferred over Al_2O_3, over Au, over NTP. This is necessary for the results to work as expected.

Setup Script

The following script was included within the *model* analysis group in Figure B.6, in the *Script* tab. It governs the adjustment of each simulation object relative to each other, as the parameters shown in Figure B.5 are adjusted.

```
y_pos = (wg_width + wg_gap)/2;
z_pos = -wg_height/2;
gap_width = wg_gap - 2*(t_alox + t_au + t_ntp); # width of gap
mesh_margin = 4e-9; # buffer
# total thickness of all despsited layers
t_coating = t_alox + t_au + t_ntp;

select("waveguide");
set("y_pos", y_pos);
set("wg_width", wg_width);
set("wg_height", wg_height);
set("wg_length", wg_length);
set("t_siox", t_siox);

select("alumina");
set("y_pos", y_pos);
set("z_pos", z_pos);
set("t_alox", t_alox);
set("wg_width", wg_width);
set("wg_height", wg_height);
set("wg_length", wg_length);

select("plasmon");
set("y_pos", y_pos);
set("z_pos", z_pos);
set("t_au", t_alox + t_au);
set("wg_width", wg_width);
set("wg_height", wg_height);
set("wg_length", wg_length);

select("ntp");
set("y_pos", y_pos);
set("z_pos", z_pos);
set("t_ntp", t_alox + t_au + t_ntp);
set("wg_width", wg_width);
set("wg_height", wg_height);
set("wg_length", wg_length);

select("mesh_slot");
set("y span", wg_gap + mesh_margin);
set("z span", wg_height +t_coating + mesh_margin);
set("z", 0);
# set("z", t_coating);

select("FDE");
set("y span", y_span);
set("z span", z_span);
y_num = y_span / y_meshsize;
z_num = z_span / z_meshsize;
set("mesh cells y", y_num);
set("mesh cells z", z_num);
```

