UNIVERSITEIT
GENT

Universiteit Gent
Faculteit Ingenieurswetenschappen
Vakgroep Informatietechnologie

Promotoren

Prof. Dr. Ir. P. Bienstman	Universiteit Gent, INTEC
Prof. Dr. Ir. R. Baets	Universiteit Gent, INTEC

Examencommissie

Prof. Dr. Ir. D. De Zutter (voorzitter)	Universiteit Gent, INTEC
Prof. Dr. P. Dubruel (secretaris)	Universiteit Gent, Department of Organic Chemistry
Prof. Dr. Ir. D. Hill	KTH, Royal Institute of Technology
Prof. Dr. Ir. L. Lagae	KU Leuven, IMEC
Prof. Dr. Ir. K. Neyts	Universiteit Gent, ELIS
Prof. Dr. Ir. R. Baets (promotor)	Universiteit Gent, INTEC
Prof. Dr. Ir. P. Bienstman (promotor)	Universiteit Gent, INTEC

Universiteit Gent
Faculteit Ingenieurswetenschappen

Vakgroep Informatietechnologie
Sint-Pietersnieuwstraat 41, B-9050 Gent, België

Tel.: +32-9-264.33.35
Fax.: +32-9-264.35.93

Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen:
Fotonica
Academiejaar 2009-2010

Dankwoord

Alhoewel mijn naam de enige is die vooraan op de kaft van die boekje prijkt, zijn er toch heel veel mensen die indirect of direct een belangrijke bijdrage tot dit werk hebben geleverd. Ze allemaal opnoemen is een onmogelijke opgave, maar laat ik toch een kleine poging wagen.

Uiteraard moet ik allereerst mijn promotoren Roel en Peter bedanken om me de kans te willen geven om een doctoraat in de fotonica aan te vatten, en mij te blijven begeleiden en sturen zodat ik het ook effectief heb kunnen afwerken. Peter, bedankt om me in te wijden in de mystieke beginselen van CAMFR. Voor ik hier begon wist ik eerlijk gezegd helemaal niets van plasmonica en het idee om metaal te gebruiken om licht te geleiden leek nogal uit de lucht gegrepen. Dankzij jullie heb ik mijn eerste stappen kunnen zetten in wat later een boeiende wetenschappelijke tak bleek te zijn.

Tijdens mijn vijfjarige doortocht in de fotonica groep in Gent heb ik uiteraard heel veel collega's zien komen en gaan. Iedereen bedanken waar ik ooit mee samengewerkt heb in al die jaren is onbegonnen werk, iedereen beschrijven die deel uitmaakt(e) van de fotonica groep nog veel moeilijker. Een heel erg collegiaal, hard werkend sympathiek zootje ongeregeld komt nog het best in de buurt. Ik wil jullie allemaal heel erg bedanken om van onze groep een plezante en ongedwongen werkomgeving te maken, en dan heb niet over de vele wetenschappelijke discussies, maar vooral over de vele uren die we al zeverend doorbrachten in de koffiekamer en in talloze cafés in de Gentse binnenstad. Altijd was wel iemand te vinden om een pintje te gaan drinken na het werk, af en toe werden er dat een paar teveel (en werd er al eens een trein gemist). Echt merci allemaal, als ik in de buurt ben spring ik zeker nog eens binnen.

Bedankt ook bureaugenootjes voor de ambiance, en dan denk ik vooral aan Jonathan en Frederik die met mij waarschijnlijk wel het langst daar helemaal vanachter in de gang gezeten hebben, elk in ons eigen hoekje. Stijn, Katrien, Tom, de rest mag dan al eens meewarig gelachen hebben met

al dat biogedoe, bedankt voor de fijne samenwerking, ik kon me geen betere collega's indenken.

Steven en Liesbet, moesten jullie er niet zijn geweest dan waren mijn ideeën gewoon fantasietjes gebleven. Bedankt om samen met mij uren naar het FIB beeldscherm te staren, me te vergeven als ik weer eens iets kapot gemaakt had in de cleanroom en altijd beschikbaar te zijn om me met raad en daad bij te staan. Jullie bijdrage in deze groep is van onschatbaar belang.

Het GOA project van de Universiteit Gent bood me de kans om in een multidisciplinaire wetenschapstak te duiken. Een heel leerrijke en boeiende ervaring, waar het me vooral duidelijk werd dat je heel veel verschillende mensen nodig hebt met verschillende achtergronden om een project tot een goed einde te brengen. De eerste keer dat we vergaderden verstonden we elkaar helemaal niet, maar ondertussen denk ik dat we al een beetje dichter bij elkaar zijn gekomen. In het bijzonder wil ik Jordi bedanken, die zijn (voor mij toch) magische chemie op mijn samples losliet.

Uiteraard wil ik graag mijn ouders bedanken, om me de kans te geven om zo ongelofelijk lang 'naar school te blijven gaan' en me gedurende al die jaren te steunen. Helemaal op het einde, maar super belangrijk, Daphne...bedankt om altijd in mij te blijven geloven, bedankt om er altijd voor mij te zijn en bedankt om samen met mij de wereld rond te trekken.

Roeselare, 17 december 2010
Peter Debackere

Table of Contents

List of Figures

List of Tables

List of Acronyms

A

AFM	Atomic force microscope
ASR	Adaptive spatial resolution

B

BCB	Benzocyclobutene
BOX	Buried oxide layer

C

CCD	Charge-coupled device
CMP	Chemical mechanical polish
CAMFR	Cavity modeling framework
CMOS	Complementary metal-oxide semiconductor

D

DALY	Disability-adjusted life years
DLL	Dynamic link library
DVS-BCB	Divinylsiloxane-bis-benzocyclobutene

DUV	Deep ultra-violet

F

FDFD	Finite difference frequency domain
FEFD	Finite element frequency domain
FMM	Fourier modal method
FWHM	Full-width at half maximum

G

GPIB	General purpose input bus

I

ICP RIE	Inductive coupled plasma reactive ion etching

L

LOC	Lab on chip

M

MHA	Mercaptohexadecanoic
MZI	Mach Zehnder interferometer

P

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDMS	Polydimethylsiloxane
PECVD	Plasma enhanced chemical vapour deposition
POC	Point of care
PML	Perfectly matched layer
PMMA	Polymethylmethacrylate

R

RCWA	Rigorouos coupled wave analysis
RF	Radio frequency
RIE	Reactive ion etching
RIU	Refractive index unit

S

SAM	Self assembled monolayer
SEM	Scanning electron microscope
SOI	Silicon-on-insulator
SNR	Signal-to-noise ratio
SPI	Surface plasmon interferometer
SPP	Surface plasmon polariton
SPR	Surface plasmon resonance

T

TE	Transverse electric
TM	Transverse magnetic

U

USB	Universal serial bus

V

VISA	Virtual instrument software architecture

Nederlandse samenvatting

HEDENDAAGSE interesse naar verschillende detectiemethodes die integreerbaar zijn op (elektronische of fotonische) chips is vooral te danken aan de recente vooruitgang in zowel de miniaturisatie van optische componenten en detectieprincipes alsook de bereikte resultaten op het vlak van microfluidica. Deze geïntegreerde systemen laten toe om sneller en goedkoper verschillende biologische molecules te detecteren dan tegenwoordig mogelijk is met dure en werkintensieve laboratorium toestellen. Het uiteindelijke einddoel is een mobiel toestel waarmee men besmettelijke of andere ziektes kan detecteren, dit wordt aangeduid als *aan-het-bed analyse* (point-of-care, POC). Noodzakelijkerwijze moeten deze systemen goedkoop zijn, maar ook accuraat, betrouwbaar, gebruiksvriendelijk en in staat om te functioneren in extreme omstandigheden. Bovendien moet dezelfde detectielimiet, specificiteit en reproduceerbaarheid gehaald worden als de huidige systemen.

Het miniaturiseren van optische componenten en het integreren van die componenten op kleine chips wordt fotonische integratie genoemd. Deze componenten kunnen fotonische, elektronische en zelf fluïdische functionaliteit bezitten. Optische sensoren werken over het algemeen volgens het evanescente golf principe, wat betekent dat de propagatieconstante van optische golfgeleidermodes sterk afhankelijk is van de brekingsindex van de omringende lagen. Naast de extreme gevoeligheid die met dit principe kan gehaald worden kan de interesse in optische sensoren nog aan een aantal andere factoren toegeschreven worden. Zo zijn ze elektrisch passief and dus ongevoelig voor elektromagnetische interferentie, waardoor ze gebruikt kunnen worden in situaties waarbij elektronische sensoren niet kunnen functioneren. Maar misschien het grootste voordeel is het potentieel om de techniek te paralleliseren: door verschillende golflengtes of polarizatietoestanden te gebruiken zouden in principe verschillende biologische markers simultaan gedetecteerd kunnen worden.

Misschien wel het meest conventionele en toegepaste optische detectieprincipe is het zogenaamde *oppervlakte plasmon resonantie* mechanisme

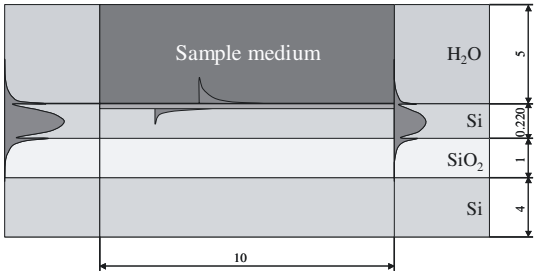
(surface plasmon resonance, SPR). De enorme gevoeligheid van deze evanescente golven voor heel erg kleine brekingsindexvariaties aan het oppervlak van een metaallaag is de hoofreden dat deze techniek vaak gebruikt wordt voor reële-tijd en label-vrije biosensoren[1]. Gemotiveerd door de visie om een lab-op-chip (LOC) component te realiseren, is er dan ook enorm veel onderzoek gebeurd om dit principe te miniaturiseren en te integreren. Verschillende geïntegreerde oppervlakte plasmon sensoren werden dan ook gedemonstreerd in de laatste tien jaar, meestal zijn die oplossingen gebaseerd op een dunne goudlaag gedeponeerd op een optische golfgeleider. Alle systemen die tot op de dag van vandaag geïntroduceerd werden, zijn echter gefabriceerd in een lage-index contrast materiaalsysteem, waardoor de typische dimensies over het algemeen te groot blijven om in aanmerking te komen voor parallellisatie en lab-op-chip toepassingen.

Een van de hoofddoelstellingen van dit werk was om het potentieel van oppervlakte plasmon modes om brekingsindexverschillen te detecteren te combineren met de unieke fotonische integratie mogelijkheden van een hoog-index contrast materiaalsysteem zoals silicium-op-isolator (SOI). Op die manier zou een gevoelige, kleine en goedkope biosensor gerealiseerd kunnen worden.

Detectie principe

Het hoge index contrast is enerzijds heel erg voordelig op het vlak van integratie, maar anderzijds sluit het ook het principe uit waarop alle tot op de dag van vandaag geïntroduceerde geïntegreerde oppervlakte plasmon sensoren steunen. Het principe is namelijk gebaseerd op fase-koppeling, als de propagatieconstante van een diëlektrische golfgeleidermode gelijk wordt aan de propagatieconstante van een oppervlakte plasmon mode koppelt optische vermogen van de golfgeleidermode naar de plasmonmode. Omdat de plasmonmode hoge verliezen vertoont leidt dit tot een grote absorptie van optisch vermogen voor een welbepaalde golflengte en brekingsindex. De diëlektrische golfgeleidermode propageert voornamelijk in het hoge-index materiaal silicium, het oppervlakte plasmon waarvan sprake propageert in een waterige oplossing waarvan de brekingsindex relatief laag is. De propagatieconstantes verschillen dan ook te veel om koppeling tussen deze

[1] Label-vrij betekent dat de biomolecules gedetecteerd worden zonder de hulp van een label. Dit label kan bijvoorbeeld een fluorescente molecule zijn, of een radio-actieve component. Label-vrije biosensoren detecteren biomolecules die binden aan specifiek gefunctionaliseerde oppervlakken door de verandering van brekingsindex van deze laag te kwantificeren.



Figuur 1: Schematische weergave van een oppervlakte plasmon interferometer, alle afmetingen in μm.

modes te verwezenlijken. Een nieuw detectieprincipe dringt zich dus op. Het concept dat bestudeerd werd in dit werk is een oppervlakte plasmon interferometer gebaseerd op een dunne goud laag, geïntegreerd in een silicium golfgeleider. Figuur 1 toont schematisch hoe dit systeem werkt. Aan de linkerkant wordt de interferometer geëxciteerd met de transversaal magnetische nulde orde mode van de silicium golfgeleider. Als deze mode de goud laag bereikt worden daar twee verschillende oppervlakte plasmon modes geëxciteerd, één aan de bovenkant en één aan de onderkant van de goudlaag. Deze modes zijn niet gekoppeld vanwege het hoge brekingsindex verschil tussen onderkant en bovenkant van de goudlaag. Hun fasesnelheid wordt volledig bepaald door de brekingsindex van de laag waardoor ze propageren. Op het einde van de goudlaag koppelen de beide oppervlakte plasmon modes terug naar de nulde orde mode van de silicium golfgeleider. In deze golfgeleider interfereren beide modes constructief of destructief, afhankelijk van het relatieve faseverschil tussen beide. Dit principe verklaart hoe deze structuur als een sensor kan gebruikt worden, door middel van interferentie wordt in de uitgangsgolfgeleider een faseverschil omgezet in een (meetbaar) intensiteitsverschil.

Het hoofddoel van dit werk is om deze structuur volledig te karakteriseren en de belangrijke factoren en ontwerpsparameters te beschrijven die van deze structuur een gevoelige sensor kunnen maken. De karakteristieken van all optische modes, de meest belangrijke uiteraard de gevoeligheid voor brekingsindexvariaties, werden onderzocht. Gepaste uitdrukkingen voor zowel bulk als oppervlakte gevoeligheid werden opgesteld, waarmee dan expliciet het potentieel om als kleine brekingsindexverschillen te detecteren berekend kan worden. Het effect van de dikte van de silicium golfgeleider, de dikte van de goudlaag lengte van de volledige structuur op de gevoeligheid en de karakteristieke transmissie werden onderzocht. Gebaseerd

op deze bevindingen werd een algemene set ontwerpsregels opgesteld om gevoelige oppervlakte plasmon interferometers in silicium-op-isolator te ontwerpen.

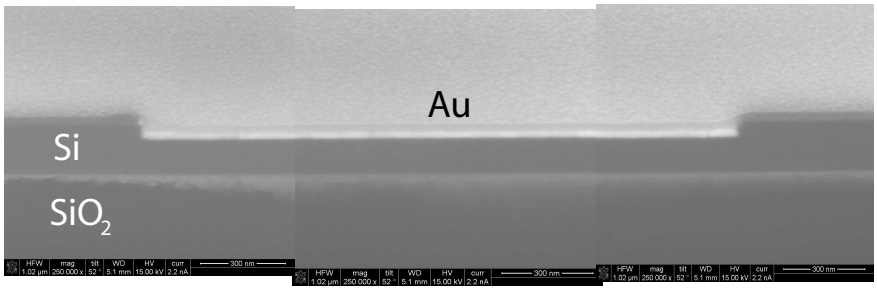
Simulatie methodes

Een van de oplossingsmethodes van de vectoriële eigenmode expansie software die in dit werkt gebruikt werd (CAMFR), is gebaseerd op vlakke golf expansie, ook wel bekend als Fourier exansie of de Fourier modale methode. Een van de nadelen van deze methode is dat ze heel traag convergeert voor geometrie¨en die dunne metaallagen bevatten, vooral voor transversaal magnetisch gepolariseerde elektromagnetische golven. Opdat we met deze methode oppervlakte plasmonen zouden kunnen berekenen werd de vlakke golf expansie uitgebreid met de adaptieve ruimtelijke resolutie methode (adaptive spatial resolution, ASR), gecombineerd met geavanceerde randvoorwaarden zoals perfect aangepaste lagen (perfectly matched layers, PML). Voor het eendimensionale geval leverde dit een gevoelig verbeterde convergentie op, het tweedimensionale geval werd geplaagd door additionele problemen waardoor de toepasbaarheid van het adaptieve resolutie algoritme beperkt bleef.

Fabricage

Het fabriceren van een structuur zoals beschreven in dit werk vereist verschillende fabricagestappen, die allemaal dienen geoptimaliseerd te worden. Om mooie, vlakke silicium oppervlakken te verkrijgen, in combinatie met scherpe overgangen aan het begin en einde van de goudlaag werd het RIE-ICP droge etsproces dan ook grondig bestudeerd. Aanhoudende problemen met de ICP-RIE machine in onze cleanroom, noopten ons ertoe om andere routes te bewandelen om de noodzakelijke oppervlaktekwaliteit te halen. Het effect van de ruwheid van het silicium oppervlak op de topologie van de goudlaag werd kwalitatief bestudeerd, waarbij aangetoond werd dat de uiteindelijke kwaliteit van de goudlaag heel sterk afhankelijk is van de kwaliteit van het geëtste silicium oppervlak.

Alhoewel de eigenschappen van golfgeleider modes van golfgeleiders in een symmetrische omgeving goed bekend zijn, is over asymmetrische omgevingen relatief weinig bekend. Sterke asymmetrie tussen de brekingsindex van de boven-en onderlaag van een golfgeleider leidt tot hinderlijke TE-TM anti-kruisingen en geassocieerde mini-stopbanden. Als gevolg functioneren



Figuur 2: FIB doorsnde van een gafabriceerde structuur. De structuur hier weergegeven werd ook experimenteel uitgemeten, resultaten zijn te vinden in hoofdstuk 7 (eerste meetsessie bulk experimenten).

adiabatische golfgeleider koppelaars matig tot zeer slecht. Dit effect, dat niet bekend was ten tijde van de eerste fabricage stappen, is een serieuze uitdaging wat fabricage betreft.

Roosterkoppelaars specifiek ontworpen voor TM polarisatie kunnen optisch vermogen van een monomode optische vezel naar een golfgeleidermode koppelen. Alhoewel deze koppelaars in theorie even efficiënt zijn als de reeds bestaande TE koppelaars presteerden ze experimenteel toch ondermaats. De effici¨entie en bandbreedte zijn aanzienlijk lager dan theoretisch voorspeld. Uiteindelijk werd dan ook geopteerd om horizontaal in te koppelen.

Data Analyse

De toegepaste data analyse van ruwe data bepaalt in hoge mate de betrouwbaarheid en gevoeligheid van een sensor. De detectielimiet wordt namelijk bepaald door de resolutie waarmee de positie van het spectrale minimum bepaald kan worden. Verschillende data-analyse methodes zoals (gewogen) geometrisch middelpuntsmethodes, polynomiale fit methodes en een Lorentziaanse fit methode werden geëvalueerd. Hun afhankelijkheid ten opzichte van transmissie grenswaarde en ruis werden bestudeerd. Uiteindelijk kon op die manier een experimentele ondergrens voor de detectielimiet bepaald worden.