Example of Lumerical Script

The following is an example of the 1D script used to calculate the conversion efficiency η from Equation (3.3). 2D scripts were also utilized, and in some instances the Stokes shift (i.e., η calculated after Equation (3.2)) was accounted for as well. These two other scripts are not included here, but are available through my GitHub page: https://github.com/joakimekern/sersmask.

```
# ===========
# Script to extract the relevant results of a waveguide as one or
# more variables are sweeped
# The sweep properties can be set under 'Sweep parameters'
# The injection axis is assumed to be x, the height of the
# waveguide is assumed to be in z, and the width in y
# The formula for the conversion efficiency is found in
# "Efficiency of evanescent excitation and collection of
# spontaneous Raman scattering near high index contrast
# channel waveguides" (2015), by A. Dhakal et al.
# ===========
clear;

#==================
# Input parameters
#==================

selected_mode = "mode1";
pump_lambda = 785e-9;
cladding_index = 1;         # Air
#cladding_index = 1.538;     # NTP
index_margin = 0.01;

# Geometry
injection_axis = "x";   # Not used
width_axis = "y";
height_axis = "z";

# Surface. Define the number of mesh cells in x and y which define
#the surface of the wg structure
meshcells_x = 10;
meshcells_y = 10;

# Sweep parameters
parameter_name = "wg_width";
parameter_scope = "::model";
#parameter_values = linspace(350, 700, 36) * 1e-9;
parameter_values = [500] * 1e-9;

# Save results
save_to_json = false;
savename = "test.json";

#==================
# Functions
#==================

# A filter seperating out the bulk region of interest
function bulk_filter(n_reg_interest, n_margin){
```

```
    # Get the refractive index and the axis lengths
    ri = real(pinch(getdata("FDE::data::material","index_x")));
    dim = size(ri);
    x = dim(1);
    y = dim(2);
    # Create the matrix to compare with the refractive index
    background_matrix = ones(x, y) * n_reg_interest;

    # Create the filter.
    # All grid points with an index value equal (almost) to the
    # background index is set to 1.
    # All other points are set to 0.
    filter = almostequal(ri, background_matrix, n_margin, n_margin);

    return filter;
}

function shift_matrix(m, num, direction){
    # Get the dimensions of the original matrix
    dim = size(m);
    x = dim(1);
    y = dim(2);
    # Create a new matrix
    m2 = ones(x, y);

    # Shift the old matrix
    if(direction == "up"){
        m2(1:x, num+1:y) = m(1:x, 1:y-num);
    }
    else if(direction == "left"){
        m2(1:x-num, 1:y) = m(num+1:x, 1:y);
    }
    else if(direction == "right"){
        m2(num+1:x, 1:y) = m(1:x-num, 1:y);
    }
    else if(direction == "down"){
        m2(1:x, 1:y-num+1) = m(1:x, num:y);
    }
    else {return 0;}

    # Invert the matrix and return. This ensures the new edges are 0
    m2 = m2 == 0;
    return m2;
}

# Create a surface filter in a region of interest. The surface is
# defined by 'xshift' number of mesh cells horizontally,
# and 'yshift' number of mesh cells vertically
function surface_filter(n_reg_interest, n_marg, xshift, yshift){
    # Initialize a filter seperating out the region of interest
    mask = bulk_filter(n_reg_interest, n_marg);

    # Make masks shifting the filter in four directions
    topmask = shift_matrix(mask, yshift, "up");
    rightmask = shift_matrix(mask, xshift, "right");
    leftmask = shift_matrix(mask, xshift, "left");
    downmask = shift_matrix(mask, yshift, "down");
    # Combine the shifted masks
```

```
    ormask = topmask + rightmask + leftmask + downmask;
    # Do boolean operations to create the final mask
    #final = mask * ormask;
    #final = final != 0;
    f1 = ormask != 4;
    f2 = ormask != 0;
    final = f1 * f2;

    return final * mask;
}

# Function to get the inverse effective mode area of a given mode.
# x and y are the cross-section axes
function get_inv_eff_mode_area(mode, x, y){
    # Get the electric field amplitude
    Enorm2 = pinch(getelectric("FDE::data::" + mode));
    # Calculate the permittivity from the RI
    index = getdata("FDE::data::material", "index_z");
    permittivity = sqrt(real(pinch(index)));

    # Get |E|^2 times the permittivity (epsilon) at each grid point
    # and integrate to get the total charge
    charge = Enorm2 * permittivity;
    charge = integrate(charge, [1, 2], x, y);

    return Enorm2 / charge;
}

# Returns the integrated luminosity of a given mode. Both the pump
# and Stokes shift can be taken into account, in which case 1
# denotes pump and 2 denotes Stokes. area is the inverse effective
# mode area. area1 and area2 should be cross-sections over the
# same area, and are thus integratable
function get_integrated_luminosity(area1, area2,
                                   lambda2, ng1, ng2, n2){
    # Initialize the constants
    kv = 1.26e23;
    constants = pi^2 / eps0^2 / kv;
    ref_indices = ng1 * ng2 / n2;

    return constants * ref_indices * lambda2^2 * area1 * area2;
}

function get_integrated_luminosity_noconst(area1, area2,
                                           filter, x, y){
    lum = integrate(area1 * area2 * filter, 1:2, x, y);
    return lum;
}

function get_int_lum_constants(mode){
    loss = getdata("FDE::data::" + mode, "loss");
    ng = real(getdata("FDE::data::" + mode, "ng"));
    neff = real(getdata("FDE::data::" + mode, "neff"));

    return {"loss": loss, "ng": ng, "neff": neff};
}

# Function to simply calculate the Dhakal prefactor, as it depends
```

```
# on whether or not the Stokes shift is taken into consideration
function dhakal_factor(ng1, ng2, neff1, neff2, l_p, l_s){
    if(ng2 == 0){
        ng2 = ng1;
        neff2 = neff1;
        l_s = l_p;
    }
    return ng1 * ng2 * l_s^2 / neff2;
}

# Given a parameter and its value, set it and calculate the number
# of modes. The wavelength  must be specified as well
function set_and_run(scope, name, value, wavelength){
    # Set the new parameter
    switchtolayout;
    setnamed(scope, name, value);
    setanalysis("wavelength", wavelength);
    # Run
    run;
    n = findmodes;

    return n;
}

function calculate_results(mode, xaxis, yaxis, n_clad, n_marg, shiftx,
  shifty, use_stokes, lambda_p, lambda_s){
    # Get the cross section
        x = pinch(getdata("FDE::data::material",xaxis));
        y = pinch(getdata("FDE::data::material",yaxis));
        # Make the filter
        filter_bulk = bulk_filter(n_clad, n_marg);
 # Always uses 1, air. For NTP, it's better to use bulk filter
        filter_surf = surface_filter(1, n_marg, shiftx, shifty);

        # Get the constants
        constants = get_int_lum_constants(mode);
        # [loss, group index, effective index]

        # Get the mode area
        area = get_inv_eff_mode_area(mode, x, y);


        # Calculate the integrated luminosity
        lum_bulk = integrate(area_pump^2 * filter_bulk, 1:2, x, y);
        lum_surf = integrate(area_pump^2 * filter_surf, 1:2, x, y);

        return [lum_bulk_p; lum_surf_p; lum_bulk_s; lum_surf_s;
            constants_p.loss; constants_p.ng; constants_p.neff;
            constants_s(1); constants_s(2); constants_s(3);
      n_modes_stokes];
}

#==================
# Script
#==================

n_val = length(parameter_values);
# Arrays to store the sweep results
```

```
integrated_luminosity_bulk = matrix(n_val);
integrated_luminosity_surf = matrix(n_val);
# Arrays to store the loss, group index, and effective index at
# the given wavelength
loss = matrix(n_val);
ng = matrix(n_val);
neff = matrix(n_val);
n_modes = matrix(n_val);

# Initialize the constant used in Dhakal et al.
kv = 1.26e23;
dhakal_const = pi^2 / eps0^2 / kv;

# Run the sweep
for(i=1; i <= n_val; i=i+1){
    ?"Now analysing parameter '" + parameter_name + "' with value '"
  + num2str(parameter_values(i)) + "'";
    # Set the new parameter
    n_modes_p(i) = set_and_run(parameter_scope, parameter_name,
   parameter_values(i), pump_lambda);

    # If no modes are found, do not throw error, simply append 0.
    if(n_modes_p(i) != 0){

        # Calculate and store the results
        res = calculate_results(selected_mode, width_axis,
                height_axis, cladding_index, index_margin,
                meshcells_x, meshcells_y, include_stokes,
                pump_lambda, stokes_lambda);
        loss_p(i) = res(5);
        loss_s(i) = res(8);
        ng_p(i) = res(6);
        ng_s(i) = res(9);
        neff_p(i) = res(7);
        neff_s(i) = res(10);
        n_modes_s(i) = res(11);

        # Calculate the conversion efficiency as in Dhakal
        if (include_stokes){
            integrated_luminosity_bulk(i) = res(3);
            integrated_luminosity_surf(i) = res(4);
            factor = dhakal_const * dhakal_factor(ng_p(i), ng_s(i),
       neff_p(i), neff_s(i), pump_lambda, stokes_lambda);
        } else {
            integrated_luminosity_bulk(i) = res(1);
            integrated_luminosity_surf(i) = res(2);
            factor = dhakal_const * dhakal_factor(ng_p(i), ng_p(i),
              neff_p(i), neff_p(i), pump_lambda, pump_lambda);
        }

        ce(i) = factor * integrated_luminosity_bulk(i);
 if(save_to_json){
            jsonsave(savename, parameter_name, parameter_values, ce,
                     integrated_luminosity_bulk,
                    integrated_luminosity_surf, loss_p, ng_p,
                     neff_p, loss_s, ng_s, neff_s, n_modes_p,
                     n_modes_s, pump_lambda, stokes_lambda,
                     dhakal_const, cladding_index);
```

```
        }
    }
}

# Plot results
#plot(parameter_values, ce);
```