Experimentele Resultaten

De gevoeligheid van de oppervlakte plasmon interferometer voor bulk brekingsindex variaties werd aangetoond en gekwantificeerd. Alhoewel de ontwerpsparameters verre van ideaal waren (door beperkingen wat fabricage betreft) werd een verschuiving van het spectrale minimum opgemeten van om en bij de $450\ nm/RIU$. De opgemeten spectra vertoonden echter allemaal heel erg hoge verliezen, de signaal-tot-ruis verhouding was dan ook nauwelijks $15\ dB$. Pogingen tot het opmeten van de gevoeligheid voor binding van biomolecules aan het gefunctionaliseerde oppervlak door middel van een biotine-avidine bindingsexperiment leverden helaas geen bruikbare resultaten op.

English summary

INTEREST in scaled-down sensing processes, combined with advances in microfluidics, is motivating various chip-based methods in which sensing can be carried out more rapidly and at lower cost via small-scale systems than with current laboratory bench-scale methods. This so called point-of-care (POC) analysis will lead to the possibility of self-contained handheld assays for detection of infectious or other diseases. These systems must be inexpensive, but also accurate, reliable and rugged while maintaining or even surpassing the sensitivity, specificity and reproducibility levels similar to those of central laboratory analyzers, but yet require very little experience or training from the end-user.

Photonic integration is the process of miniaturizing optical components and integrate them onto small chips containing photonic, electronic and possibly fluidic functionality. Their potential for integration into a Lab-on-chip device is due to the fact that optical waveguide modes can be made very sensitive to refractive index changes, a principle known as evanescent wave sensing. The success of optical sensors is due to a number of distinct advantages that optical transduction mechanisms seem to offer over other mechanisms. They are electrically passive and thus free of any electrical interference. They can be used in high-voltage applications and harsh environments where electronic circuitry would not survive. Moreover, optical sensors lend themselves easily to multiplexing capabilities by using different wavelengths or polarization states of light.

Perhaps the most established and widely used optical detection principle is the surface plasmon resonance mechanism (SPR). The high sensitivity of this technique to surface phenomena makes it ideal for use in real-time and label-free biosensors where very small changes in refractive index must be detected. Driven by the vision of a lab-on-chip, extensive work has been done to miniaturize and integrate SPR biosensors. In the past decade, several integrated optical SPR sensors have been demonstrated [1, 2, 3] in which thin gold films serving as a platform for the attachment of sensing films are deposited on top of an integrated optical waveguide

system. However, all integrated SPR sensors that have been investigated so far are fabricated in a material system with a low refractive index contrast, keeping typical dimensions of waveguides and optical components too large for miniaturization and consequent lab on chip applications.

One of the main objectives of this work was to combine the proven potential of surface plasmon modes with the unique photonic integration capabilities of a high index contrast material system, such as Silicon-on-insulator (SOI). This combines the sensitivity of SPR with the high level of integration and low-fabrication cost of SOI devices.

Sensing Principle

While the high index contrast is advantageous for integration purposes, it also inherently includes the restriction that the integrated surface plasmon principle used in prior art, does no longer function. For this principle to work, phase matching between a waveguide mode and a surface plasmon mode is required. The waveguide mode is generally confined to a high refractive index medium ($n_{Si} \approx 3.45$), the surface plasmon mode will propagate in an aqueous solution with a low refractive index ($n_{H_2O} \approx 1.3105$). Phase matching these two modes has proven to be unfeasible. So, a novel detection principle had to be envisioned. The result of this work was the conception of a surface plasmon interferometer, consisting of a thin layer of gold embedded in a silicon wire waveguide. The working principle of this interferometer is schematically depicted in figure 3. Upon reaching the gold-clad wire waveguide layer, a dielectric TM-polarized mode guided by the silicon waveguide excites two surface plasmon modes, one at the upper and one at the lower interface of the gold layer. Due to the highly asymmetric cladding layers these modes are not coupled. Therefore, their phase velocities are entirely determined by the refractive index of the upper and lower dielectric. At the end of this section, interference of the two surface plasmon modes results in a dielectric mode launched in the output waveguide. This explains the sensing functionality of the interferometer: a change in the refractive index of the medium above the gold layer results in a phase difference between the two surface plasmon modes and consequently, in a change of output intensity.

To fully develop the potential of this device is the main focus of this work. To this purpose, a lot of effort has been spent to truly understand the different physical mechanisms that determine the sensing potential of a surface plasmon interferometer. The characteristics of all optical modes

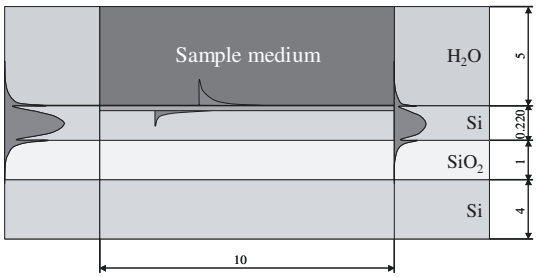


Figure 3: Schematical setup of a surface plasmon interferometer, all dimensions in μm

in the structure, the most important being their dependence on either bulk refractive index changes and/or adlayers that represent the adsorption of biological molecules, were determined. Appropriate expressions for the bulk and surface sensitivity were formulated, with these expression the potential for sensing can be calculated. The effect of cladding layer thickness, gold thickness and sensor device length on the sensitivity and the overall transmission has been described and based on these findings a set of general design rules for surface plasmon interferometers in SOI were derived.

Simulation methods

The in-house developed fully vectorial mode solving software (CAMFR) used in this work, has among its mode solvers one based on plane-wave expansion, also known as Fourier expansion or the Fourier modal method (FMM), which, especially for TM polarization of metallic lamellar gratings, is plagued by slow convergence. In order to simulate surface plasmon waveguide structures the FMM method was improved by incorporating the adaptive spatial resolution technique and combining this technique with perfectly matching layers (PML) boundary conditions. In the one dimensional case, this resulted in substantial improvement in convergence rate, the two dimensional case (for the simulation of waveguide cross-sections) suffered from additional problems, limiting the applicability of the adaptive spatial resolution algorithm.

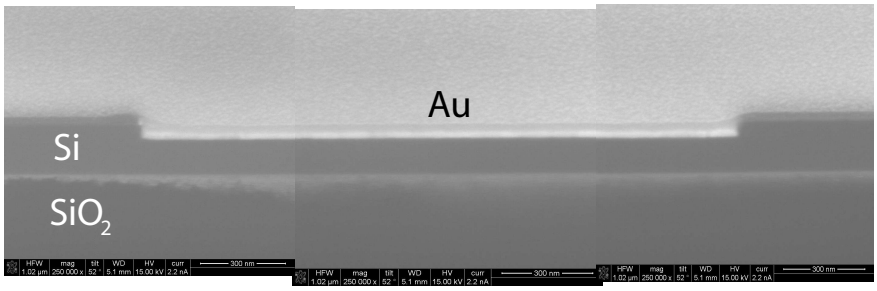


Figure 4: FIB cross-section of a fabricated sample (the sample depicted here is the sample that was measured in measurement run 1, see chapter 7 for experimental results).

Fabrication

Several fabrication runs, each iteration featuring incremental improvements over the previous run were carried out. The RIE-ICP dry etching process parameters were tweaked in order to obtain low-roughness Si surfaces, with sharp discontinuities at the begin and end of the Au region. Due to the fact that the RIE-ICP machine in our clean room is not a dedicated device for Si etching, eventually other routes had to be explored in order to obtain the required surface quality. The effect of Si surface roughness on the resulting Au layer after deposition was qualitatively studied, again proving the importance of obtaining nice low-roughness surfaces prior to gold deposition.

Although the modal characteristics of waveguide modes in wire waveguides are well understood for a symmetrical cladding environment [2], the situation proved to be somewhat different for asymmetrical cladding layers. Strong asymmetry results in considerable TE-TM anti crossings and associated mini-stopbands, limiting the performance of adiabatic couplers. These effects, unknown at the time, are severe challenges for the fabrications.

In and outcoupling of optical power from an single mode optical fiber can be done using grating couplers that were specifically optimized for TM-polarization. These couplers, although potentially as efficient as TE grating couplers, a trusted and reliable technology, proved to be somewhat more unreliable in experimental situations. The efficiency and 3 dB bandwidth was considerably lower than theoretical estimates. The horizontal incoupling

[2] The denomination symmetric/antisymmetric refers to the refractive indices of the materials above and below the metal strip, 'symmetric' meaning that both refractive indices are the same, 'asymmetric' pointing to a difference in these indices.

technique of end-fire coupling was eventually used.

Data Analysis

A sensor is only as reliable and sensitive as its data-analysis methods quality, since the resolution with which the position of the spectral minimum can be determined, determines the detection limit. Various data-analysis schemes such as (weighted) centroid algorithms, polynomial curve fitting and Lorentzian curve fitting were studied. Their dependence on threshold baseline and Gaussian noise were evaluated, leading to an experimentally obtainable detection limit.

Sensing experiments

Proof-of-principle refractometric sensing was obtained. Although not optimized (due to fabrication restrictions) in the sense of having a very large bulk sensitivity S_{bulk} and/or surface sensitivity $S_{surface}$, the wavelength shift as a function of bulk refractive index can be as high as $\approx 450\ nm/RIU$. Obtained waveguide spectra showed high overall losses, and a signal-to-noise ratio of only $\approx 15dB$. The surface sensitivity $S_{surface}$ for a biotin-avidin binding experiment could not be determined, this remains an open question.

References

[1] J. Čtyrocký, J Homola, P.V. Lambeck, S. Musa, H.J.W.M. Hoekstra, R.D. Harris, J.S. Wilkinson, and N.M. Usievich, B.and Lyndin. *Theory and Modelling of Optical Waveguide Sensors Utilising Surface Plasmon Resonance*. Sensors and Actuators B -Chemical, 54:66–73, 1999.

[2] R. D. Harris, B. J. Luff, J. S. Wilkinson, J. Piehler, A. Brecht, G. Gauglitz, and R. A. Abuknesha. *Integrated optical surface plasmon resonance immunoprobe for simazine detection*. Biosensors & Bioelectronics, 14(4):377–386, 1999.

[3] J Dostálek, J Čtyrocký, J Homola, E Brynda, M Skalský, P Nekvindová, J Špiriková, J Škvor, and J Schröfel. *Surface plasmon resonance biosensor based on integrated optical waveguide*. Sensors and Actuators B -Chemical, 76:8–12, 2001.

1 Introduction

THE scope of this work is to study, design, fabricate and characterize a label-free biosensor, using sensitive and quantitative optical detection methods based on the integration of surface plasmon waveguides in the Silicon-on-Insulator platform. This combines the maturity and sensitivity of surface plasmon detection principles with the tremendous promise of photonic integration of the SOI material system. This could possibly pave the way for miniaturization, leading to very compact diagnostic chips with nano-array configurations, such that multi parameter analyses of the same biological sample or even a single cell become possible [1].

1.1 Context

1.1.1 Why?

Substantial progress has been made in public health in recent years, but this progress has not been equal in developed and developing countries. The benefits and services that we take for granted are often lacking, not present or inadequate in many developing countries [2, 3]. The developing world does not have access to many of the best medical diagnostic technologies, simply because these technologies were designed for air-conditioned laboratories, refrigerated storage of chemicals, a constant supply of calibrators and reagents, stable electrical power, highly trained personnel and rapid

transportation of samples.

For every public health triumph such as the eradication of smallpox, other infectious diseases such as tuberculosis and malaria have re-emerged, accompanied by new diseases (see figure 1.1). Infectious diseases are responsible for more than 40% of all deaths worldwide. Breaking this number down between developed and development countries reveals that in the poorest countries more than half the deaths are due to infectious diseases, whereas in the richest countries this number drops below 5% [4, 5]. In some cases, poor healthcare clinics have the drugs to treat a certain illness but just lack the diagnostic tools and required expertise to identify patients who should receive the drugs, or they fail to pinpoint the agent or pathogen responsible for the illness. Global health, poverty and development are interdependent. Endemic poverty is a significant impediment to improving health [6].

Pathogens can be in food, in water or airborne, and early detection of these pathogens is extremely important to both minimize the spreading of infectious diseases and rapidly diagnose infections for better, faster and often less expensive treatment. For example food borne illnesses are often caused by bacterial pathogens (91%), e.g. Escherichia coli[1], Salmonella [2], Staphylococus aureus[3], Campylobacter coli[4] and Bacillus Cereus[5] can easily contaminate food and water. These bacterial pathogens are quite dangerous and can even cause death, even a very low amount can lead to infection (10 E. coli or Salmonela bacteria are enough to make one sick). Moreover,

[1] E. coli is a Gram negative bacterium that is commonly found in the lower intestine of warm-blooded organisms (endotherms). Most E. coli strains are harmless, but some virulent strains , can cause can cause gastroenteritis, urinary tract infections, and neonatal meningitis, and are occasionally responsible for costly product recalls [7].

[2] Salmonella is a group of bacteria that can cause diarrheal illness in humans. Infected persons develop diarrhea, fever, and abdominal cramps 12 to 72 hours after infection. Though most subjects recover without treatment, in some patients, the Salmonella infection may spread from the intestines to the blood stream, and then to other body sites and can cause death [7]

[3] Staphylococcus aureus is a common bacterium found on the skin and in the noses of up to 25% of healthy people and animals. It has the ability to make seven different toxins that are frequently responsible for gastrointestinal illnesses [7].

[4] Campylobacters are mainly spiral-shaped, S-shaped or curved, rod-shaped bacteria. They are a major cause of diarrhoeal illness in humans and are generally regarded as the most common bacterial cause of gastroenteritis worldwide. In almost all developed countries, the incidence of human campylobacter infections has been steadily increasing for several years. The reasons for this are unknown [8].

[5] Bacillus Cereus is a spore forming bacteria which causes an increasing number of cases of food poisoning mainly due to rice from take-away food premises (the so-called *fried rice syndrome*) [7].

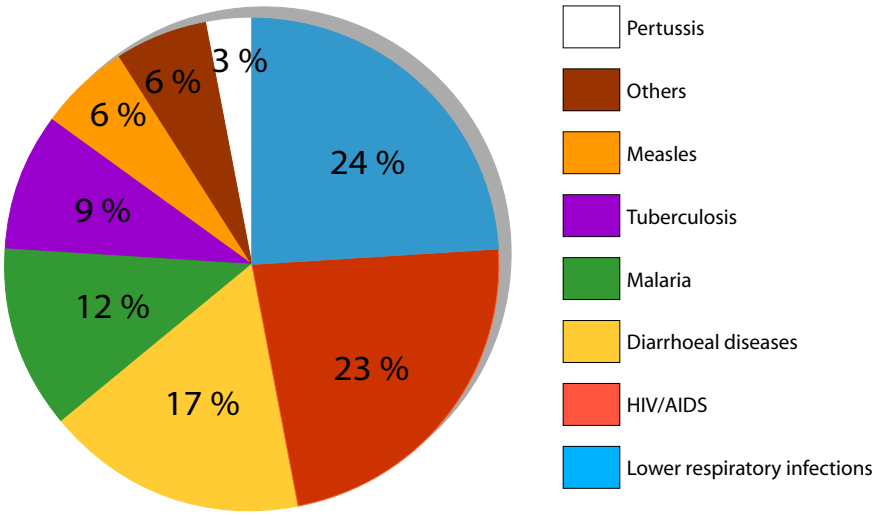


Figure 1.1: Disability-adjusted life years (DALYs) for infectious and parasitic diseases. To properly reflect the full impact of a disease, disease burdens can be measured in DALYs by adding the years of life lost by a persons premature death to the time lived with a disability. Infectious and parasitic diseases accounted for almost 30% of all DALYs and 15 million deaths each year worldwide. Shown are the infectious and parasitic diseases responsible for the DALYs in 2005 (figure taken from [9])

the occurrence of infections due to bacterial pathogens does not seem to decrease.

For optimal use in low-resource settings, it is important for tests to be rapid, simple to use (that is, requiring little in the way of facilities, equipment or training), low-cost or cost-effective, easily interpretable, and stable when transported and stored under extreme conditions. Improved point-of-care (POC) methods are needed to return same-day test results so that patients can receive appropriate therapy while they are still at the clinic. Multiplexed tests are especially needed to accurately identify the agent causing a disease that could have multiple causative agents, such as acute lower respiratory infections, diarrheal diseases, acute onset fevers and sexually transmitted infections [9].

Other, more innocent, bacterial infections might not lead to dangerous situations, they however constitute a large health care cost since simple infections like tonsillitis (inflammation of the throat) and inflammation of the urinary system lead to more than 25 million calls at medical doctors each year in the USA alone. Maintaining control of the rising health care costs in the United States has had a serious impact on diagnostic testing. Reim-

bursement for the cost of laboratory tests has resulted in lower investments in highly trained medical technologists and the purchase of less expensive methodologies. Diagnostic tests are sought by medical providers that are simple to perform, are at least as sensitive and specific as conventional assays, and are inexpensive [10].

Interest in moving to a more patient-centric POC/home-testing approach is on the rise in the developed world [11, 12]. Lab-on-a-chip (LOC) technology suits both developing- and developed-world applications. Many researchers believe we will soon have microfluidics-based POC and home-care devices that can perform assays at sensitivity, specificity and reproducibility levels similar to those of central laboratory analyzers, but yet require little user input other than the insertion of the sample.

1.1.2 Lab-on-Chip and Biosensing

A lab-on-a-chip is a device that integrates one or several laboratory functions on a single chip of only millimeters to a few square centimeters in size [13]. The laboratory functions can be very diverse in nature, like cell lysis, sample amplification (e.g. PCR) and dilution, sample cleaning, fluorescent labeling, process monitoring and optical detection. These are all processes that are currently being carried out in a laboratory environment, which is slow, suffers from high fluid volume consumption (of water, reagents and sample volumes) and is rather costly, both in equipment cost as in the use of highly-trained personnel. Labs-on-chip, on the contrary, would have a fast response time, allow for massive parallelization due to compactness and would need only very low fluid volumes. Furthermore, and this is the real advantage for developing countries, the technology to produce labs-on-a-chip can be massively parallel so that low-cost fabrication is a real and viable option, up to the point where one can envision these chips to be disposables.

This work does not have the ambition nor the scope of solving all the lab-on-chip issues, rather, it focuses on a very small (but nevertheless important) part of a LOC device, optical detection, also know as biological sensing. The term biosensor covers a wide range of devices that measure the presence or concentration of biological molecules, biological structures, micro-organisms, etc., by translating a biochemical interaction at the probe surface into a quantifiable physical signal.

Although the preceding section mainly focused on the detection of bacterial pathogens, the possible application field is quite extensive, involving such areas as the medical industry for drug development, diagnostics at the local doctor and home medical tests (point-of-care diagnostics). Online screening of water quality for both household and industrial use, environmen-

tal monitoring of both air and water quality, conditions in storage facilities and other indoor facilities, . . .

Better sensors to detect biomolecules in a multi parameter fashion are obviously crucial for the further development of biotechnology and the total analysis of a biological system. There are three issues that require nanotechnology solutions: large-scale multi parameter analysis, high sensitivity, and the ability to get real time quantitative results. New nanotech-based sensors should have maximum sensitivity to allow measurements of biological samples or at a cell base with a large dynamic range and a high accuracy, high throughput, and in an array setup for multi parameter analysis.

Nowadays, typical detection systems use antibodies that are labeled using radioactivity, enzymes or fluorescence. When such a labeled molecule binds to a surface, its presence can be detected using straightforward luminometry, fluorimetry, spectrometry, radiometry or confocal microscopy. While this approach is extremely sensitive, with the detection limit down to a single molecule [14], it is far from ideal: fluorescent or radio labeling can involve expensive and or hazardous labeling procedures, moreover it is a laborious, time-consuming operation [15]. Furthermore, labeling may change the reactivity and/or specificity of the labeled biomolecules, reducing both the qualitative (detectability, specificity and selectivity) and quantitative (kinetics and thermodynamic parameters, concentration analysis) information of biological assays. Due to the fact that the number of fluorophores on each molecule cannot be precisely defined, the method also suffers from signal bias.

It would be much more elegant to detect the presence of biomolecules directly, i.e. without an intermediate labeling step. This way target molecules are detected in their natural form. This type of detection is relatively easy and cheap to perform, and allows for quantitative and kinetic measurement of molecular interaction. Additionally, since label-free detection schemes measure refractive indices, the detection signal does not scale down with sample volume. This is particularly attractive when ultra small detection volumes are involved.

The fundamental principle underlying label-free detection method is in theory very simple (Fig. 1.2). A biomolecule binds to a chemically adapted surface containing detector groups (e.g. mRNA binds a complementary DNA strand or an antigen its corresponding antibody). Since the target analytes have a different refractive index than the buffer solution (e.g. the refractive index of a protein is 1.5 vs, 1.33 for buffer solution), this results in a refractive index change near the surface, which can be detected optically as the sensing transduction signal. Most label-free sensors use light concentrated near the surface with an evanescent field exponentially decaying in

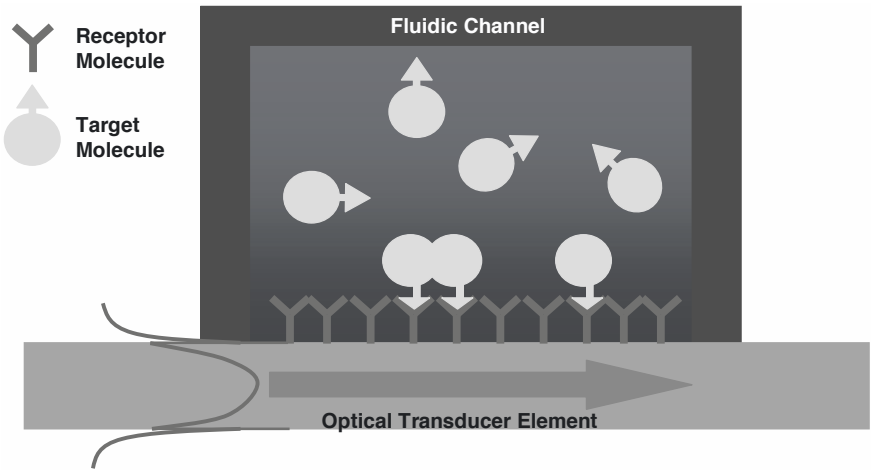


Figure 1.2: Schematical depiction of label-free biosensing principles.

the bulk solution with a characteristic decay length a few tens up to a few hundreds of nanometers. Thus, those analytes that are not captured by the biorecognition molecules will not have an influence on the sensor signal, providing a means to distinguish between target molecules that bind to the surface and molecules that are present in the solution but do not bind to the surface [16].

1.1.3 Surface Plasmon Biosensors (SPR)

Today, biosensors using the excitation of the surface plasmons are generally termed surface plasmon resonance (SPR). They provide rich information on the specificity, affinity, and kinetics of biomolecular interactions and/or the concentration levels of an analyte of interest from a complex sample [15]. The analysis is done in real-time and without requiring fluorophore labeling. Numerous commercial systems are currently available; however, their designs do not differ significantly from the original concept described by Liedberg [17] to demonstrate SPR biosensing.

In SPR biosensing, the adsorption of a targeted analyte by a surface bioreceptor is measured by tracking the change in the conditions of the resonance coupling of incident light to a surface plasmon wave. The existence of this surface plasmon wave is dictated by the electromagnetic properties of a metal, typically gold, and a dielectric interface (see chapter 2). The resonant coupling appears as a dip in the reflectivity of the light spectrum, which is traditionally tracked by measuring the wavelength, the incident angle or the intensity of the reflected light (Fig. 1.3). Free space coupling of light to the

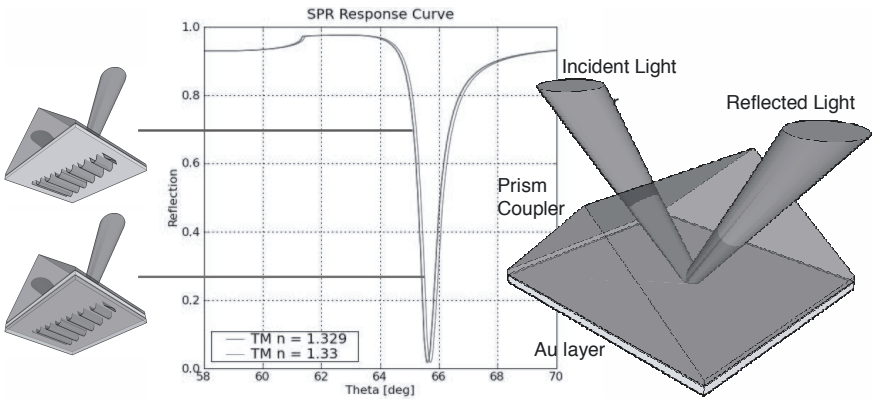


Figure 1.3: Tracking the surface absorption by SPR in the prism coupled configuration (Kretchmann configuration). Simulation of the resonance position as a function of reflectance angle for two distinct refractive indices of the dielectric at the gold interface.

surface plasmon requires, for electromagnetic reasons, a high-index prism or a periodic grating surface [18, 19].

The ability of SPR to detect low concentrations of analyte lies in the strong electromagnetic enhancement of the surface plasmon. Commercial SPR biosensors are generally capable of detecting $1pg/mm^2$ of absorbed analytes. This detection limit is strongly dependent on many parameters, but is particularly dependent on surface functionalization. In comparing detection limits between reported SPR biosensors, one must be cautious as the values are often described independently of the surface functionalization chemistry or for a specific application. For a more relevant assessment, the detection limit, where it is available, will be expressed in term of detectable refractive index unit (RIU) change. This value strictly reflects the performance of the optical configuration, the measurement approach and the data analysis algorithm. Commercial systems are reported with detection limits typically within 1×10^{-5} to 1×10^{-6} RIU. Today, the key challenge in the SPR biosensor development lies not primarily in the integration of the various components of the biosensor (sampling handling, control electronics, etc.) but on providing robust integrated SPR biosensors that are as or more sensitive ($< 10^{-6} RIU$) than their current counterparts [20]. Detection limits of the different surface plasmon resonance based biosensors are listed in table 1.1 , the Sentata Spreeta approach [21] is the classical surface plasmin resonance biosensor based on the Kretschmann configuration.

Due to the enormous integration potential of DUV-lithography the fabrication cost of an integrated surface plasmon biosensor can be very low,

SPR Approach	DL(RIU)	Reference
Fiber and waveguide SPR		
-Micro-prism	8×10^{-4}	[22]
-Polarization-maintaining fiber	4×10^{-6}	[23]
-Multi-mode waveguide	4×10^{-3}	[24]
-DBR-waveguide	1.28×10^{-5}	[25]
-Dual-LED	5.2×10^{-4}	[26]
-Y-splitter and SU8 microfluidics	6×10^{-4}	[27]
Silicon SPR		
-SPR-Schottky	12×10^{-5}	[28]
Imaging SPR		
-Polarization contrast	2×10^{-6}	[29]
Sentata Spreeta	1×10^{-6}	[21]

Table 1.1: Detection limit comparison of the different SPR approaches, table taken from [20].

rough estimates for the different fabrication steps are given in table 1.2. These estimates are typical for CMOS research laboratories (such as IMEC) and could be further reduced in a CMOS production line.

Low cost of production, a compact design, reusability and increased functionality (multiple analyte detection, temperature control, etc.) are as well sought after, besides the requirement for high sensitivity.

1.2 Integrated Surface Plasmon Biosensors: State-of-the-Art

Numerous examples of novel surface plasmon biosensor designs that improve upon the traditional and popular prism-coupled SPR can be found in the current literature.

Todays commercial systems or bench-top SPR biosensors are generally based on the traditional prism-coupled SPR configuration (Kretchmann's configuration, the basic transduction method of this approach is schematically depicted in Fig. 1.3). The approach is simple, sensitive and robust; however, it is not amenable to miniaturization and integration due to the fact that this is not an in-plane technology, requiring high alignment accuracy of the comprising optical components. Sensing area is ultimately limited by the diffraction-spot size of the excitation wavelength and the propagation length of the surface plasmon on the gold surface. The progress made over the past

Wafer cost	300
DUV lithography mask set	25000
Deep Etch	
- Lithography	1000 / lot
- Etching	1000/lot
- Resist Strip	1000/lot
Shallow Etch	
- Lithography	1000 / lot
- Etching	1000/lot
- Resist Strip	1000/lot
Dicing	100/wafer
Number of chips per wafer ($10mm^2$)	12500
Number of wafers per lot	23
Estimated price per chip	0.402

Table 1.2: Estimate of the price per chip in a CMOS research lab environment, all prices in euro.

20 years has been described extensively in review papers by Homola [30, 31]. Recently, much of the development of SPR is directed toward providing an integrated, low-cost, reusable and sensitive biosensor. Several approaches are possible to improve and integrate old and novel technological approaches to SPR.

A first attempt to integrate SPR were optical fibers. Early fiber-SPR biosensors use single mode fiber with polished and metal-coated tip. The metal is functionalized for the detection of any gases or biomolecules [32]. Changes in the surface plasmon resonance are measured by either collecting the back reflected light from the fiber or the diffracted light from the polished end. Recently, to improve the coupling of light to the surface plasmon mode, a micro-cone was etched in the tip of the fiber using micro-machining technologies forming a micro-prism [22]. Here a sensitivity of 8×10^{-4} RIU was reported. It was also discovered that moving from the optical spectral regime to the infrared spectral regime could improve the sensitivity with an order of magnitude [33]. Other recent improvements focus on using polarization-maintaining fiber for more stable and sensitive measurements [23].

Despite these recent improvements, the integration of fiber-SPR to other biosensor components raises some design issues. The insertion of the fiber into a planar substrate containing a fluidic circuit and chambers requires precise alignments and complex assembly.

The proposed alternative approach to fiber-based SPR has been the planar optical waveguide structure. Thin-film technologies have long been used to develop planar optical waveguides consisting of a multilayer structure of varying optical refractive index that can support a propagating light mode. Over 20 years ago, Kreuwel demonstrated that these structures be readily modified with a bio-functionalized metal layer for the excitation and propagation of a surface plasmon wave [34]. Variations in the transmission power of the guided light are monitored as it is coupled upon phase-matching to the surface plasmon propagating on the external gold surface. Since then, extensive theoretical and experimental work has also been done in this area, leading to the demonstration of several integrated optical SPR sensors [35, 36, 37], in which thin gold films serving as a platform for the attachment of sensing films are deposited on top of an integrated optical waveguide system. However, all integrated SPR sensors that have been investigated so far are fabricated in a material system with a low refractive index contrast [35, 36, 37, 38], keeping typical dimensions of waveguides and optical components too large for miniaturization and consequent lab on chip applications.

Waveguide-SPR biosensors are particularly well suited for large scale and planar integration. Complex optical functions can be incorporated into a single substrate. For example, an integrated waveguide Mach Zehnder interferometer SPR device has been reported [39]. The resonance coupling of light to the SP induces a significant phase change to the incident light, which can then be measured by a heterodyne or interferometry approach [27]. The reported integrated device consists of a waveguide channel that is divided into two arms, a sensing arm and a reference arm, by means of Y-splitter. Both arms are recombined to produce a phase-dependent interference between the two guided modes. The device also integrates an SU-8 microfluidic system[6] and preliminary results show a detection limit of $6 \times 10^{-4}\ RIU$. In this design, we see that the surface plasmon excitation is translated into a quantifiable output signal by means of interferometric

[6]SU-8 is a commonly used epoxy-based negative photoresist. It is a very viscous polymer that can be spun or spread over a thickness ranging from $0.1\ \mu m$ up to $2\ mm$ and still be processed with standard contact lithography. SU-8 has been shown to be biocompatible. It is chemically stable and no background has been noticed from the material itself in analytical applications and it has been shown to be compatible with most chemicals applied in analytical applications. The transparency of the material enables optical detection as well. However, some fluorescence signal is emitted from SU-8 itself. Drawbacks of the material include a relatively high coefficient of thermal expansion (CTE) that may cause stresses to wafers if SU-8 is used with materials with widely different CTE values at elevated temperatures. However, the good patterning ability of SU-8 enables wafer-level batch fabrication of accurately defined microfluidic components [40].

principles. Our work takes that principle one step further by using surface plasmon waves generated on either side of a thin gold layer as the two branches of an interferometer.

1.2.1 Integration in Silicon-on-Insulator

Significant breakthroughs often happen when the power of unrelated disciplines are combined in yet unimaginable ways. Discovery at the intersection of biology and silicon technology can help bring together the better of two worlds: significant technical and manufacturing advances of the CMOS-industry combined with the power of biology. Silicon-on-insulator (SOI) consists of a thin Silicon layer on top of an oxide cladding layer deposited on a bare Silicon wafer. Silicon-on-insulator is a very good material for optical waveguides. With its Silicon core $n = 3.45$ and its oxide cladding ($n = 1.45$) it has a high vertical refractive index contrast [41]. Also, both the Silicon and the oxide are transparent at the telecom wavelengths of $1.3\mu m$ and $1.55\mu m$ [42]. The oxide thickness was chosen to be $2\mu m$ to reduce substrate leakage. The thickness of the core was chosen to be $220\ nm$, in order to keep the slab waveguide single mode for the TM polarization and to comply with standard fabrication technology available in-house [43].

1.2.2 This Work

As stated earlier, the main goal in integrated biosensor research is towards miniaturization and parallelization, while maintaining the same specifications when it comes to sensitivity. Mass-fabrication of cheap biosensor chips is another goal. Working with a high refractive index material system such as silicon-on-insulator is a straight-forward approach to meet the requirements for high-level integration and high-throughput fabrication.sensor research is towards miniaturization and parallelization, while maintaining the same specifications when it comes to sensitivity. Mass-fabrication of cheap biosensor chips is another goal. Working with a high refractive index material system such as silicon-on-insulator is a straight-forward approach to meet the requirements for high-level integration and high-throughput fabrication.

A high index contrast material system puts fundamental limits to the resonant excitation of surface plasmons at a gold water surface. Phase-matching of guided waveguide modes with surface plasmon modes can be obtained, but only for modes with an effective index that is much lower than the refractive index of the waveguide. Moreover, the operating spectral range of such a device will be limited and set by the conditions for phase-matching,

as is the case in conventional SPR waveguide sensors.

In this work, we propose a novel configuration for a biosensor in silicon-on-insulator (SOI) [44]. The basic element of our sensor is a surface plasmon interferometer consisting of a thin layer of gold embedded in the silicon membrane.

The first chapter is a general introduction into the world of surface plasmon polariton modes in general. The second chapter describes the simulation software that was specifically written for this work as a means to simulate surface plasmon polariton modes propagating on thin gold stripes. In chapter three we will completely describe the working principle of a surface plasmon interferometer. We will describe and define design rules for this new type of device and calculate theoretical limits of detection. The chapter ends with an extensive overview of the sensitivity of this devices and methods to improve this. In chapter four we will take a look at the fabrication steps required to fabricate the surface plasmon interferometer. The incoupling problem of TM polarization into single-mode TM waveguides is discussed here. Raw spectral data needs to be interpreted and be reduced to one single value (generally the wavelength of minimal transmission) characterizing either analyte refractive index or binding events at the surface. The resolution with which this value can be determined ultimately defines sensor sensitivity, therefore chapter five looks at several data analysis methods. Chapter six concludes this work with the proof-of-principle measurements, the bulk sensing results discussing of experimental detection limit, and finally, the ultimate goal, a biosensing experiment.

1.3 Publications

The scientific output of our research has been published in a number of international journals and has been presented at several national and international conferences. The following list gives an overview.

1.3.1 Patent Applications

1. P. Debackere, S. Scheerlinck, P. Bienstman, R. Baets, Integrated surface mode biosensor *PCT/EP2007/003441*

1.3.2 International Journals

1. P. Debackere, R. Baets, P. Bienstman, Bulk Sensing Experiments Using a Surface Plasmon Interferometer *Optics Letters, 34(18), p. 2858-2860, 2009*

2. D. Van Thourhout, G. Roelkens, R. Baets, W. Bogaerts, J. Brouckaert, P. Debackere, P. Dumon, S. Scheerlinck, J. Schrauwen, D. Taillaert , F. Van Laere, J. Van Campenhout, Coupling Mechanisms for a Heterogeneous Silicon Nanowire Platform. *Semiconductor Science and Technology, 23, p. 64004, 2008*

3. P. Debackere, S. Scheerlinck, P. Bienstman, R. Baets, Surface Plasmon Interferometer in Silicon-on-Insulator: Novel Concept for an Integrated Biosensor: Reply, *Optics Express, 15(21), p. 13651-13653, 2007*

4. P. Debackere, P. Bienstman, R. Baets, Adaptive Spatial Resolution: Application to Surface Plasmon Waveguide Modes, *Optical and Quantum Electronics, 38(9-11), p. 857-867, 2006*

5. P. Debackere, S. Scheerlinck, P. Bienstman, R. Baets, Surface Plasmon Interferometer in Silicon-on-Insulator: Novel Concept for an Integrated Biosensor, *Optics Express, 14(16), p. 7063-7072, 2006*

1.3.3 International Conference Proceedings

4. P. Debackere, P. Bienstman, R. Baets, Experimental Characterization of Biosensor based on Surface Plasmon Nano Interferometer, *Leos Annual Meeting, United States, p.TuY2, 2008*

5. G. Roelkens, W. Bogaerts, D. Taillaert, P. Dumon, L. Liu, S. Selvaraja, J. Brouckaert, J. Van Campenhout, K. De Vos, P. Debackere, D. Van Thourhout, R. Baets, Silicon Nanophotonics: towards VLSI Photonic Integrated Circuits, *URSI (invited), United States, 2008*

6. K. De Vos, P. Debackere, R. Baets, P. Bienstman, Label-free Biosensors on Silicon-on-Insulator Optical Chips based on Microring Cavities and Surface Plasmon Interferometry, *ICTON 2008 (invited), Greece, p. 88-91, 2008*

7. P. Debackere, K. De Vos, S. Scheerlinck, P. Bienstman, R. Baets, Silicon-on-Insulator as a Platform for Biosensors, *Frontiers Research Meeting, Toulouse, 2007*

8. P. Vandersteegen, Angel Ullan Nieto, Carl Van Buggenhout, S. Verstuyft, P. Bienstman, P. Debackere, Kristiaan Neyts, R. Baets, Employing a 2D Surface Grating to Improve Light Out-coupling of a Substrate Emitting Organic LED, In *Photonics West 2007: Integrated Optoelectronics Devices, 6486, United States, p.64860H ,2007*

9. P. Debackere, D. Taillaert, S. Scheerlinck, K. De Vos, P. Bienstman, R. Baets, Si based Waveguide and Surface Plasmon Sensors, *Photonics West 2007: Silicon Photonics II (invited), 6477, United States, p.647719, 2007*

10. P. Debackere, S. Scheerlinck, P. Bienstman, R. Baets, A Biosensor based on Surface Plasmon Interference *LEOS Benelux Annual Symposium, Netherlands, 2006*

11. P. Debackere, S. Scheerlinck, P. Bienstman, R. Baets, Surface Plasmon Interferometer in Silicon-on-Insulator: Novel Concept for an Integrated Biosensor, *Group IV Photonics , Canada, p.7-10, 2006*

12. P. Debackere, P. Bienstman, R. Baets. Improved ASR Convergence for the Simulation of Surface Plasmon Waveguide Modes *ICTON (COSTP 11 training school), p.We.P.15, 2006*

13. P. Debackere, P. Bienstman, R. Baets. Improved ASR Convergence for the Simulation of Surface Plasmon Waveguide Modes *OWTNM 2006, Italy, p.14 , 2006*

References

[1] C. Zandonella. *Cell nanotechnology: The tiny toolkit*. Nature, 423(6935):10–12, 2003.