Appendix C

Additional Results Computed in Lumerical

The plots depicted in this part of the appendix are included to show how Lumerical was utilized to arrive at the simulation parameters described in Chapter 4. The Tables 4.1 and 4.2 should be consulted for an explanation of the variables used here.

The plots in Figure C.7 show how the conversion efficency changes for different values of the waveguide dimensions (gap and width), and different thicknessess of the alumina and gold layers. The color of the graphs denote which data is being plot. NTP refers to the inclusion or exclusion of the 1 nm modelled NTP layer, while the terms "Bulk" and "Surface" refer to two different computation algorithms used in the preliminary calculations. "Bulk" means that the cladding integral in Equation (3.3) was over the entire cladding region in the simulation domain, while "Surface" means that the integral was over an area close (on the scale of nm) to the surface. Ultimately, "Bulk" was chosen, and all results depicted in Chapter 4 used that computing algorithm. This was for simplicity, because the results were relatively similar, and the matrix filter used to capture the surface region was not fully developed and tested.

Figures C.8 and C.9 explore the differences between including and excluding the Stokes shift. When a graph is denoted with "Stokes" Equation (3.2) was used to compute η, and when a graph is denoted with "Nostokes" Equation (3.3) was used. Clearly, the difference in including the Stokes shift is not particularly large, as stated by Dhakal *et al.* [13]. Therefore, all results depicted in Section 4.3 excluded the Stokes shift.