[2] P. von Lode. *Point-of-care immunotesting: approaching the analytical performance of central laboratory methods*. Clin. Biochem., 38:591–606, 2005.

[3] World Bank. *World Bank World Development Report 1993: Investing in Health*, 1993.

[4] D. M. Morens, G. K. Folkers, and A. S. Fauci. *The challenge of emerging and re-emerging infectious diseases*. Nature, 430:242–249, 2004.

[5] A. S. Fauci. *Infectious diseases: considerations for the 21st century*. Clin. Infect. Dis., 32:675–685, 2001.

[6] National Research Council Committee on Science and Technology in Foreign Assistance. *The fundamental role of science and technology in international development: An imperative for the US agency for international development*, 2006.

[7] `http://www.cdc.gov/nczved/dfbmd/`.

[8] `http://www.who.int/mediacentre/factsheets/fs255/en/`.

[9] P. Yager, T. Edwards, E. Fu, K. Helton, K. Nelson, R. M. Tam, and B. H. Weigl. *Microfluidic diagnostic technologies for global public health*. Nature, 442:412–418, 2006.

[10] B. Robertson and K. A. Nicholson. *New microbiology tools for public health and their implications*. Annu. Rev. Public Health, 26:281–302, 2005.

[11] C. P. Price. *Regular review: point of care testing*. Brit. Med. J., 332:1285–1288, 2001.

[12] C. D. Chin, V. Linder, and S. K. Sia. *Lab-on-a-chip devices for global health: Past studies and future opportunities*. Lab Chip, 7:41–57, 2007.

[13] H. Andersson and A. van den Berg. *Microtechnologies and nanotechnologies for single-cell analysis*. Current Opinion in Biotechnology, 15(1):44–49, 2004.

[14] X. Fan, I. M. White, S. I. Shopova, H. Zhu, J. D. Suter, and Y. Sun. *Sensitive optical biosensors for unlabeled targets: a review*. Anal. Chem. Acta, 620(8):8–26, 2008.

[15] D. R. Shankaran and N. Miura. *Trend sin interfacial design for surface plasmon resonance based immunoassays*. Journal of Physics D - Applied Physics, 40(23):7187–7200, 2007.

[16] W. Lukosz. *Integrated optical chemical and direct biochemical sensors*. Sensors and Actuators B -Chemical, 29:37–50, 1995.

[17] B. Liedberg, C. Nylander, and I. Lundstrom. *Surface-plasmon resonance for gas detection and biosensing*. Sensors and Actuators, 4(2):299–304, 1983.

[18] H. Raether. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer Tracts in Modern Physics 111. Springer-Verlag, Berlin, 1988.

[19] J. R. Sambles, G. W. Bradberry, and Fuzi Yang. *Optical excitation of surface plasmons: an introduction*. Contemporary Physics, 32(3):173–183, 1991.

[20] X. D. Hoa, A. G. Kirk, and M. Tabrizian. *Towards integrated and sensitive surface plasmon resonance biosensors: A review of recent progress*. Biosensors & Bioelectronics, 23(2):151–160, 2007.

[21] `http://www.sensata.com/contact/spreeta-analytical-sensor.htm`.

[22] K. Kurihara, H. Ohkawa, Y. Iwasaki, O. Niwa, T. Tobita, and K. Suzuki. *Fiber-optic conical microsensors for surface plasmon resonance using chemically etched single-mode fiber*. Analytica Chimica Acta, 523(2):165–170, 2004.

[23] M. Piliarik, J. Homola, Z. Manikova, and J. Ctyroky. *Surface plasmon resonance sensor based on a single-mode polarization-maintaining optical fiber*. Sensors and Actuators B -Chemical, 90(1-3):236–242, 2003.

[24] P. Stöcker, B. Menges, U. Langbein, and S. Mittler. *Multimode waveguide mode surface plasmon coupling: a sensitivity and device realizability study*. Sensors and Actuators A -Physical, 116(2):224–231, 2004.

[25] C. W. Lin, K. P. Chen, C. N. Hsiao, S. M. Lin, and C. K. Lee. *Design and fabrication of an alternating dielectric multi-layer device for surface plasmon resonance sensor*. Sensors and Actuators B -Chemical, 113(1):169–176, 2006.

[26] A. Suzuki, J. Kondoh, Y. Matsui, S. Shiokawa, and K. Sozuki. *Development of novel optical waveguide surface plasmon resonance (SPR) sensor with dual Light Emitting Diodes*. Sensors and Actuators B -Chemical, 106:383–387, 2005.

[27] P. I. Nikitin, A. A. Beloglazov, V. E. Kochergin, M. V. Valeiko, and T. I. Ksenevich. *Surface plasmon resonance interferometry for biological and chemical sensing*. Sensors and Actuators B -Chemical, 54(1-2):43–50, 1999.

[28] P. I. Nikitin and A. A. Beloglazov. *A multipurpose sensor based on surface plasmon polariton resonance in a Schottky structure*. Sensors and Actuators A -Physical, 42(1-3):547–552, 1994.

[29] M. Piliarik, H. Vaisocherova, and J. Homola. *Towards parallelized surface plasmon resonance sensor platform for sensitive detection of oligonucleotides*. Sensors and Actuators B -Chemical, 121(1):187–193, 2007.

[30] J. Homola. *Present and Future of Surface Plasmon Resonance Biosensors*. Anal. Bioanan. Chem., 377:528–539, 2003.

[31] J. Homola, S. Y. Sinclair, and G Gauglitz. *Surface plasmon resonance sensors: review*. Sensors and Actuators B -Chemical, 54:3–15, 1999.

[32] E. Fontana. *A novel gold-coated multimode fiber sensor*. IEEE Transactions on Microwave Theory and Techniques, 50(1):82–87, 2002.

[33] J. F. Masson, Y. C. Kim, L. A. Obando, W. Peng, and K. S. Booksh. *Fiber-optic surface plasmon resonance sensors in the near-infrared spectral region*. Applied Spectroscopy, 60(11):1241–1246, 2006.

[34] H. J. M. Kreuwel, P. V. Lambeck, J. v Gent, and T. J. A. Popma. *Surface plasmon dispersion and luminescence quenching applied to*

planar waveguide sensors for the measurement of chemical concentrations. Proceedings of the SPIE - The International Society for Optical Engineering—Proceedings of the SPIE - The International Society for Optical Engineering, 798:218–24, 1987.

[35] J. Čtyrocký, J Homola, P.V. Lambeck, S. Musa, H.J.W.M. Hoekstra, R.D. Harris, J.S. Wilkinson, and N.M. Usievich, B.and Lyndin. *Theory and Modelling of Optical Waveguide Sensors Utilising Surface Plasmon Resonance*. Sensors and Actuators B -Chemical, 54:66–73, 1999.

[36] R. D. Harris and J. S. Wilkinson. *Waveguide surface plasmon resonance sensors*. Sensors and Actuators B -Chemical, 29:261–267, 1995.

[37] J. Homola, J. Čtyrocký, M. Skalský, J. Hradilová, and P. Kolářová. *A Surface Plasmon Resonance Based Integrated Optical Sensor*. Sensors and Actuators B -Chemical, 38-39:286–290, 1997.

[38] J Dostálek, J Čtyrocký, J Homola, E Brynda, M Skalský, P Nekvindová, J Špiriková, J Škvor, and J Schröfel. *Surface plasmon resonance biosensor based on integrated optical waveguide*. Sensors and Actuators B -Chemical, 76:8–12, 2001.

[39] B. Sepúlveda, J. Sánchez del Rí o, M. Moreno, F. J. Blanco, K. Mayora, K. Domínguez, and L. M. Lechuga. *Optical biosensor microsystems based on the integration of highly sensitive Mach-Zehnder interferometer devices*. J. Opt. A: Pure Appl. Opt., 8:S561–S566, 2006.

[40] K. Y. Lee, N. LaBianca, S. A. Rishton, S. Zolgharnain, J. D. Gelorme, J. Shaw, and T. H.-P. Chang. *Micromachining applications of a high resolution ultrathick photoresist*. volume 13, pages 3012–3016. AVS, 1995.

[41] W. Bogaerts, D. Taillaert, B. Luyssaert, P. Dumon, J. Van Campenhout, P. Bienstman, D. Van Thourhout, R. Baets, V. Wiaux, and S. Beckx. *Basic structures for photonics integrated circuits in Silicon-on-Insulator*. Optics Express, 12(8):1583–1591, 2004.

[42] P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout, and R. Baets. *Low-loss SOI photonics wires and ring resonators fabricated with deep-UV lithography*. IEEE Photonics Technology Letters, 16(5):1328–1330, 2004.

[43] W. Bogaerts, P. Dumon, D. Taillaert, V. Wiaux, S. Beckx, B. Luyssaert, J. Van Campenhout, D. Van Thourhout, and R. Baets. *SOI Nanophotonic Waveguide Structures Fabricated with Deep UV Lithography ,anophotonic waveguide structures fabricated with deep UV lithography*. Photonics and Nanostructures: Fundamentals and Applications, 2(2):81–86, 2004.

[44] W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, and D. Van Thourhout. *Nanophotonics waveguides in silicon-on-insulator fabricated with CMOS technology*. Journal of Lightwave Technology, 23(1):401–412, 2005.

2

Surface Plasmon Polaritons

2.1 Metal Properties

IT is generally known that metals are good conductors of electricity and heat and are good reflectors of radiation. This is a rather loose definition of a metal which basically relates to the ability of the free electrons in the metal to respond to externally imposed electromagnetic fields. Free electrons (although obviously still physically constrained within the metal) are able to respond to the incident radiation without scattering, giving an ideal metallic response. Such a response, which completely excludes $\mathbf{E}$ from the metal must therefore have $\epsilon_r = \pm\infty$. As there is no such thing as perfection, such a material does not actually exist, for the free electrons inside a metals cannot respond infinitely quickly to an imposed electromagnetic field. Electrons have a tiny but finite mass and they do suffer resistive scattering from lattice vibrations (phonons), lattice defects and the surface itself. These effects lead to damping of the externally applied oscillation, hence the dielectric constant ϵ will be composed of a negative, real part ϵ' and an imaginary part ϵ''.

Increasing the frequency of the externally imposed field then leads to the fact that the electrons have an increasingly more difficult time of responding. Ultimately, at a high enough frequency, a low enough wavelength and thus sufficiently high energy, the metal becomes transparent and behaves more like a dielectric. The frequency for which metals are no longer metals (for

many metals in the ultra-violet region) is called the plasma-frequency.

2.2 Plasmons: Electromagnetic Theory

Eigenmodes of collective oscillations of the quasi-free electrons in metals are called plasmons. Since electrons carry charge, these oscillations are inherently associated with an electromagnetic field. A theoretical description of (surface) plasmons has to include this interplay of charges and fields[1]. Three types of plasmon modes can be distinguished. They emerge by solving Maxwell's equations with appropriate boundary conditions for either bulk material, a planar metal-dielectric interface or a metallic particle. In the literature these three types are commonly referred to as bulk plasmons, surface plasmons and particle plasmons, respectively.

In the classical model of a metal the positive ions are replaced by a uniform background medium with a positive charge density equal to the average charge density of all electrons but of opposite sign. In this model all electrons can be described as a gas with fluctuating density, which cause periodic Coulomb attraction or repulsion of electrons, giving rise to longitudinal collective oscillations of the electronic gas or *plasma oscillations*. A quantum of these oscillations is called a *plasmon*, or bulk plasmon since it exists in a bulk metal. The plasma frequency of a metal represents the resonant frequency for bulk plasmons.

Surface plasmons were first observed experimentally back in 1902, when Wood illuminated a metallic diffraction grating with polychromatic light and discovered narrow dark bands in the spectrum of the diffracted light, which he referred to as anomalies [1]. Theoretical work by Fano [2] led to the conclusion that these anomalies were associated with the extinction of electromagnetic surface waves on the surface of the diffraction grating. Another 17 years later Thurbadar observed a large drop in reflectivity when illuminating thin metallic films on a substrate, but did not link this effect to surface plasmons. Otto finally explained Thurbador's result in 1968 and showed that the drop in reflectivity in the attenuated total reflection spectrum was due to the resonant excitation of surface plasmon modes on the metal surface. In the same year Kretchman and Raether reported excitation of surface plasmon modes in another configuration, paving the way for an

[1] Quasi particles resulting from strong coupling of electromagnetic waves with an electric or magnetic dipole-carrying excitation are generally called polaritons in physics, they describe the crossing of the dispersion of light with any interacting resonance. The coupling of a plasmon quasi-particle and a photon is therefore denominated as a surface plasmon polariton (SPP). So surface plasmon and surface plasmon polariton are interchangeable names for the same physical phenomenon.

efficient and convenient method for the excitation of surface plasmon modes, introducing plasmonics into modern optics [3].

In this work we will only deal with surface plasmon modes and will hence limit the theoretical description of plasmons to this specific case.

2.2.1 Theoretical Derivation of Surface Plasmon Modes

Surface plasmons are electromagnetic modes bound to metal-dielectric interfaces, involving charges in the metal and electromagnetic fields in both media. The field intensity in the metal as well as the dielectric medium decays exponentially normal to the surface. In the plane of the interface, the field intensity and charge distribution in the metal propagates as a longitudinal wave along the surface. Hence, surface plasmons propagate along the interface with electromagnetic fields, energy and charges highly localized within the interface area. Their properties depend strongly on the exact properties of both the metal (complex dielectric function, corrugations, roughness) and the dielectric (refractive index). The sensitivity to the local refractive index of the dielectric close to the interface is used to detect chemical binding by taking a liquid as dielectric. When solvents present in this liquid bind to the metal-dielectric interface, they change the refractive index at the interface, which leads to a measurable change in surface plasmon properties.

To understand surface plasmons, we will take a look at the results of simple electrodynamical theory applied to an ideal model interface. Obviously this approach neglects many real effects due to the exact surface properties, screening effects and more. However, the main features of surface plasmons can be understood in this simple model.

We consider an interface in the yz-plane between two half-infinite spaces (shown in Fig. 2.1), labeled dielectric and metal, materials of which the optical properties are described by their complex frequency-dependent dielectric functions. We ignore magnetic materials. Surface polaritons can only be excited at such an interface if the dielectric displacement $\mathbf{D}$ of the electromagnetic mode has a component normal to the surface which can induce a surface charge density σ,

$$(\mathbf{D_2} - \mathbf{D_1}).\mathbf{x} = \sigma \tag{2.1}$$

S-polarized light propagating along the z-direction possesses only electric field components, $\mathbf{E}$, parallel to the surface , i.e. transversal electric (TE) waves have $\mathbf{E} = (\mathbf{0}, \mathbf{E_y}, \mathbf{0})$, and hence are unable to excite surface plasmon polaritons. Only p-polarized light (transversal magnetic TM) modes with

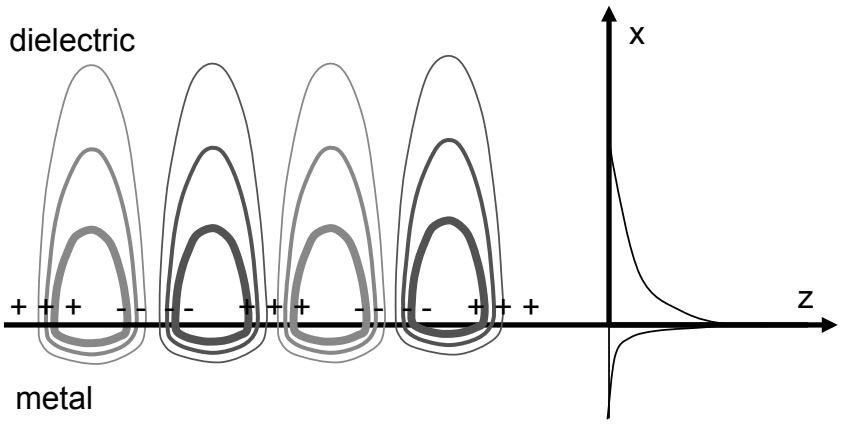


Figure 2.1: The charges and $\mathbf{E}$ *field of a surface plasmon propagating on a surface in the* x *direction. The exponential dependence of the* E_z *field is seen on the right*

$\mathbf{E} = (\mathbf{E_x}, \mathbf{0}, \mathbf{E_z})$, or, equivalently, $\mathbf{H} = (\mathbf{0}, \mathbf{H_y}, \mathbf{0})$, can couple to such modes.

One can also explain this by appealing to the underlying physics of surface plasmons. A surface plasmon polariton consists of two inhomogeneous plane waves (one in the dielectric and one in the metal), both having the same propagation constant. The diagram in Fig. 2.2a represents the fields of an SP, with the E fields originating on the positive charges and terminating on negative ones. If the continuity of the tangential component of the H-field ($H_{\|}$) is assumed, then a negative ϵ_m ensures the continuity of the normal D component ($D_{\perp}$), while $E_{\|}$ can be made continuous by a proper choice of propagation constant. In contrast, the diagram presented in Fig. 2.2b represents a physical impossibility: the absence of magnetic charges in nature means that the H field needs to be divergence free everywhere and in particular at the metal-dielectric interface. However, since $H_{\|}$ will now have opposite directions above and below the surface, it cannot satisfy the requisite boundary condition. This is why surface plasmon polaritons, out of necessity, be TM-polarized [4].