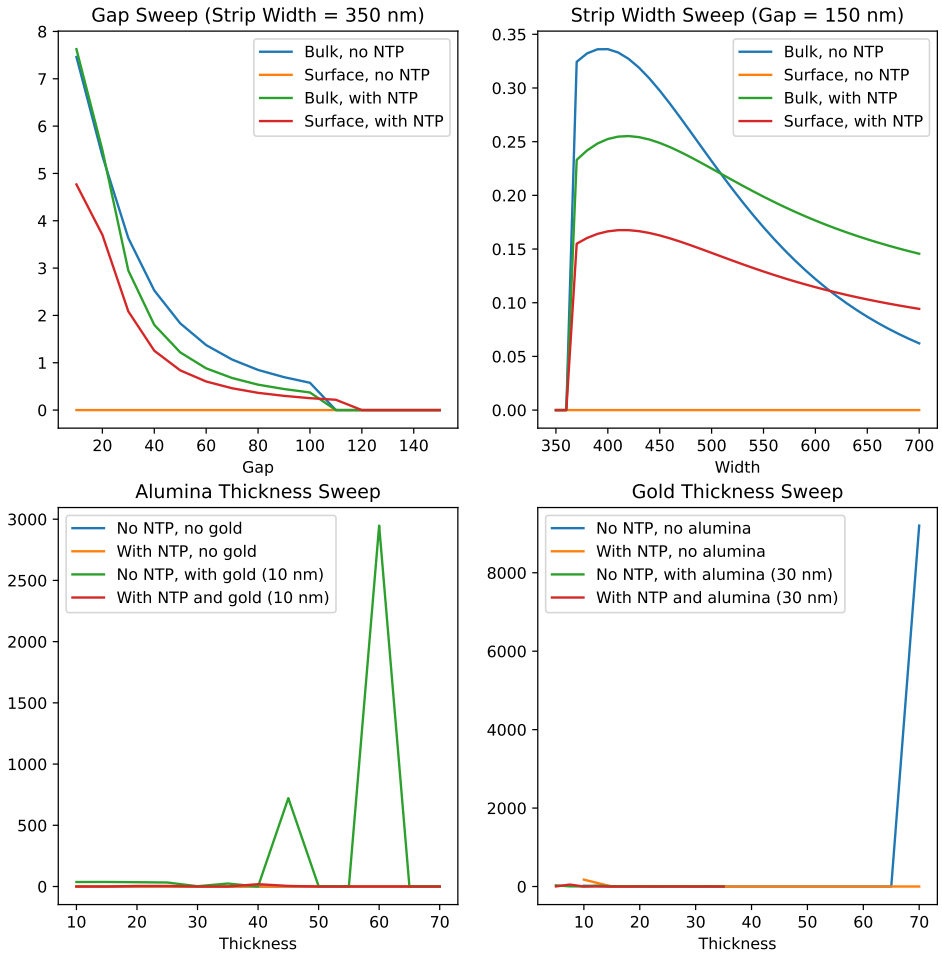


Figure C.7: Preliminary sweeps of the conversion efficiency η. The y axes depict η in units of steradians, while the x axes depict the thickness of either the gold or the alumina layer, given in nm.

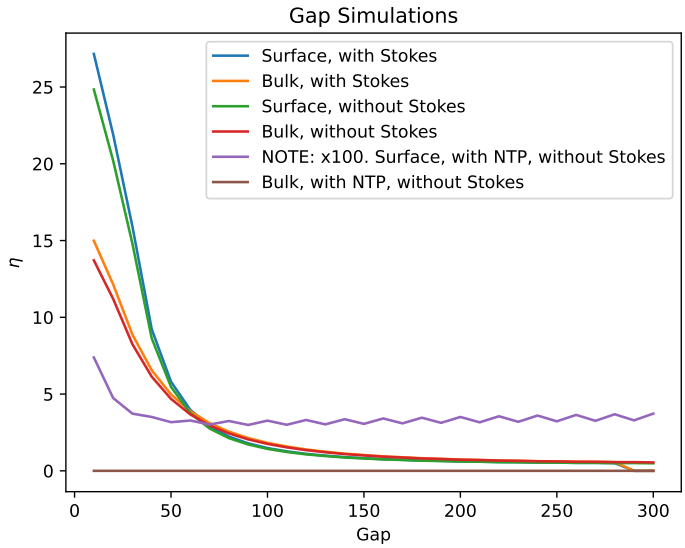


Figure C.8: Preliminary sweeps of the conversion efficiency for different gap sizes. The y axis is given in steradians, and the x axis in nm.

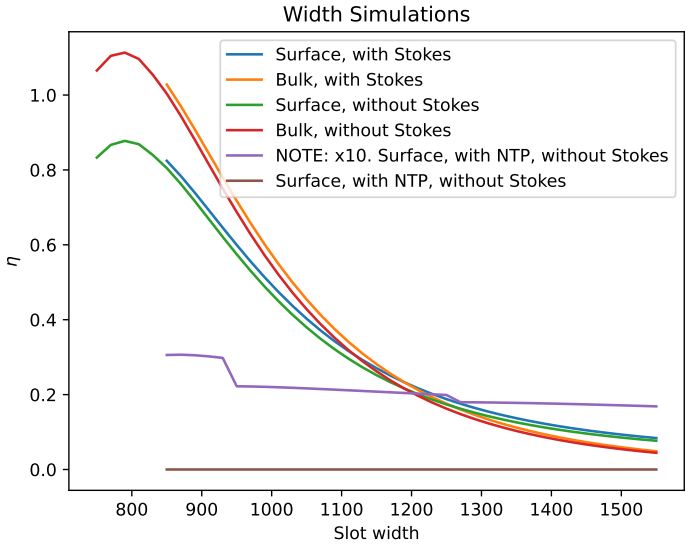
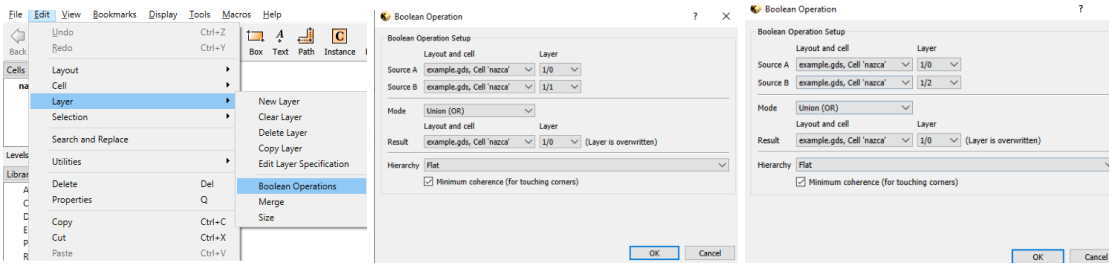


Figure C.9: Preliminary sweeps of the conversion efficiency for different waveguide widths (total slot with, i.e., $2r + g$). The y axis is given in steradians, and the x axis in nm.