The resulting surface electromagnetic wave, will have the following general form [5, 6]

$$\mathbf{E}_1 = \mathbf{E}_{10} e^{-j(k_{x1}x + k_{z1}z - \omega t)} \tag{2.2}$$

$$\mathbf{H}_1 = \mathbf{H}_{10} e^{-j(k_{x1}x + k_{z1}z - \omega t)} \tag{2.3}$$

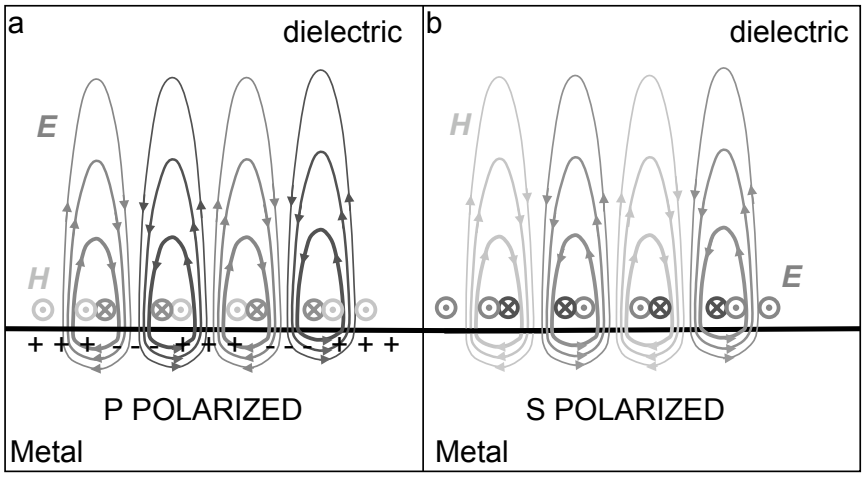


Figure 2.2: (a) The SPs E-fields originate on positive charges and terminate on negative ones. Continuity of $H_{\|}$ and a negative ϵ_m ensure the continuity of $D_{\perp}$, while $E_{\|}$ becomes continuous.(b) A physical impossibility, since the divergence-free nature of the H-field requires $H_{\|}$ to have opposite directions above and below the surface, thus prohibiting the continuity of $H_{\|}$ at the boundary.

in medium 1; $x < 0$,and

$$\mathbf{E}_2 = \mathbf{E}_{20} e^{-j(k_{x2}x - k_{z2}z - \omega t)} \tag{2.4}$$

$$\mathbf{H}_2 = \mathbf{H}_{20} e^{-j(k_{x2}x - k_{z2}z - \omega t)} \tag{2.5}$$

in medium 2, $x > 0$ where k_{z1} and k_{z2} are the wavevectors in the z-direction, k_{x1} and k_{x2} those in the x-direction, i.e. normal to the interface, and ω is the angular frequency. Both fields must fulfill Maxwell's equations, which in an isotropic, nonmagnetic, homogeneous medium are given by

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\mu_0 \frac{\partial}{\partial t} \mathbf{H}(\mathbf{r}, t) \tag{2.6}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \epsilon_0 \epsilon_r \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t) \tag{2.7}$$

$$\nabla . \mathbf{H} = 0 \tag{2.8}$$

$$\nabla . \mathbf{E} = 0 \tag{2.9}$$

Electromagnetic boundary conditions dictate that the tangential components

of $\mathbf{E}$ and $\mathbf{H}$ have to be equal at the interface, i.e.

$$E_{z1} = E_{z2} \tag{2.10}$$
$$H_{y1} = H_{x2} \tag{2.11}$$
$$\tag{2.12}$$

From Equation 2.10 it follows immediately that $k_{z1} = k_{z2} = k_z$. On the other hand, from Equations 2.2,2.4 and 2.7 we can derive that

$$k_{x1}H_{y1} = \epsilon_0\epsilon_{r1}\omega E_{z1} \tag{2.13}$$
$$k_{x2}H_{y2} = -\epsilon_0\epsilon_{r2}\omega E_{z2} \tag{2.14}$$

This can only lead to a nontrivial solution if

$$\frac{k_{x1}}{k_{x2}} = -\frac{\epsilon_{r1}}{\epsilon_{r2}} \tag{2.15}$$

This equation indicates that surface electromagnetic modes can exist at interfaces between two media with dielectric constants of opposite sign. For a material in contact with a dielectric medium which has a positive dielectric constant ϵ_{rd}, this can be fulfilled for a whole variety of possible elementary excitations provided their oscillator strength is sufficiently strong to result, for a narrow spectral range, in a negative ϵ_r. Here we are dealing with the interface between a metal with its complex relative dielectric function ($\epsilon_{rm} = \epsilon'_{rm} + \epsilon''_{rm}$) and a dielectric material (ϵ_{rd}), hence, with coupling the collective plasma oscillations of the nearly free electron gas in a metal to an electromagnetic field. The condition that ϵ'_{rm} ought to be negative is largely fulfilled by several metals in the visible and the near-infrared parts of the spectrum where ϵ'_{rm} generally has a large negative real part (the small imaginary part being largely associated with absorption and scattering losses in the metal). For example, gold at a wavelength of $830\ nm$ has a relative permittivity of $\epsilon_{rm} \approx -29 + 2.1j$ [7].

Combining Maxwell's curl equations 2.6 and 2.7 with equations 2.13 and 2.14 lead to

$$k_z^2 - k_{x1}^2 = \epsilon_{r1}\left(\frac{\omega}{c}\right) \tag{2.16}$$
$$k_z^2 - k_{x2}^2 = \epsilon_{r2}\left(\frac{\omega}{c}\right) \tag{2.17}$$
$$\tag{2.18}$$

which combined with equation 2.15 leads to the dispersion relationships (i.e. the energy-momentum relation) for surface plasmons at a metal/dielectric interface:

$$k_z = \left(\frac{\omega}{c}\right)\sqrt{\frac{\epsilon_d\epsilon_m}{\epsilon_d + \epsilon_m}}, \tag{2.19}$$

where ϵ_d is the dielectric function of the dielectric, ϵ_m the dielectric function of the metal [2], k_z is the propagation constant of the surface plasmon mode and k_0 is the free-space wave-vector. The relative permittivity of a dielectric is usually only weakly dispersive so that most of the interesting physics arises from the intriguing behavior of the relative permittivity of metals. Assuming that $\epsilon''_m < |\epsilon'_m|$, we can write the complex wave-vector as

$$k'_z = \frac{\omega}{c}\left(\frac{\epsilon'_m \epsilon_d}{\epsilon'_m + \epsilon_d}\right)^{1/2} \tag{2.20}$$

$$k''_z = \frac{\omega}{c}\left(\frac{\epsilon'_m \epsilon_d}{\epsilon'_m + \epsilon_d}\right)^{3/2} \frac{\epsilon''_m}{2(\epsilon'_m)^2} \tag{2.21}$$

This dispersion relations places an additional condition on ϵ_d and ϵ_m, namely that in addition to the requirement that $\epsilon_d \epsilon'_m < 0$ [8] (which is more or less a given since medium m is a metal), it is also required that $\epsilon_d < |\epsilon'_m|$ so that the mode can propagate. Solving the equation for k_{xi} leads to purely imaginary values so we see that the surface plasmon is a bound, non-radiative evanescent wave whose field amplitude is maximum at the interface and which is decaying exponentially into the dielectric (and into the metal). The mode is propagating as a damped oscillatory wave.

An example of the dispersion of surface plasmons at a gold-air interface is shown in Fig. 2.3 together with the dispersion of plane electromagnetic waves in air. At low frequencies the surface mode lies very close to the light line and is predominantly light-like, it is in this region that is is best described as a polariton. As the frequency rises, the mode moves further away from the light line, gradually approaching the asymptotic limit, the surface plasmon resonant frequency, which - if the metal can be described using the Drude model (see appendix A)- is given by

$$\omega_{SPP} = \frac{\omega_p}{\sqrt{1+\epsilon_d}}, \tag{2.22}$$

where ω_p is the plasma frequency of the metal. This occurs when the relative dielectric function of the metal and dielectric are of the same magnitude but opposite sign, producing a pole in the dispersion relation 2.19. At this point the group velocity of the surface plasmon modes decreases sharply until it reaches the limit of almost stationary waves. The high energy modes, with a frequency above the plasma frequency ω_p should be neglected because gold, or any other metal for that matter, does not behave like a metal at those

[2] In order not to overload notation we will omit the subscript r for relative dielectric functions

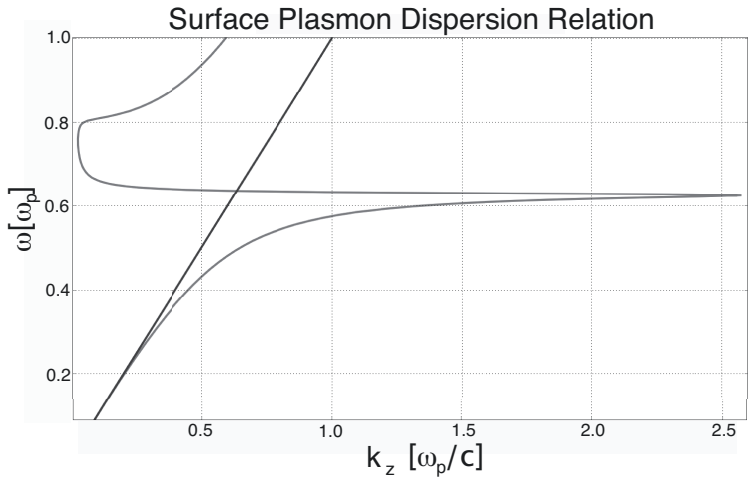


Figure 2.3: The dispersion relation of non-radiative surface plasmon modes, right of the light line $\omega = ck_z$(blue). At large k_z the value of ω approaches $\omega_p/\sqrt{1+\epsilon_r d}$.

frequencies. The optical properties are dominated by excitations of electrons from deeper lying bands. These modes also propagate into the bulk and are thus not true surface modes.

The surface plasmons discussed above are coupled oscillations of electron density and electromagnetic field. The correct term for this is surface plasmon polariton. Surface plasmons are, strictly speaking, the (unphysical) electron density oscillation modes without the coupling with the field. Their energy corresponds to the limiting stationary value in Fig. 2.3. Because the distinction of surface plasmons and surface plasmon polaritons is not central in the context of this work we will speak of surface plasmons when referring to the coupled mode.

2.2.2 Spatial extension of the Surface Plasmon Fields

The field amplitudes of the surface plasmon wave decrease exponentially as $exp(-|k_{xi}||x|)$, normal to the surface. the value of the skin-depth at which point the field amplitude falls to $1/e$, becomes

$$x_{skin} = \frac{1}{|k_{xi}|} \tag{2.23}$$

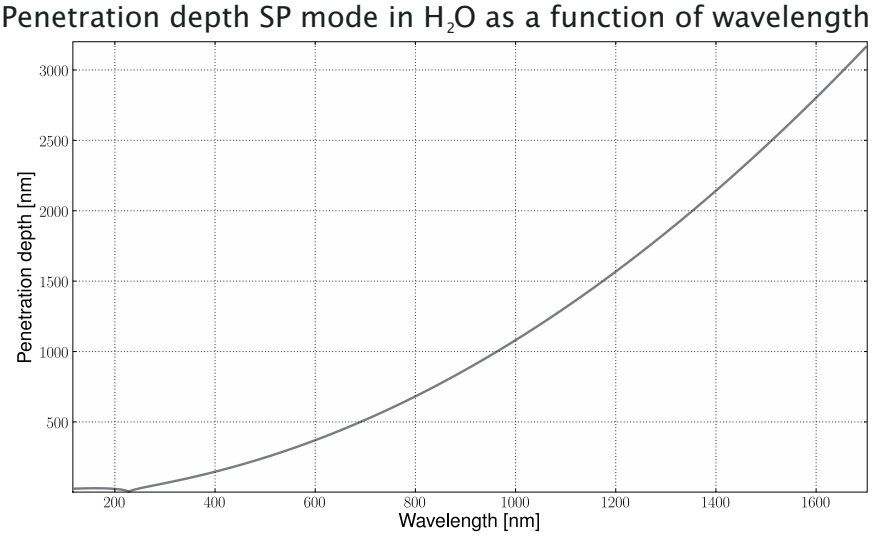


Figure 2.4: Variation of skin depth of the dielectric medium (air) with free-space wavelength in the visible and near-infrared, Au - air interface.

or, in the dielectric

$$x_{skin} = \frac{\lambda}{2\pi}\left(\frac{\epsilon'_m + \epsilon_d}{\epsilon_d^2}\right)^{1/2} \tag{2.24}$$

and in the metal

$$x_{skin} = \frac{\lambda}{2\pi}\left(\frac{\epsilon'_m + \epsilon_d}{\epsilon_m'^2}\right)^{1/2} \tag{2.25}$$

The penetration depth in the dielectric is plotted in figure 2.4 as a function of wavelength. From this figure we see that for wavelengths in the visible spectral region the penetration depth into the dielectric is less than the free-space wavelength whilst at the infrared end it is more than the free-space wavelength. This increase, relative to the free-space wavelength, arises because as one moves to longer wavelengths the metal is a better conductor and the mode has a wave-vector that is closer to the free-space wave-vector (and is thus more light like)

The short wavelength limit, so large k_z vectors, has a skin depth approximately equal to $1/k_z$ leading to a very strong localization of the fields near the surface in both media. In this special case both electric field components (E_z and E_x) are approximately equal, $E_x = jE_z$ in the dielectric and $E_x = -jE_z$ in the metal. In the long surface plasmon wavelength

limit (low k_z vector), the field in the dielectric has a quite strong transverse component E_x, compared to the longitudinal component E_z, namely $E_x/E_z = -j|\epsilon'_m|^{1/2}$ and extends far in the dielectric region, thus resembling a guided photon field or Zenneck-Sommerfeld wave [9][3]. In the metal, E_x is small since $E_x/E_z = j|\epsilon'_m|^{-1/2}$, thus producing a longitudinal field.

The penetration depth of the field into the dielectric gives us a measure of the length scale over which the SPP mode is sensitive to refractive index changes, for example, for the binding of specific biomolecules in a biosensor. We can see that, although the sensitivity of the SPP to changes in the dielectric medium falls of exponentially, the distance over which the fall-off takes place is quite large on a molecular scale. Associated with the localization of the field near the metal there is also a field enhancement, owing to the increase in local photonic mode density [11]. This enhanced field has many consequences, but the one that is most important here is that this is the reason why they are so sensitive to changes at the surface.

In this work, we will deal mainly with SPP modes propagating on the Au-H_2O and Au-Si interface, the skin depth of these SPP's in the dielectric will determine the sensitivity, while the skin depth in the metal will determine the mutual coupling of two surface plasmon mode propagating on either interface of a thin Au film. One can see from figures 2.5 and 2.6 that the skin in the metal is approximately equal to $23\ nm$, while the skin depth in both dielectrics is a few orders of magnitude larger, thus supporting the these that these modes are light-like.

2.2.3 Propagation length

The surface plasmon propagation length δ_{SPP} can be found from the imaginary part of the surface plasmon polariton wave-vector (equation 2.21). δ_{SPP} is the distance over which the power/intensity of the mode falls to $1/e$ of its initial value, is given by $\delta_{SPP} = 1/2k''_z$

$$\delta_{SPP} = \lambda \frac{(\epsilon'_m)^2}{2\pi\epsilon''_m} \left(\frac{\epsilon'_m + \epsilon_d}{\epsilon'_m \epsilon_d} \right)^{3/2} \tag{2.26}$$

[3] J. Zenneck [10], in 1907, was the first to analyze a solution of Maxwell's equations that had a surface wave property. This so-called Zenneck wave is simply a vertically polarized plane wave solution to Maxwell's equations in the presence of a planar boundary that separates free space from a half space with a finite conductivity. For large conductivity - this depends on the frequency and dielectric constant, too - such a wave has a Poynting vector that is approximately parallel to the planar boundary. The amplitude of this wave decays exponentially in the directions both parallel and perpendicular to the boundary (with differing decay constants).

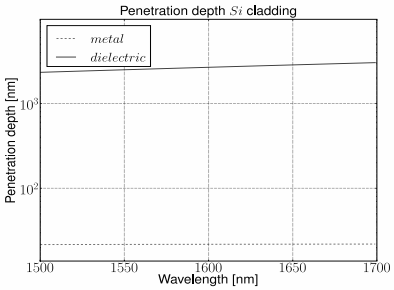


Figure 2.5: Variation of skin depth of Si and Au with free-space wavelength in the range 1500-1700 nm.

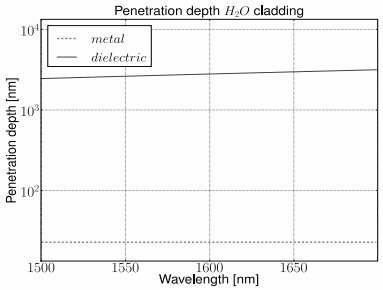


Figure 2.6: Variation of skin depth into H_2O and Au varies with free-space wavelength in the range 1500-1700 nm.

When the metal is low-loss, and the condition $|\epsilon'_m| \ll \epsilon_d$ is satisfied, we can approximate the propagation length as

$$\delta_{SPP} \approx \lambda \frac{(\epsilon'_m)^2}{2\pi\epsilon''_m} \tag{2.27}$$

From this equation we can see that for a long propagation length we require a large negative real part of the relative permittivity ϵ'_m and a small imaginary part ϵ''_m, i.e. we need a low loss metal, as one would expect. The propagation lengths calculated using Eqn. 2.27 and based on the Drude parameters of gold (taken from [12, 13]) are shown in figure 2.7. Several implications can be drawn from this data. First, if one is considering trying to construct photonic components or circuits based on SPP's (as is the intent of this PhD) then the propagation lengths represent an upper limit on the size of the structures one can contemplate using. The above data clearly shows that the absorptive nature of metals places severe restrictions on the potential circuit size. Second, since the propagation length is significantly larger that the wavelength of the surface plasmon polariton, wavelength-scale gratings and other periodic surface structures can be used to manipulate surface plasmon polariton modes since the modes are able to interact over many periods of such a structure.

As a comparison, and a means for future reference, the propagation lengths for a surface plasmon mode propagating in a Si substrate and a H_2O substrate were calculated using eqn. 2.26 and are shown in figure 2.8 in the wavelength region 1500 to 1700 nm. While the propagation length in H_2O is quite considerable, one can see that the surface plasmon mode propagating on the Au-Si interface is very lossy.

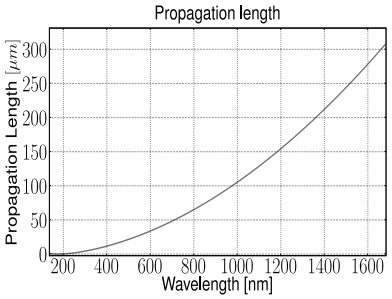


Figure 2.7: Propagation length of the surface plasmon mode propagating on a $Au-air$ interface. The data was computed using the Drude model, parameters taken from [12, 13]. The relative permittivity of the dielectric was equal to 1.

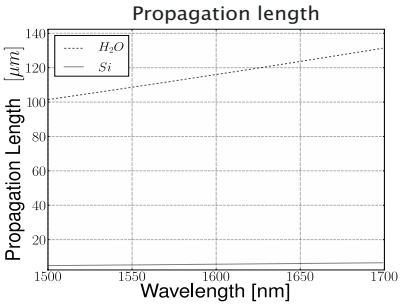


Figure 2.8: Propagation length of the surface plasmon polariton propagating on the $Au - Si$ interface and a $Au - H_2O$ interface. The data was computed using the Drude model, parameters taken from [12, 13]. The relative permittivity of the dielectric was equal to 1.

2.2.4 Field Enhancement

In order to excite surface plasmons the wave-vector of the incoming beam has to match the wave-vector of the surface plasmon. This is possible by passing the light through a block of glass to increase the wave-vector and evanescently couple this light to the surface plasmon on the metal surface.

Since the SPP field is concentrated close to a metal surface, it is significantly enhanced at the surface. The field enhancement depends on the dielectric constant of the metal and the radiative losses associated with a surface roughness and/or a metal film thickness. The maximum possible field enhancement on a smooth surface can be estimated from

$$\frac{|E_{SPP}^{x=0}|^2}{|E_{light}|} = \frac{2}{\epsilon_d}\frac{|\epsilon_m'|^2}{\epsilon_m''}\frac{a}{1+\epsilon_m'} \tag{2.28}$$

where $a^2 = |\epsilon_m'|(\epsilon_s - 1 - \epsilon_s)$, E_{light} is the electric field of the incident light, and ϵ_s is the dielectric constant of the prism through which the metal film is illuminated to excite SPP's. This enhances, for example, the intensity by more than two orders of magnitude for a 60 nm thick silver film illuminated with red light [14], for a 50 nm thick Au-film, illuminated through a SiO_2-prism at a wavelength of 1.55 μm, the intensity of the field is enhanced $\approx$ 150 times.

.

References

[1] R. W. Wood. *On a remarkable case of uneven distribution of light in a diffraction grating spectrum*. Philosophical Magazine, 4:396, 1902.

[2] U. Fano. *The theory of anomalous diffraction gratings and of quasi-stationary waves on metallic surfaces (Sommerfeld's waves)*. Journal of the Optical Society of America, 31(3):213–222, 1941.

[3] B. Liedberg, C. Nylander, and I Lundström. *Biosensing with surface plasmon resonance – How it all started*. Biosensors & Bioelectronics, 10:i–ix, 1995.

[4] A. R. Zakharian, J. V. Moloney, and M. Mansuripur. *Surface plasmon polaritons on metallic surfaces*. In 17th Annual Magnetic Recording Conference (TMRC 2006), pages 845–850, Pittsburgs, PA, 2006.

[5] J. J. Burke, G.I. Stegeman, and T. Tamir. *Surface-polariton waves guided by thin, lossy metal films*. Physical Review B, 33(8):5186–5201, 1986.

[6] S. A. Maier. *Plasmonics: Fundamentals and Applications 1*. Springer, 2007.

[7] W. L. Barnes. *Surface-plasmon polariton length scales: a route to subwavelength optics*. J.Opt. A.: Pure Appl. Opt., 8:S87–S93, 2006.

[8] E. Burstein, W.P. Chen, and A. Hartstein. *Surface polaritons - propagating electromagnetic modes at interfaces*. Journal of Vacuum Science Technology, 11(6):1004–1019, 1974.

[9] H. Raether. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer Tracts in Modern Physics 111. Springer-Verlag, Berlin, 1988.

[10] J. Zenneck. *Über die Fortpflanzung ebener elektromagnetischer Wellen einer ebenen Leiterflache und ihre Beziehung zur drahtlosen Telegraphie*. Ann. der Physik, 23:846–866, 1907.

[11] W. L. Barnes. *Fluorescence near interfaces: The role of photonics mode density*. Journal of Modern Optics, 45:661–699, 1998.

[12] P. G. Etchegoin, E. C. Le Ru, and M. Meyer. *An analytical model for the optical properties of gold*. The Journal of Chemical Physics, 125:164705, 2006.

[13] P. G. Etchegoin, E. C. Le Ru, and M. Meyer. *Erratum: "An analytic model for the optical properties of gold"[J. Chem.Phys. 125,164705 (2006)]*. The Journal of Chemical Physics, 167:189901, 2007.

[14] A. V. Zayats and II Smolyaninov. *Near-field photonics: surface plasmon polaritons and localized surface plasmons*. Journal of Optics A: Pure and Applied Optics, 5:S16–S50, 2003.

3
Simulation Methods

AS we have seen in the previous chapter, surface plasmons can be described by macroscopic electromagnetic theory, i.e. Maxwell's equations, since at optical frequencies, the electron mean free path in the metal is much shorter than the plasmon wavelength. In macroscopic electromagnetic theory, bulk material properties such as dielectric constants, are used to describe objects irrespective of their size, which is a good approximation in our case since the waveguides described here a quite large. In this chapter,we give an overview of the problems plaguing simulation techniques for plasmonic devices, and explain in intricate detail the solutions proposed here. All materials are assumed to be non-magnetic ($\mu = \mu_0$) and are characterized by their bulk dielectric constant ϵ_r.

Numerical modeling of plasmonic devices involves several challenges which need to be addressed. First, as mentioned above, plasmonic devices can have arbitrary geometries. Several techniques are specific for one type of geometrical configuration and are therefore not appropriate for modeling of arbitrary plasmonic devices. In this thesis we will be dealing mostly with planar geometries.

The second problem is related to the complex and frequency-dependent nature of the dielectric constant of the metal. Several simulation techniques which are limited to lossless, non-dispersive materials are not applicable to plasmonic devices. In addition, in time-domain methods the dispersion properties of metals have to be approximated by suitable analytical expressions [1, 2].

Third, in surface plasmons propagating along the interface of a metal and a dielectric, the field is concentrated at the interface, and decays exponentially away from the interface in both the metal and dielectric regions [3]. For numerical methods based on discretization of the fields on a numerical grid, a very fine grid resolution is required at the metal-dielectric interface to adequately resolve the local fields [3]. The nanoscale feature sizes of plasmonic devices pose an extra challenge to numerical simulation techniques.

Several different simulation techniques are currently being used to study surface plasmons. Frequency-domain techniques (such as FDFD [4], FEFD [5], eigenmode expansion) can treat arbitrary material dispersion but require a large number of simulations to obtain the broadband response. Keeping simulation time within reasonable limits is required to use this technique for broadband simulations.

In time-domain techniques the entire frequency response can be obtained with a single simulation. However, the treatment of dispersion requires additional computational and storage cost and the analytical approximations used result in substantial error [6]. These techniques will not be discussed here.

3.1 Fourier Modal Method and Eigenmode Expansion

Eigenmode expansion is also a frequency-domain technique, applicable to planar geometries where we can divide the structure in a number of layers where the refractive index profile does not change in the z-direction. The fields in each layer are then written as a sum of the local eigenmodes of that particular layer. This leads to a very compact representation of the field and therefore shorter computation times.

One of the mode solvers of the in-house developed software CAMFR [7, 8] is a two-stage eigenmode expansion tool. In the first stage of the method a coarse estimate of the propagation constant is calculated using the Fourier Modal Method (FMM). This method has proven its reliability for modeling grating diffraction problems. It is a frequency-domain method based on computation of the grating modes as eigenvectors and on Fourier expansion of both the permittivity and the electromagnetic fields inside the grating region. With a few minor modifications however, the technique can also be used for modeling aperiodic structures, specifically planar waveguide geometries for TE and TM polarization [9]. In a second stage these estimates are refined by letting them converge to the true solution of a transcendental equation.

The method requires a discrete set of modes, and in order to achieve this, the structure we want to model is typically enclosed in a metal box. Unfortunately, this can create parasitic reflections: radiation that would otherwise travel freely towards infinity is now completely reflected at the metal boundaries, returns to the structure that we want to study, and disturbs the simulation results. To alleviate this problem, advanced boundary conditions can and should be implemented, the most well-known are perfectly matched layers (PML). These layers can be thought of as layers with a real refractive index, but with a complex thickness. This complex thickness provides for reflection less absorption of the incident field, regardless of incidence angle, wavelength or polarization. Using these advanced boundary conditions not only improves the accuracy of the model, but also speeds up the computation time, as the metal walls can now be placed much closer to the structure [10]. The Fourier Modal Method can also incorporate PML boundary conditions.

Due to the combination of Fourier Modal Method estimates and exact propagation constants calculated using a transcendental equation, the accuracy of the FMM can easily be determined by comparing the coarse estimates of the propagation constants to the actual values. For metallic structures, the first step of the algorithm (FMM) is crucial, since finding the zeros in the complex plane of the transcendental equation requires reliable estimates [11]. However, the classical FMM is notoriously slow for thin metallic layers and TM polarization [12, 13]. In this chapter we will focus on increasing the speed, reliability and convergence rate of the Fourier Modal Method for structures containing thin metallic layers.

3.2 Adaptive Spatial Resolution in 1D

The Fourier Modal Method formulated by Moharam and Gaylord [14, 15] is known to be slowly converging, especially for TM polarization of metallic lamellar gratings. The correct factorization rules for discontinuous functions derived by Li [16, 17] improved the convergence rate for TM polarization [12, 13]. For thin metallic gratings however, the convergence of the method, even when applying the correct factorization rules, was still problematic. Reformulating the eigenproblem in a new basis can improve convergence rates, as has been shown by Granet [18], who first introduced the concept of adaptive spatial resolution (ASR). The method consists of a coordinate transformation, which allows a spatially adaptive resolution, increasing the resolution in the neighborhood of the discontinuities of the permittivity function. The original technique was later extended to multilevel profiles [19].

3.2.1 Theory

The Fourier Modal Method is based on a series representation of the electromagnetic field in the region that contains the grating. This series representation must then satisfy the appropriate boundary conditions [14, 13]. Consider the case of a lamellar grating configuration as shown in Fig. 3.1. A piece-wise homogeneous system with thickness h, characterized by a refractive index $n_2(x)$, periodic in the x direction with period d and independent of the y-coordinate separates the vacuum (n=1) and a homogeneous isotropic medium with refractive index n_3. This structure is illuminated by a p-polarized monochromatic plane wave with a vacuum wavelength λ, angular frequency ω and wave number k at normal incidence. The unknown function is the y-component of the magnetic field $H_y(x,z)$. In each region, which we refer to by the subscript i, $i \in \{1,2,3\}$, the field can be represented by the following superposition of eigenmodes:

$$H_{iy}(x,z) = \sum_q \left[A_{iq}^+ e^{-ikr_{iq}z} + A_{iq}^- e^{ikr_{iq}z} \right] \phi_{iq}(x) \tag{3.1}$$

where A_{iq}^+ and A_{iq}^- are modal-field amplitudes and r_{iq} and $\phi_{iq}(x)$ are modal eigenvalues and modal eigenfunctions, which are determined by the boundary-value problem

$$L_i(x)\phi_{iq}(x) = k^2 r_{iq}^2 \phi_{iq}(x) \tag{3.2}$$

$$\phi_{iq}(x+d) = \phi_{iq}(x), \tag{3.3}$$

where $L_i(x)$ is a second order differential operator derived from Maxwell's equations and only depending on the x variable.

$$L_i(x) = n_i^2(x) \left(\frac{\partial}{\partial x} \left(\frac{1}{n_i^2(x)} \frac{\partial}{\partial x} \right) + k^2 \right) \tag{3.4}$$

In the vacuum and the substrate, the H_{iy} functions are given by Rayleigh expansions, i.e. plane waves. The only remaining problem is to compute the eigenmodes $\phi_{2q}(x)$ and the eigenvalues r_{2q}^2 in the grating region. For that purpose we first expand the eigenmodes in terms of plane waves

$$\phi_{2q}(x) = \sum_m \phi_{2mq} e^{-ika_m x}, \tag{3.5}$$

and then project equation 3.2 onto the Fourier basis $e^{-i2\pi m(x/d)}$. We thus obtain a matrix for which the eigenvalues and eigenvectors are being sought. The above described theory is known as the classical Fourier Modal Method.

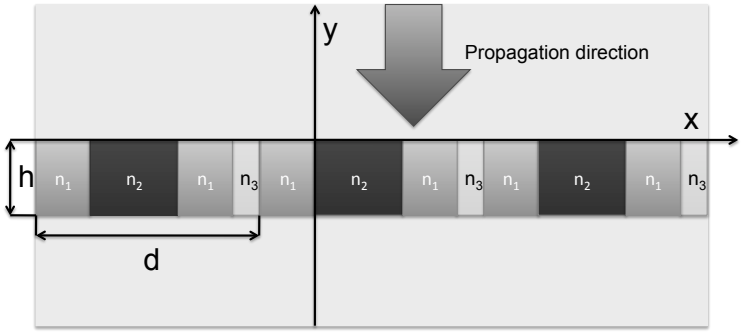


Figure 3.1: Setup for the Fourier Modal Method

3.2.2 Coordinate Stretching Formalism

In order to solve the convergence problems, we first need to take a look at what is causing them. The main reason for slow convergence is that spatial resolution remains uniform within a grating period whatever the permittivity may be [18, 19]. The parametric Fourier Modal Method solves this problem by choosing a coordinate transformation in such a way that the mapping of space fits the variation of the periodic function of interest. In order to do so we need a function $x = x(u)$ so that spatial resolution is increased at the metallic interfaces. Around these points a given variation Δu of u should result in a smaller variation Δx of x. In other words, the coordinate transformation function maps non-uniformly spaced points along the x axis in the physical domain to uniformly spaced points in the transferred u domain. We present here the coordinate x as a function of u and the transition points are denoted by x_l in the x space and by u_l in the u space. Between the transitions x_{l-1} and x_l we use the function $x_l(u)$ for the mapping between different domains. The original formulation of the ASR method, applied to lamellar gratings, was described in Granet *et al* [18, 20] and is given by

$$x_l(u) = u - \frac{\eta a_3}{d} \sin\left(\frac{2\pi(u - x_{l-1})}{a_3}\right), \tag{3.6}$$

with $a_3 = x_l - x_{l-1}$ and d the total thickness of the structure. This coordinate transform and the resulting transformed (real part of the) dielectric profile can be seen in figure 3.2 and 3.3. The metallic layer has a thickness of $40\ nm$ ($\epsilon_r = -19 - 0.5301j$), dielectric cladding layers were $1\ \mu m$ thick ($\epsilon_r = 4$).

Vallius *et al* [19] improved on this original formulation by allowing a different spacing of the transitions in x and u without discontinuities of the

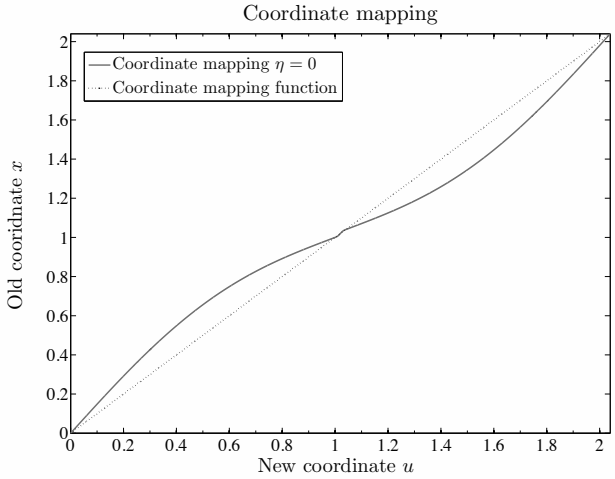


Figure 3.2: Coordinate transformation original adaptive spatial resolution coordinate transform

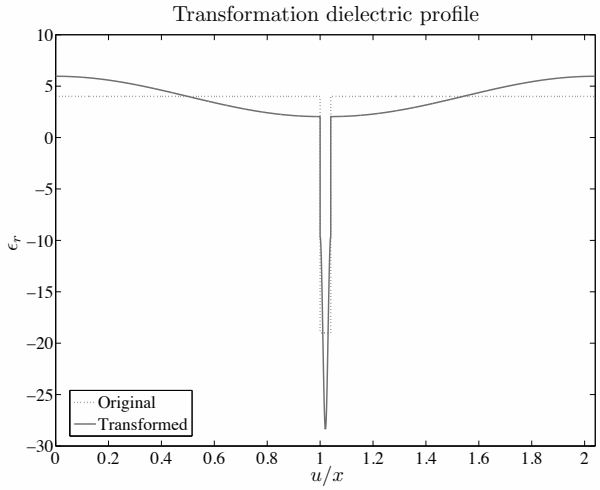


Figure 3.3: Transformation of the dielectric index profile, the original profile corresponds with the x-axis, the transformed profile with the u-axis

coordinate transform function. The formalism was also modified so that the resolution near the discontinuities is increased even more (by omitting the dependence on the total thickness d). The coordinate transform function is given in equation 3.7, the method is illustrated in figure 3.4 and 3.5

$$x_l(u) = a_1 + a_2 u - \frac{n a_3}{2\pi} \sin\left(\frac{2\pi(u - u_{l-1})}{u_l - u_{l-1}}\right), \tag{3.7}$$

with

$$a_1 = \frac{u_l x_{l-1} - u_{l-1} x_l}{u_l - u_{l-1}} \tag{3.8}$$

$$a_2 = \frac{x_l - x_{l-1}}{u_l - u_{l-1}} \tag{3.9}$$

$$a_3 = x_l - x_{l-1} \tag{3.10}$$

This introduces another degree of freedom (besides the value of the numerical parameter η), one can now also choose the position of the discontinuities in the new basis. Two fairly obvious possibilities were investigated in our work. The first possibility is to divide the period d in l equal sections, we shall refer to this method as the uniform solver. The second possibility is to chose the transitions so that the peak height of the transformed dielectric index profile is equal in each section, this will be referred to as the adaptive solver.

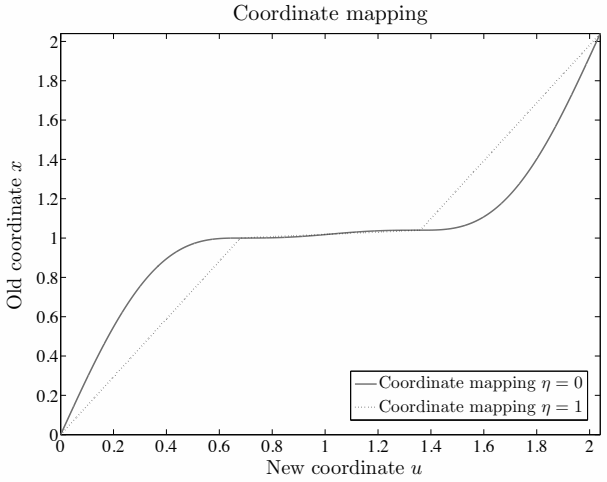


Figure 3.4: Coordinate transformation modified spatial resolution coordinate transform [19]

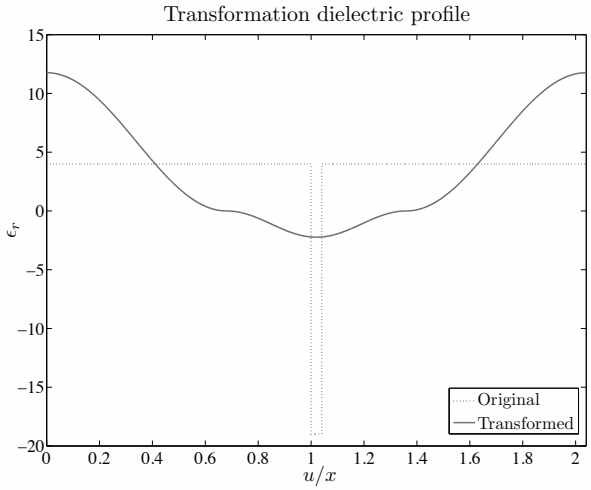


Figure 3.5: Transformation of the dielectric index profile, the original profile corresponds with the x-axis, the transformed profile with the u-axis

Our work takes the adaptive resolution principle even one step further by allowing the period d' in the new basis to be different from the original period d, which allows us to stretch the entire structure. Our method uses the same coordinate transformation as before (equation 3.7), only this time there is no restriction with respect to the maximum value of the new coordinate axis u. The method is illustrated in figure 3.6 and 3.7, where d' is divided in a number of equal sections, this shall be referred to as the extended uniform solver.