(a) (b) (c)

Figure D.10: Example of how to do the Boolean operations in KLayout. (a) shows how to open the *Boolean operations* interface. (b) and (c) shows the two operations that need to be executed.

Appendix D

Waveguide Design Code

Class to design a slot waveguide using the Python library Nazca (note that it is not in the pip libraries, and must be downloaded and installed separately). The code is included in this thesis only for completeness, since the recommended way to access this code is through my GitHub profile at https://github.com/joakimekern/sersmask.

A single class is defined in *sersmask.py*, the *SlotWaveguide*. Its arguments are all the relevant parameters needed to create a slot waveguide intended for SERS sensing. An example of how to use the class is found in *example_mask.py*.

Note however, that every design must be subject to XOR logical operation on layers (1, 0), (1, 1), and (1, 2) to produce the final mask. Nazca does not provide the possibility of Boolean operations, so these operations must be done manually. One way to do this is with KLayout, which is a free mask file editing software. Opening the final mask with Klayout, go to Edit → Layers → Boolean Operations. Then select 1/0 as source A and 1/1 (or 1/2) as source B, XOR as mode and 1/0 as result, see Figure D.10. Doing the Boolean operation twice (once on 1/1 and once on 1/2) produces the final mask.

sersmaks.py

Class declaration.

```
import nazca as nd


class SlotWaveguide:

    # Name of the cross-sections.
    # Shared by all 'SlotWaveguide' objects
    _xs_wg = 'wg'
    _xs_t = 'taper'
    _xs_to = 'taper_out'
    _xs_tg = 't_gap'
    _xs_alox = 'alox'
    _xs_au = 'au'
```

```
    # Initialize the cross-sections
    nd.add_xsection(name=_xs_wg)
    nd.add_xsection(name=_xs_t)
    nd.add_xsection(name=_xs_tg)
    nd.add_xsection(name=_xs_alox)
    nd.add_xsection(name=_xs_au)
    # Keyword to put the next element on top of the previous
    putback = 'putback'

    def __init__(self, width: float,
                 gap: float,
                 length: float,
                 buffer: float,
                 label=False,
                 label_height=3.0,
                 accuracy=0.001,
                 point=(0, 0),
                 entrance=0.0,
                 taper=False,
                 taper_length=None,
                 taper_access=None,
                 taper_width=None,
                 taper_buffer=None,
                 taper_out=False,
                 taper_out_width=None,
                 taper_out_length=None,
                 taper_out_buffer=0.0,
                 bend=False,
                 bend_angle=None,
                 bend_exit=None,
                 bend_sep=None,
                 alumina=False,
                 alumina_length=100.0,
                 gold=False,
                 gold_length=50.0,
                 metal_width=None,
                 out_length=0.0)
                 -> None:

        # Position attributes
        self.start = point
        self.end = point

        # General attributes
        self.asdict = {
            'width': width,
            'gap': gap,
            'length': length,
            'slot': 2 * width + gap,
            'buffer': buffer + width + gap / 2,
            # taper_access is depreciated
            'input': entrance if not taper_access or entrance != 0
                     else taper_access,
            # bend_exit is depreciated
            'output': bend_exit if out_length == 0 and bend_exit
                      else out_length,
            'start': self.start,
            'end': self.end,
```

```
        'label': label,  # Not implemented
        'label_height': label_height  # Not implemented
    }

    # Taper attributes
    self.taper = {'length': taper_length if taper else 0,
                  'width': taper_width if taper else 0,
                  'buffer': taper_buffer if taper else 0}
    self.taperout = {'length': taper_out_length if taper_out else 0,
                     'width': taper_out_width if taper_out else 0,
                    'buffer': taper_out_buffer if taper_out else 0}

    # Euler bend attributes
    self.bend = {'angle': bend_angle if bend else 0,
                 'sep': bend_sep if bend else 0}

    # Metal mask attributes. The 'min' conditions ensures
    # that the metal masks are within the active region
    alumina_length = min(alumina_length, length) if alumina else 0
    gold_length = min(gold_length, length) if gold else 0
    self.metals = {'alumina': alumina_length,
                   'gold': gold_length,
                   'length': alumina_length + gold_length,
                   'width': metal_width if metal_width
                    else self.asdict['buffer']}
     # Store shapes to be inserted with the 'put' method.
    self.shapes = []
    self._xs_active = self._xs_wg if not self.taper['length']
      else self._xs_t  # Currently active cross-section

    # Make the waveguide and taper layers. Boolean 'XOR'
    # operations (in for instance KLayout) must be used on
    # the layers to make the slot. Using XOR on layers
    # means that sections with an odd number of overlapping
    # layers will be preserved, while an even number will
    # be removed.
    self._initialize_layers(accuracy=accuracy)

    # Add the entrance
    self._add_straight(length=self.asdict['input'], \
      width=self.taper['width'] if self.taper['width'] else 0)

    # Add a taper at the input
    if self.taper['length']:
        self._taper(self.taper['length'], self.taper['width'],
                    self.asdict['slot'],
                    gapw1=0, gapw2=self.asdict['gap'],
                    xs=self._xs_t, xsg=self._xs_tg)

    # Active region
    self._active(self.asdict['length'],
                 metal_len=self.metals['length'],
                 alox_len=self.metals['alumina'],
                 au_len=self.metals['gold'],
                 xs_wg=self._xs_wg,
                 xs_alox=self._xs_alox,
                 xs_au=self._xs_au)
```

```
        # Add the taper after the active region
        if self.taperout['length']:
            self._taper(self.taperout['length'],
                        self.asdict['slot'],
                        self.taperout['width'],
                        gapw1=self.asdict['gap'],
                        gapw2=0,
                        xs=self._xs_to,
                        xsg=self._xs_tg)

        # Add the Euler bend to shift the output from the input.
        # The cross-section used is the one currently active
        if self.bend['angle']:
            self._euler(angle=self.bend['angle'],
                        ysep=self.bend['sep'],
                        width=self.taperout['width'] if self.taperout['width']
                        else 0)

        # Add the output section
        self._add_straight(length=self.asdict['output'],
                           width=self.taperout['width'] /
                            if self.taperout['width'] else 0)
        # self._add_straight(xs=xs_wg, length=marg)
        # Insert structure
        self._put()

    def _initialize_layers(self, accuracy=0.001) -> None:
        """ Initializes all layers, with the given accuracy
            (in um). The active region contains the slot, and
            possibly metal masks. The taper cross-sections
            are initialized as well """

        for i, xs in enumerate((self._xs_wg, self._xs_alox,
                self._xs_au)):
            # Add the three waveguide layers
            nd.add_layer2xsection(xsection=xs, layer=(1, 0),
                  accuracy=accuracy, growx=self.asdict['gap'] / 2)
            nd.add_layer2xsection(xsection=xs, layer=(1, 1),
                  accuracy=accuracy, growx=self.asdict['slot'] / 2)
            nd.add_layer2xsection(xsection=xs, layer=(1, 2),
                  accuracy=accuracy, growx=self.asdict['buffer'])
            # Add the alumina layer
            if i != 0 and self.metals['alumina']:
                nd.add_layer2xsection(xsection=xs, layer=(2, 0),
                  accuracy=accuracy, growx=self.metals['width'])
            # Add the gold layer
            if i == 2 and self.metals['gold']:
                nd.add_layer2xsection(xsection=xs, layer=(3, 0),
                  accuracy=accuracy, growx=self.metals['width'])
        # Add the main taper layer (for both input and output) and the gap layer
        if self.taper['length'] or self.taperout['length']:
            nd.add_layer2xsection(xsection=self._xs_t,
              layer=(1, 0), accuracy=accuracy)
            nd.add_layer2xsection(xsection=self._xs_to,
              layer=(1, 0), accuracy=accuracy)
            nd.add_layer2xsection(xsection=self._xs_tg, layer=(1, 2),
              accuracy=accuracy)
        # Add the taper buffer layers
```

```
        if self.taper['length']:
            nd.add_layer2xsection(xsection=self._xs_t, layer=(1, 1),
              accuracy=accuracy, growx=self.taper['buffer'])
        if self.taperout['length']:
            nd.add_layer2xsection(xsection=self._xs_to, layer=(1, 1),
              accuracy=accuracy, growx=self.taperout['buffer'])

    def _add_straight(self, length=0.0, width=0.0, xs=None, layer=None,
                      edge1=None, edge2=None, edgepoints=50,
                      name=None) -> None:
        """ Simple function to append straight shapes to 'self.shapes'.
            It is practically the 'nd.strt()' function
            but notice that the default width and length are 0.
            In addition, the default cross-section is '_xs_wg' """

        xs = self._xs_active if not xs else xs
        self._xs_active = xs
        self.shapes.append(nd.strt(length=length,
                           width=width, xs=xs, layer=layer,
                           edge1=edge1, edge2=edge2,
                            edgepoints=edgepoints, name=name))

    def _add_taper(self, length: [int, float], startwidth: [int, float],
                   endwidth: [int, float],
                  xs=None) -> None:
        """ Identical to '_add_straight()', but this function adds a taper """

        xs = self._xs_active if not xs else xs
        self._xs_active = xs
        self.shapes.append(nd.taper(length=length, \
              width1=startwidth, width2=endwidth, xs=xs))

    def _add_euler_bend(self, xs=None, width=0, angle=45) -> None:
        """ Identical to '_add_straight()',
            but this function adds an Euler bend """

        xs = self._xs_active if not xs else xs
        self._xs_active = xs
        self.shapes.append(nd.euler(angle=angle, \
                xs=xs, width1=width, width2=width, radius=50))

    def _active(self, length: [int, float], metal_len=0.0,
                alox_len=0.0, au_len=0.0, xs_wg=None,
                xs_alox=None, xs_au=None):
        """ Adds the active region, including the metals
            masks if they exist"""

        marg = (length - metal_len) / 2
        alox_marg = (alox_len - au_len) / 2

        # Add the active region
        self._add_straight(xs=xs_wg, length=marg)
        if metal_len:
            self._metals(au_len, alox_marg, \
            xs_alox=xs_alox, xs_au=xs_au)
        self._add_straight(xs=xs_wg, length=marg)

    def _taper(self, length: [int, float], width1: [int, float],
```

```
              width2: [int, float],
              gapw1=None, gapw2=None, xs=None,
              xsg=None) -> None:
        """ Adds a taper region with the specified start and end
            widths. Supports adding a taper gap if defining
            the widths 'gapw1' and 'gapw2'. The main taper will
            be put in the cross-section 'xs' and the gap
            in 'xgs'. """

        if gapw1 or gapw2:
            # Keyword to put the gap taper on top of the main taper
            self.shapes.append(self.putback)
            self._add_taper(length, gapw1, gapw2, xs=xsg)
        # Make main taper
        self._add_taper(length, width1, width2, xs=xs)

    def _euler(self, xs=None, width=0, angle=45, ysep=0.0) -> None:
        """ Adds an Euler bend to shift the direction of the guided wave """

        xs = self._xs_active if not xs else xs
        self._xs_active = xs
        # First Euler bend
        self._add_euler_bend(xs=xs, width=width, angle=angle)
        # Increased vertical separation of input and exit if needed
        self._add_straight(length=ysep, width=width)
        # Second Euler bend
        self._add_euler_bend(width=width, angle=-angle)

    def _metals(self, gold_length: [int, float],
                margin: [int, float],
                xs_alox=None, xs_au=None) -> None:
        """ Add the metal masks """

        if self.metals['alumina']:
            self._add_straight(length=margin, xs=xs_alox)
        if self.metals['gold']:
            self._add_straight(length=gold_length, xs=xs_au)
        if self.metals['alumina']:
            self._add_straight(length=margin, xs=xs_alox)

    def _put(self) -> None:
        """ Insert all shapes to the design. Handles gap tapers
              put on top of main tapers using 'put_back' """

        put_back = False
        backpoint = nd.cp.xy()

        for i, shape in enumerate(self.shapes):
            if i == 0:
                shape.put(self.asdict['start'])
            elif put_back:
                shape.put(backpoint)
                put_back = False
            elif shape != self.putback:
                # Store this point if the next shape should be
                # put on top. Note that this is bad Nazca practice
                if i > 0 and self.shapes[i - 1] == self.putback:
                    backpoint = nd.cp.xy()
```