The above coordinate transform (equation 3.7) defines a cartesian coordinate system (u, y, z) with a scale factor along the x-axis. We then introduce a new vector basis $\mathbf{e}_u = (dx/du)\mathbf{e}_x$, the other two unit vectors $\mathbf{e}_y$ and $\mathbf{e}_z$ being unchanged. The newly defined function of u is periodic with period d' which can be different than the period of the permittivity function and its derivative is minimal at the discontinuities. It can then be shown that the L_i operator becomes

$$L_i(u) = \frac{1}{b_i(u)} \left(\frac{dx}{du} + \frac{\partial}{\partial u} \left(\frac{1}{a_i(u)} \frac{\partial}{\partial u} \right) \right) \tag{3.11}$$

with

$$b_i(u) = \frac{h(u)}{n_i^2(u)}, \quad a_i(u) = h(u) n_i^2(u), \quad h(u) = \frac{dx}{du} \tag{3.12}$$

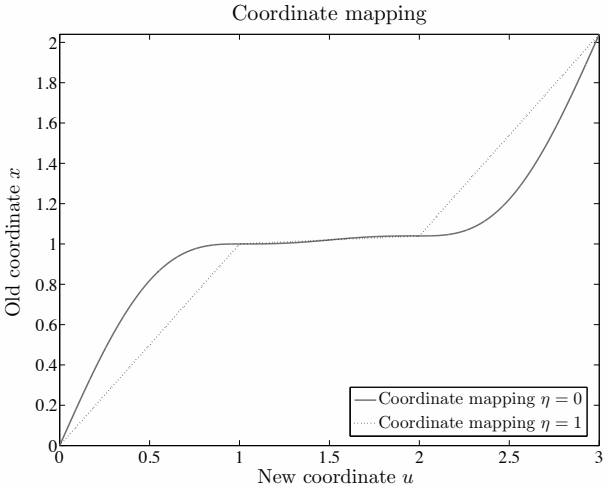


Figure 3.6: Coordinate transformation used in this work

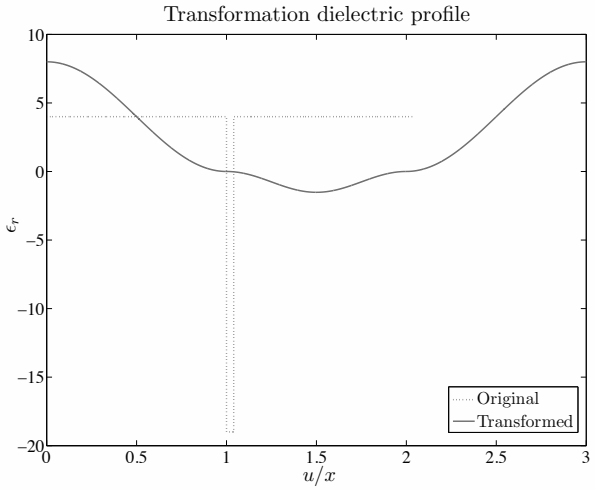


Figure 3.7: Transformation of the dielectric index profile, the original profile corresponds with the x-axis, the transformed profile with the u-axis

We can obtain a matrix relation by following the same way as we did in the classical Fourier Modal Method. The matrix for which the eigenvalues are being sought is given by

$$[\mathbf{L}_i] = [\mathbf{b}_i]^{-1} \left[[\mathbf{h}] - [a][\mathbf{a}_i]^{-1}[a] \right] \tag{3.13}$$

where $[\mathbf{h}]$ is the matrix formed by the Fourier coefficients of h such that

$$h_{mn} = h_{m-n} = \frac{1}{2d} \int_{-d'}^{d'} h(u) e^{-i2\pi(m-n)\frac{u}{d'}} du \tag{3.14}$$

Similar definitions hold for $[\mathbf{a}_i]$ and $[\mathbf{b}_i]$, $[a]$ is the diagonal matrix formed by a_q, defined by

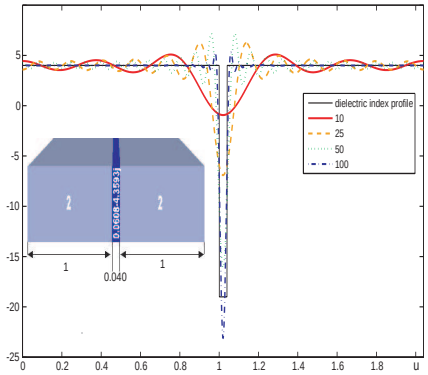
$$a_q = q\frac{\lambda}{d'} \tag{3.15}$$

The main advantage of this change of coordinate is that we no longer need the Fourier coefficients of the exact refractive index profile, but of the transformed refractive index profile $h(u)\epsilon_r(u)$ (as was already depicted in figures 3.2,3.4 and 3.6). The choice of coordinate mapping thus has a direct consequence on the convergence of the Fourier series. As an illustration we depicted the transformed dielectric profile $h(u)\epsilon_r$ for a thin metallic layer ($40\ nm$, $\epsilon_r = -19 - 0.5301j$), surrounded by a dielectric cladding layer ($1\ \mu m$, $\epsilon_r = 4$) and the Fourier approximation truncated after 10 (red), 25 (orange), 50 (green) and 100 (blue) modes. Only the real part of the dielectric function is shown in these figures. One can clearly see that the uniform solver and the extended uniform solver show the most promising convergence because the discontinuities in the refractive index profile have been smoothed out to a large extent.

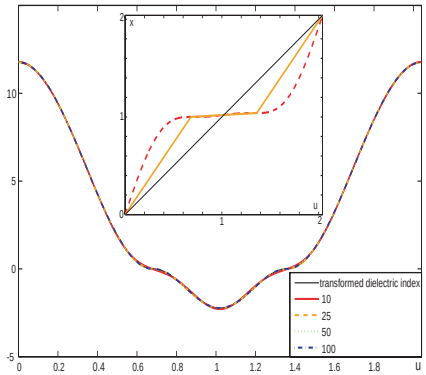
3.2.3 Example and Discussion

We have investigated the convergence properties of the new method and compared them to the original formulation of the FMM and Granet's formulation [18].

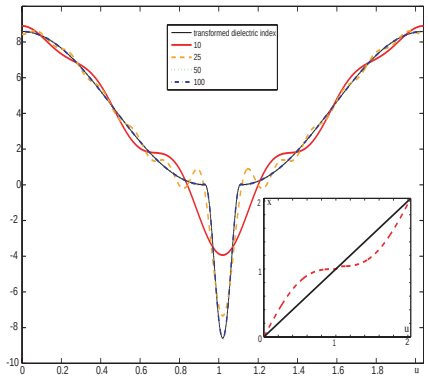
The setup consists of a thin metallic layer (width $40nm$, $n = 0.0608 - 4.3593j$) surrounded by dielectric cladding layers (width $1\mu m$, $n = 2$). If the metallic layer is thin enough, the surface plasmon modes guided by the interfaces become coupled due to field tunneling through the metal, thus creating supermodes. In general, only two purely bound TM modes are guided by a such a structure. In the plane perpendicular to the direction of propagation the electric field of these modes is comprised of a single component, normal to the interface, having either a symmetric or asymmetric



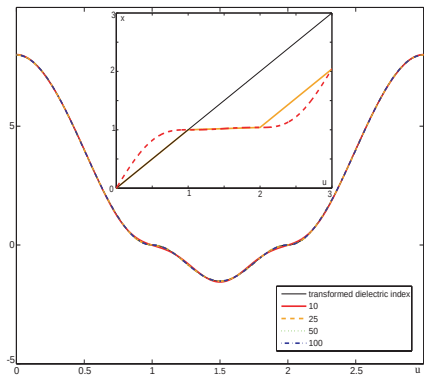
(a) Profile of the relative dielectric constant of a thin metallic layer, the configuration is shown in the inset



(b) Transformed relative dielectric profile of the uniform solver algorithm, the inset depicts the coordinate transformation for this configuration



(c) Transformed relative dielectric profile of the adaptive solver algorithm, the inset depicts the coordinate transformation for this configuration



(d) Transformed relative dielectric profile of the extended uniform algorithm, the inset depicts the coordinate transformation for this configuration

Figure 3.8: Comparison of the different coordinate transformation schemes

distribution across the waveguide. While the symmetric mode extends into the cladding layers, the asymmetric mode is tightly bound to the metallic waveguide and will be extremely lossy. In order to test the convergence of both new methods we made use of the symmetries and placed the entire structure between two magnetic walls, thus selecting the asymmetric plasmon polariton mode. In order to reduce reflections from the top magnetic wall we have also incorporated PML in the simulations implemented as a complex cladding thickness. We calculated the eigenvalues by using a growing number of the truncation order and compared the convergence of the real part of the first eigenvalue. The result of this simulation is plotted in Fig. 3.9. Due to the small feature size the classical Fourier Modal Method converges slowly. Since the old parametric formulation [18] does not allow the spreading of the small features in the u space, the Gibb's phenomenon will prevent small detail to be resolved correctly, which has a negative impact on the convergence of this method. Due to the presence of PML boundary conditions the uniform algorithm fails to converge. This is due to the fact that the algorithm reduces the distance between the guiding metallic layer and the boundaries and thus the PML has a detrimental influence on the convergence of the guided modes. By stretching the entire structure the thickness of the cladding layers can be maintained and the convergence is optimal as can be seen in Fig. 3.9

3.3 Adaptive Spatial Resolution in 2D

Motivated by the drastic increase in convergence rate in one dimension, an obvious next step is to seek a similar increase in performance when we go from one to two dimensions. The Fourier Modal Method can be used in 1D as well as in 2D [21, 22], though complexity of the algorithm increases substantially. The mathematical motivation can be found in Hugonin et al. [23]. The numerical performance however has been plagued by slow convergence performance in the past, mainly due to Gibb's phenomena caused by Fourier expanding discontinuous functions. So if the same phenomenon arises in $2D$ as in $1D$, perhaps adaptive spatial resolution could be the cure?

For the purpose of explaining how Fourier modal methods can be used in 2D mode solvers, consider the waveguide shown in figure 3.10. This structure is invariant in the z-direction perpendicular to the figure. The solution of the problem consists in looking for a modal solution $\phi(x, y)e^{-jk_0 n_{eff} z}$, where k_0 is the free-space wavevector at the wavelength considered and n_{eff} is the effective index.

Analysis of the isolated waveguide structure is performed by introducing an artificial 'periodization' along both the x and the y coordinate. Provided

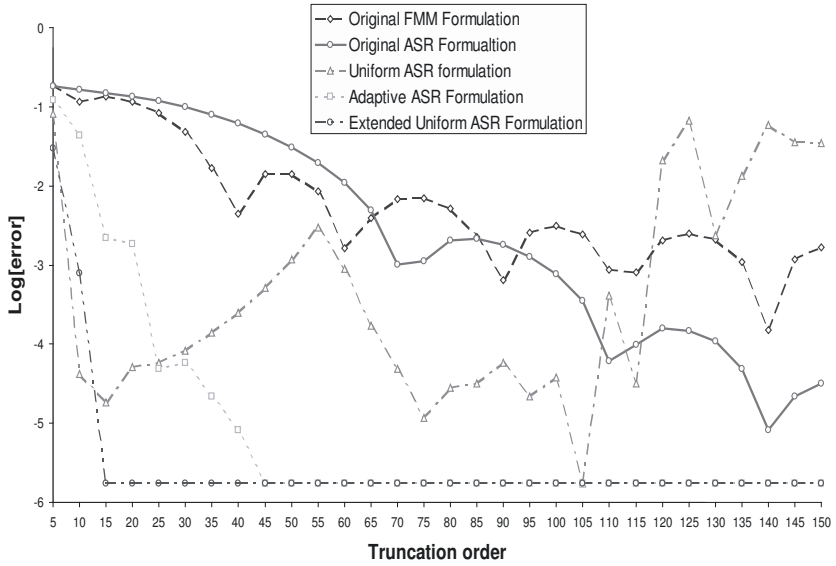


Figure 3.9: Convergence of the first order eigenmode of a metallic waveguide with PML boundary conditions

that the artificial period is large enough, the fields of bounded modes are zero at the boundary of the computational box. The unknown field distribution is then expanded in a two-dimensional Fourier series. This reduces Maxwell's equations to a standard eigenvalue problem whose eigenvalues and eigenvectors yield the modal propagation constants $\beta = k_0 n_{eff}$ and the Fourier expansion coefficients of the modal fields $\phi(x, y)$ [23, 9]. By introducing some absorption (i.e. PML) at the boundaries of the computational box, not only the guided modes, but also the radiation fields in the cladding can be expanded into Fourier series. As a consequence, this method can be used as an eigenmode solver for both guided and radiation modes and in conjunction with an analytic integration in the z-direction by use of S-matrix scheme, can solve many different problems related to the diffraction and propagation of guided waves by intricate non-periodic structures [9, 24, 25]. Mapping is done by transforming the dielectric permittivity ϵ and the magnetic permeability μ, while retaining the form of Maxwell's equations untouched, exploiting the general covariance of electromagnetics. The coordinate system, although appropriate for the geometry of the problem presented here, maps uniformly spaced points along both the x and y direction. By defining a slightly modified coordinate system, we can provide a denser mesh near the discontinuities, thus expressing the permittivity as a smoother function

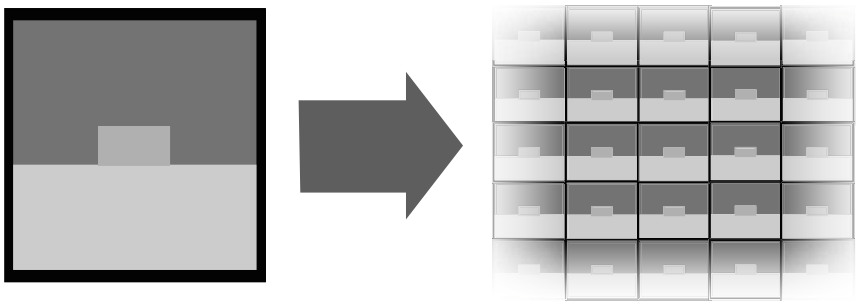


Figure 3.10: Schematic representation of the artificial periodization used in 2D FMM problems. On the left, the wire waveguide geometry 'isolated in space' (thick black lines at the boundaries represent PMLs), on the right the two dimensional periodization used in the Fourier Modal Method to solve this problem

of the new coordinates. Although smoother functions, the new formulation will still contain products of discontinuous functions and one must take special care in obtaining the Fourier coefficients of these functions in terms of the Fourier components of each individual function.

3.3.1 Theory

Electromagnetic boundary conditions dictate that the tangential component of $\mathbf{E}$ and $\mathbf{H}$ are continuous, whereas for the $\mathbf{B}$ and $\mathbf{D}$ fields the normal component is continuous. It is therefore natural to distinguish between contravariant and covariant components of the electromagnetic field. One has to keep in mind the basic properties of vector components in tensor theory: the covariant components of a vector are the perpendicular projections along the covariant basis vectors, while the contravariant components are the perpendicular projections along the contravariant base vectors. By using the Einstein convention we can write Maxwell's equations as [26]

$$\varepsilon^{ijk}\partial_j E_k = -j\omega B^j \tag{3.16}$$

$$\varepsilon^{ijk}\partial_j H_k = +j\omega D^j \tag{3.17}$$

where ∂_i represents the partial derivative $\partial/\partial x^i$ versus x^i, ε^{ijk} is the Levi-Civita symbol. The benefit of this notation lies in the fact that these equations do not change form under arbitrary reversible transformations from and to Cartesian coordinates. The constitutive relations should also be written in

this frame-independent way as

$$D^i = \epsilon^{ij} E_j \quad \text{where} \quad \epsilon^{ij} = \epsilon\sqrt{g}g^{ij} \tag{3.18}$$

$$B^i = \mu^{ij} H_j \quad \text{where} \quad \mu^{ij} = \mu\sqrt{g}g^{ij} \tag{3.19}$$

$$\tag{3.20}$$

Formulating Maxwell's equations in a generally covariant manner has the advantage that the equations do not change form under arbitrary reversible coordinate transformation, since all metric information is contained in the permeability ϵ and the permeability μ, which are functions of the metric tensor. The metric tensor for the earlier described coordinate transformation can be written down as

$$\left[g_{ij}\right] = \begin{bmatrix} -(\partial_u f)^2 & 0 & 0 \\ 0 & (\partial_v g)^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.21}$$

where f and g are the functions that map the new coordinate system on the old coordinate system according to

$$x = f(u) \qquad y = g(v) \qquad z = z \tag{3.22}$$

The contravariant components of the metric tensor can then be calculated by inverting the previous matrix

$$\left[g^{ij}\right] = \frac{1}{\sqrt{g}} \begin{bmatrix} \partial_v g/\partial_u f & 0 & 0 \\ 0 & \partial_u f/\partial_v g & 0 \\ 0 & 0 & \partial_u f \partial_v g \end{bmatrix} \tag{3.23}$$

3.3.2 Li's rules of factorization

In the here considered periodical structure the electromagnetic field components will be expanded in Floquet harmonics, the permittivity and permeability (if required) will be written as a Fourier series, therefore the correct factorization rules for discontinuous functions are of the utmost importance. Suppose that f and g are two periodic functions of x (no connection with the earlier defined functions $f(u)$ and $g(v)$), and h is their product defined by $h(x) = f(x)g(x)$. We would now like to express the Fourier coefficients of the function h as a function of the Fourier coefficients of the functions f and g.

- If the functions f and g have no concurrent discontinuities, then the Fourier coefficients of f can be found through the so-called Laurent-rule: $[h_m] = [f_{m-n}][g_n]$ where $[f_{m-n}]$ is a Toeplitz matrix where the (m, n)th element is the Fourier coefficient f_{m-n}

- If f and g have pairwise-complimentary discontinuities, while h is continuous, then $[f_{m-n}]$ must be replaced by the inverse matrix $[(1/f)_{m-n}]^{-1}$; where $1/f$ is the reciprocal function of f. This way the continuity of the product is preserved at a point of complementary discontinuity.