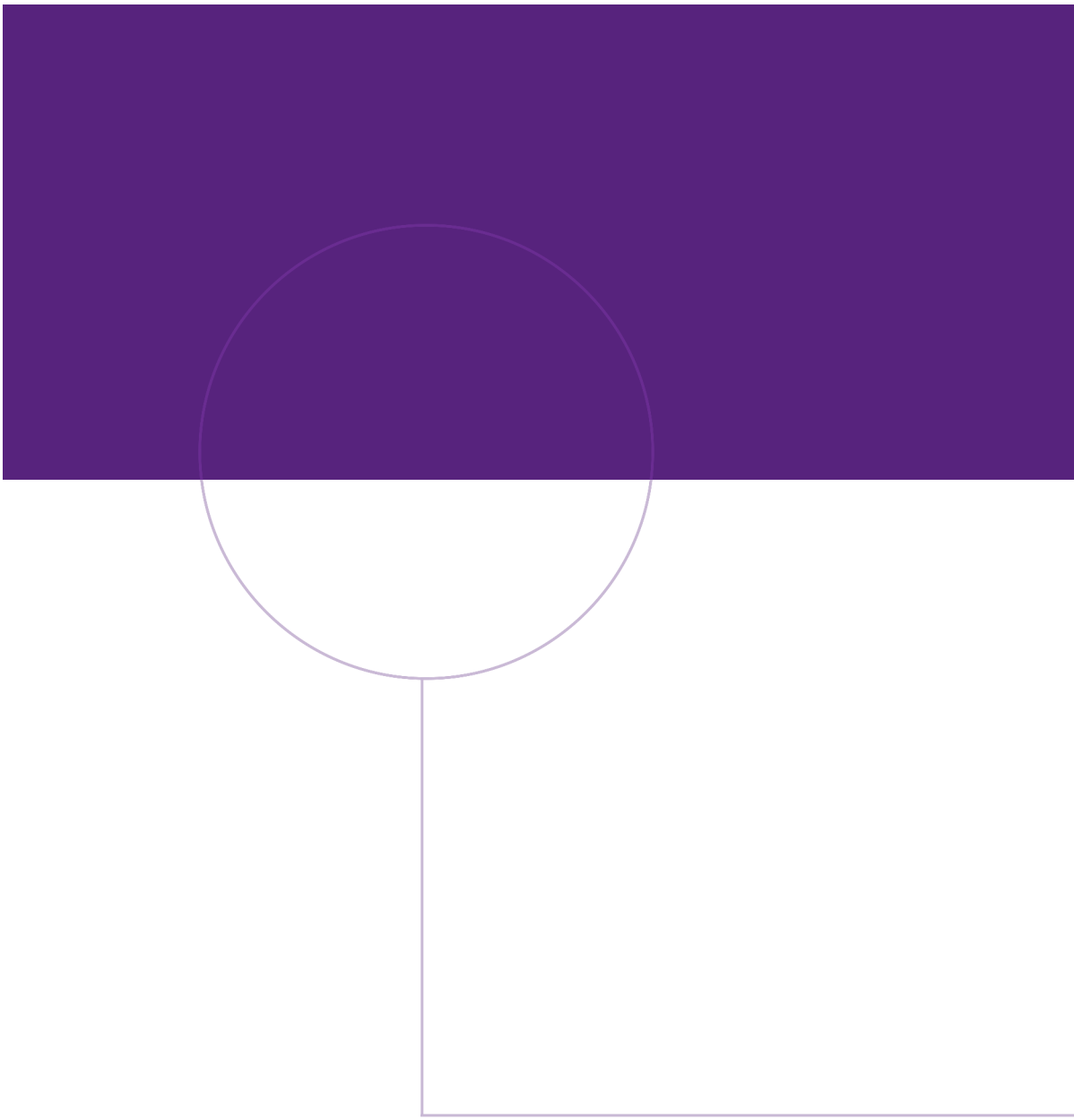
```
                    put_back = True
                shape.put()
        # Store endpoint
        self.end = nd.cp.xy()
        self.asdict['end'] = nd.cp.xy()
```

example_mask.py

An example script which makes 8 different waveguides at different positions

```
# NOTE: Not in the pip libraries. See installation guide at:
# https://nazca-design.org/installation/
import nazca as nd
from sersmask import SlotWaveguide

# Specify the file to which the design is written
export_path = ''
export_name = 'example.gds'
export = export_path + export_name

# Active region properties. All units are in microns
wg_width = 0.5      # Width of each rail in the slot
wg_gap = 0.5        # Distance between the rails
wg_length = 100     # Length of the active region
buffer = 3          # Buffer region separating the waveguide
#                     structure from the SiN thin film

# Multiple waveguides
number_of_waveguides = 8
waveguide_separation = 50   # In microns

# Set optional properties. All units are in microns
prop = {
    'entrance': 50,     # Input length
    'out_length': 100,  # Output length
    # The start point, i.e., the westernmost point of the waveguide
    'point': (0, 0),
    'label': False,     # Not implemented
    'label_height': 4,  # Not implemented
    # Which components should be included?
    'taper': True,      # Include input taper
    'taper_out': True,  # Include output taper
    'bend': True,       # Include Euler bend
    'alumina': False,   # Include alumina layer mask
    'gold': False,      # Include gold layer mask
    # Taper properties
    'taper_length': 100,    # Length of the taper
    'taper_width': 20,      # Start width
    # Buffer region between the taper and SiN thin film
    'taper_buffer': 5,
    'taper_out_length': 50, # Length of the output taper
    'taper_out_width': 10,  # End width of the output taper
    'taper_out_buffer': 10, # Output buffer
    # Euler bend properties
    'bend_angle': 30,
    # Additional vertical seperation between the Euler bends
    'bend_sep': 10,
    # Metal properties
    'alumina_length': 100,
    'gold_length': 50,
}

# Pack the waveguide properties
args = [wg_width, wg_gap, wg_length, buffer]
```

```
kwargs = prop

# Store the waveguide properties
waveguides = {i: [args.copy(), kwargs.copy()] \
              for i in range(number_of_waveguides)}

# Modify the waveguide properties.
# Removing these declarations result in the creation of
# identical waveguides
waveguides[1][1]['taper'] = False

waveguides[2][1]['taper_out'] = False

waveguides[3][1]['bend'] = False

waveguides[4][1]['taper'] = False
waveguides[4][1]['taper_out'] = False

waveguides[5][1]['taper_out'] = False
waveguides[5][1]['bend'] = False

waveguides[6][1]['bend'] = False
waveguides[6][1]['taper'] = False

waveguides[7][1]['taper'] = False
waveguides[7][1]['taper_out'] = False
waveguides[7][1]['bend'] = False

if __name__ == '__main__':
    # Iterate through the waveguides and place them in the design
    for i, (a, k) in waveguides.items():
        wg = SlotWaveguide(*a, **k)
        # Place the next waveguide above the previous
        if i < number_of_waveguides - 1:
            waveguides[i + 1][1]['point'] = (wg.start[0], \
                  wg.end[1] + waveguide_separation)
    nd.export_gds(filename=export)
```



NTNU

Norwegian University of
Science and Technology