The factorization rules for periodic functions of two variables have, to the best of our knowledge, not yet been mathematically proven [26]. Following Li [16, 27] , we assume that the same rules must be applied to the functions f,g and h as if they were functions of two independent variables x and y. Suppose that f is a function of x and y, then the following notation will be used throughout this paper:

1. $[[f]_x]_y$: Laurent's rule with respect to x and y
2. $\left[[1/f]_x^{-1}\right]_y$: Inverse rule with respect to x, Laurent's rule with respect to y
3. $\left[[1/f]_y^{-1}\right]_x$: Inverse rule with respect to y, Laurent's rule with respect to x

Fourier transformation of the constitutive relations 3.18,we get

$$D^u_{mn} = \sum_{j,l} \left[\left[\frac{1}{\epsilon} \frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]_{v,nl} E_{u,jl} \tag{3.24}$$

Contravariant electromagnetic vector components are continuous along their tangential directions and discontinuous according to their normal direction, covariant electromagnetic vector components display an inverse behavior. So D^u is continuous along u, discontinuous along v, E_u is continuous along v and discontinuous along u, while the function $\epsilon\partial_v g/\partial_u f$ is discontinuous in both the u and v direction. Because of this behavior we first use the inverse rule to find the Fourier components of $\epsilon\partial_v g/\partial_u f$ along u, and then use the Laurent's rule to find the Fourier components along v. The same line of thought can be applied to all the other components of D^v, B^u,B^v,E_v,H_u and H_v. E_z and H_z both vary continuously along the u and v directions, while D^z and B^z exhibit discontinuities along both directions, so for these equations Laurent's factorization rule can be applied.

$$D^v_{mn} = \sum_{j,l} \left[\left[\frac{1}{\epsilon} \frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,mj} \right]_{u,nl} E_{v,jl} \tag{3.25}$$

$$D^z_{mn} = \sum_{j,l} \Big[[\epsilon \partial_u f(u) \partial_v g(v)]_{u,mj} \Big]_{v,nl} E_{z,jl} \tag{3.26}$$

$$B^u_{mn} = \sum_{j,l} \left[\left[\frac{1}{\mu_0} \frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]_{v,nl} H_{u,jl} \tag{3.27}$$

$$B^v_{mn} = \sum_{j,l} \left[\left[\frac{1}{\mu_0} \frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,mj} \right]_{u,nl} H_{v,jl} \tag{3.28}$$

$$B^z_{mn} = \sum_{j,l} \Big[[\mu \partial_u f(u) \partial_v g(v)]_{u,mj} \Big]_{v,nl} H_{z,jl} \tag{3.29}$$

Important to note here is that the order of Fourier factorization is not unique, one can just as easily use the Fourier factorization

$$D^u_{mn} = \sum_{j,l} \left[\left[\epsilon \frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,mj} \right]^{-1}_{u,nl} E_{u,jl} \tag{3.30}$$

$$D^v_{mn} = \sum_{j,l} \left[\left[\epsilon \frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]^{-1}_{v,nl} E_{v,jl} \tag{3.31}$$

$$D^z_{mn} = \sum_{j,l} \Big[[\epsilon \partial_u f(u) \partial_v g(v)]_{v,mj} \Big]_{u,nl} E_{z,jl} \tag{3.32}$$

$$B^u_{mn} = \sum_{j,l} \left[\left[\mu_0 \frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,mj} \right]^{-1}_{u,nl} H_{u,jl} \tag{3.33}$$

$$B^v_{mn} = \sum_{j,l} \left[\left[\mu_0 \frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]^{-1}_{v,nl} H_{u,jl} \tag{3.34}$$

$$B^z_{mn} = \sum_{j,l} \Big[[\mu_0 \partial_u f(u) \partial_v g(v)]_{v,mj} \Big]_{u,nl} H_{z,jl} \tag{3.35}$$

This means that the Fourier factorization of a doubly periodic function is not unique in truncated Fourier space. The numerical investigation of the convergence behavior of the possible cases will be presented later on.

We have now transformed the relation between the contravariant components of $\mathbf{D}$ and $\mathbf{B}$ and the covariant components of $\mathbf{E}$ and $\mathbf{H}$ from real space to Fourier space in such a way that the electromagnetic boundary

conditions across the permittivity discontinuities are best preserved. We will now Fourier analyze the equations 3.16. To do so, we expand the fields into Floquet-Fourier series, that is, series with $e^{j\alpha_n u}$ as basis functions in the u-direction and $e^{j\beta_n v}$ in the v-direction. Practically this means that the operator ∂_u (∂_v) is equivalent to (left) multiplication by a diagonal matrix whose elements are $j\alpha_n$ ($j\beta_n$). Keeping the suppressed z dependence in mind, the operator ∂_z can be replaced by e^{jk_z}. By writing the E_z and H_z components as a function of the E_u, E_v, H_u and H_v components, and substituting these expressions in the 4 remaining equations we can find the following 4 equations

$$\omega\epsilon_0 k_z E_{u,mn} = k_0^2 \sum_{jl} \left[\left[\frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,nl} \right]_{u,mj} H_{v,jl} + \alpha_m \sum_{jl} \left[\left[\epsilon_r \partial_u f(u) \partial_v g(v) \right]_{u,mj} \right]^{-1}_{v,nl} \left[-\alpha_j H_{v,jl} + \beta_l H_{u,jl} \right] \tag{3.36}$$

$$\omega\epsilon_0 k_z E_{v,mn} = -k_0^2 \sum_{jl} \left[\left[\frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]_{v,nl} H_{u,jl} + \beta_n \sum_{jl} \left[\left[\epsilon_r \partial_u f(u) \partial_v g(v) \right]_{u,mj} \right]^{-1}_{v,nl} \left[-\alpha_j H_{v,jl} + \beta_l H_{u,jl} \right] \tag{3.37}$$

$$\omega\mu_0 k_z H_{u,mn} = -k_0^2 \sum_{jl} \left[\left[\frac{1}{\epsilon_r} \frac{\partial_v g(v)}{\partial_u f(u)} \right]^{-1}_{v,nl} \right]_{u,mj} E_{v,jl} - \alpha_m \sum_{jl} \left[\left[\epsilon_r \partial_u f(u) \partial_v g(v) \right]_{u,mj} \right]^{-1}_{v,nl} \left[-\alpha_j E_{v,jl} + \beta_l E_{u,jl} \right] \tag{3.38}$$

$$\omega\mu_0 k_z H_{v,mn} = k_0^2 \sum_{jl} \left[\left[\frac{1}{\epsilon_r} \frac{\partial_u f(u)}{\partial_v g(v)} \right]^{-1}_{u,mj} \right]_{v,nl} E_{u,jl} - \beta_n \sum_{jl} \left[\left[\epsilon_r \partial_u f(u) \partial_v g(v) \right]_{u,mj} \right]^{-1}_{v,nl} \left[-\alpha_j E_{v,jl} + \beta_l E_{u,jl} \right] \tag{3.39}$$

We can write this as the following two matrices

$$F = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad G = \begin{bmatrix} E & F \\ G & H \end{bmatrix} \tag{3.40}$$

where,

$$A = [\alpha]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.41}$$

$$B = k_0^2\left[\left[\frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u - [\alpha]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.42}$$

$$C = -k_0^2\left[\left[\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v + [\beta]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.43}$$

$$D = -[\beta]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.44}$$

$$E = -[\alpha]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.45}$$

$$F = -k_0^2\left[\left[\frac{1}{\epsilon_r}\frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u + [\alpha]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.46}$$

$$G = k_0^2\left[\left[\frac{1}{\epsilon_r}\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v - [\beta]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.47}$$

$$H = [\beta]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.48}$$

Since the calculation order of Fourier elements is to some extent arbitrary, one can also use the following matrix elements in the matrices F and G

$$A = [\alpha]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.49}$$

$$B = k_0^2\left[\left[\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v^{-1} - [\alpha]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.50}$$

$$C = -k_0^2\left[\left[\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v + [\beta]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.51}$$

$$D = -[\beta]\left[[\epsilon_r \partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.52}$$

$$E = -[\alpha]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.53}$$

$$F = -k_0^2\left[\left[\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v^{-1} + [\alpha]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.54}$$

$$G = k_0^2\left[\left[\frac{1}{\epsilon_r}\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v^{-1} - [\beta]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\beta] \tag{3.55}$$

$$H = [\beta]\left[[\partial_u f(u)\partial_v g(v)]_u\right]_v^{-1}[\alpha] \tag{3.56}$$

these two matrices consist of Fourier elements that were calculated the same way for each individual submatrix. The advantages of the latter

approach compared to the former will become apparent in the following sections.

One can now write Maxwell's equations as

$$k_0^2 k_z \begin{bmatrix} E_u \\ E_v \end{bmatrix} = [F]\,[G] \begin{bmatrix} E_u \\ E_v \end{bmatrix} \tag{3.57}$$

The above presented method resulted in a matrix eigenvalue problem of infinite order in Fourier space. A practical approach toward solving this problem is to solve the associated problem of finite order obtained by truncating the infinite Fourier matrices, provided of course that the approximate solutions tend to the exact solutions as the Fourier order is increased [28]. Mathematically, this method is called the reduction method. If the truncation is described as $-M \leq m \leq M$ and $-N \leq n \leq N$, where both M and N are integer, the matrix $[F]\,[G]$ from which eigenvalues and eigenvectors need be calculated is of rank $2(2M+1)(2N+1)$. For simplicity reasons (and because it's an obvious choice) we take $M = N$, and denote this integer the truncation order.

3.3.3 Coordinate stretching formalism

We use the same coordinate stretching functions as before for the u direction, i.e. equation 3.7. Similar equations can be written down for the coordinate stretching function in the v direction. The above relation defines a Cartesian coordinate system (u, v, z) with a scale factor along the x-axis and y-axis. We then introduce a new vector basis $\mathbf{e}_u = (dx/du)\mathbf{e}_x$ and $\mathbf{e}_v = (dy/dv)\mathbf{e}_x$, the third basis vector remains unchanged. The newly defined functions of u and v are periodic with period d' and d'' which can be different than the period of the permittivity function and its derivative is minimal at the discontinuities.

3.3.4 Integration sequence

Since the Fourier decomposition of a double-periodic function in 2D-Fourier space is in no respect unique, we can study the dependence of the convergence rate of the Fourier-Modal method on the different integration orders. In figure 3.11 we have plotted the different integration orders, since there are 4 different matrices which we need to calculate, and each can be calculated in two different ways, there are 4^2 possible combinations. If we describe the two integration order possibilities and 0 and 1, the binary number 0011 e.g. represent one single integration sequence. This is how the numbers in figure 3.11 correspond to the integration sequence of the different matrix

	A	B	C	D
1	$\left[\left[\frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u^{-1}$	$\left[\left[\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v^{-1}$	$\left[\left[\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u^{-1}$	$\left[\left[\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v^{-1}$
0	$\left[\left[\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v$	$\left[\left[\frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u$	$\left[\left[\frac{1}{\epsilon_r}\frac{\partial_u f(u)}{\partial_v g(v)}\right]_u^{-1}\right]_v$	$\left[\left[\frac{1}{\epsilon_r}\frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u$

Table 3.1: Illustration of the different possibilities based on changing the integration sequence.

elements. The number is the decimal value of the binary number $ABCD$, where A, B, C, D correspond to the matrix elements given in table 3.1

From figure 3.11 we can see that these 16 different combinations can be reduced to 4 distinct combinations, apparently convergence behavior is dominantly influenced by the matrix elements of $\frac{1}{\epsilon_r}\frac{\partial_u f(u)}{\partial_v g(v)}$ and $\frac{1}{\epsilon_r}\frac{\partial_v g(v)}{\partial_u f(u)}$. Using these results we have decided to use the matrices F and G in the form of equations 3.49-3.56.

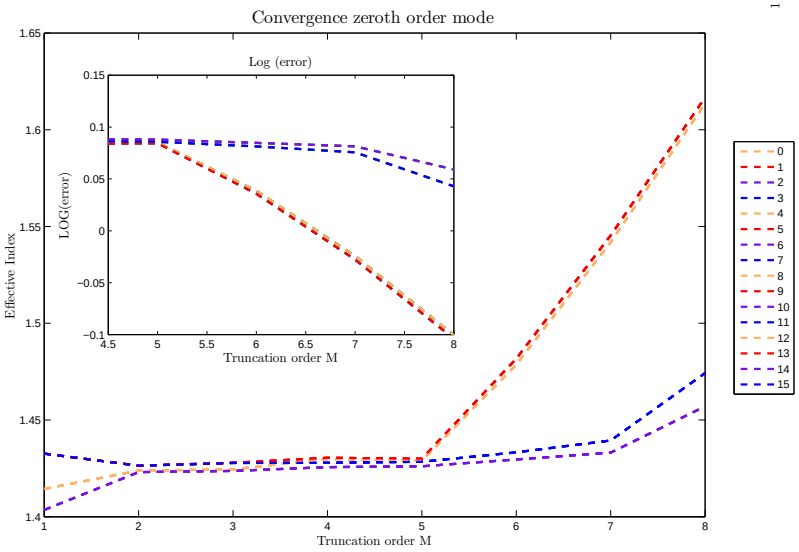


Figure 3.11: Dependence of convergence rate of the FMM on the integration order

3.3.5 Example and Discussion

The adaptive spatial resolution algorithm requires the approximation of three different transformed dielectric index profiles, being $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$, $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$. This condition contains some severe restrictions for

the coordinate stretching functions $f(u)$ and $g(v)$, since both the original stretching function and the reciprocal stretching function need to be piecewise continuous and never tend to infinity at any point. This means that the numerical parameter η should be limited to the interval $0 \leq \eta < 1$, since for $\eta = 1$, the function $\partial_u f(u)$ and the function $\partial_v g(v)$ become 0 at the discontinuity points, at which point their reciprocal functions tend to infinity.

Several new schemes can be devised for choosing the discontinuities in the new uv basis, however we will limit ourselves to three fairly straightforward choices. The first method uses the same discontinuity points in the new uv-basis, the second method uses uniformly spaced discontinuities both along the u and the v axis, the third method also uses uniformly spaced discontinuities but also stretches the entire structure in both the u and v direction. Other, less obvious choices are possible, but have not been investigated here.

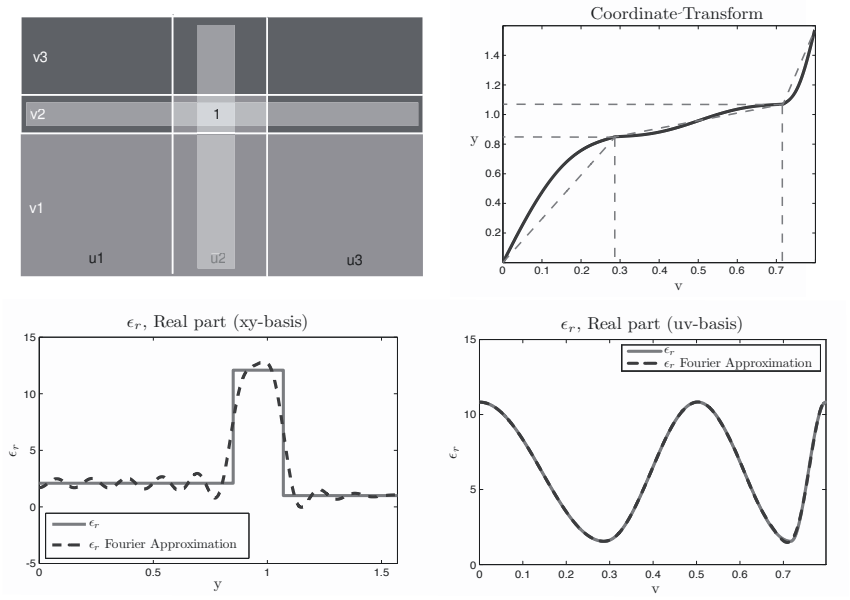


Figure 3.12: Schematical representation of the dielectric permittivity profile of the numerical example under consideration Top left: vertical slices $u1$, $u2$ and $u3$,horizontal slices $v1$,$v2$,$v3$. Top right: Coordinate transform $y(v)$ used in method 1. Bottom left: ϵ_r (slice $u2$) and Fourier approximation ($M = 10$) Bottom left: $\epsilon_r \partial_v g(v)$ (slice $u2$) and Fourier approximation

Figure 3.13a shows the original dielectric index profile of the numerical example considered here. The example consists of a high index contrast single mode waveguide ($n = 3.47$), on top of a lower index substrate layer

($n = 1.45$), the top cladding is air ($n = 1$). Figures 3.13a, 3.13c and 3.13e represent the three different dielectric permittivity functions that need to be Fourier expanded, figures 3.13b, 3.13d and 3.13f depict the three different Fourier decompositions (for these plots the numerical parameter η was chosen to be 0, so no adaptive resolution).

Let us first consider what happens with the dielectric permittivity profile if we apply the coordinate transform given by method 1. Figures 3.14a, 3.14c and 3.14e depict the three elements of the relative dielectric permittivity tensor $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$, figures 3.14b, 3.14d and 3.14f depict their respective Fourier approximation, the integration sequence is indicated in the figures. The numerical factor governing the adaptive spatial resolution algorithm η is chosen to be 0.75 in plotting these figures. One can clearly see that although the function $\epsilon_r \partial_u f(u) \partial_v g(v)$ has been transformed from a difficult to expand discontinuous function into an easier to approximate function, the same does not hold for the functions $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$. In fact, convergence of the Fourier approximation to the actual values is far worse.

Convergence of the zeroth order mode is given in figure 3.15. One can see that the adaptive spatial resolution algorithm does not converge substantially faster than the original algorithm.

Now consider what happens when we use method 2. Again, the functions $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$ and their respective Fourier approximations are plotted in figures 3.16a,3.16c, 3.16e and 3.16b, 3.16d and 3.16f.

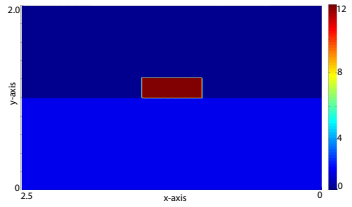
Convergence behavior for this method is plotted in figure 3.17

For method 3, the dielectric permittivity functions and their respective Fourier approximations are plotted in figures 3.18a,3.18c, 3.18e and 3.18b, 3.18d and 3.18f

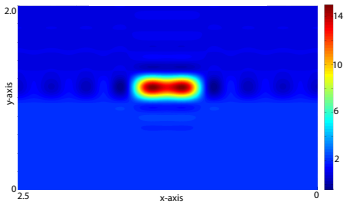
The convergence rate for this method is given in figure 3.19. Again one can see that the convergence rate is not substantially better for the adaptive spatial resolution method than for the original method

3.3.6 Conclusion

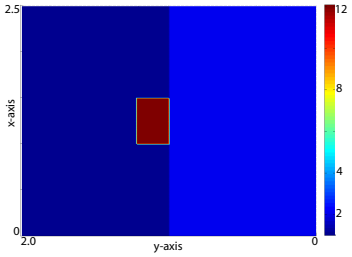
While the ASR method certainly has its merits for $1D$ structures, improving the convergence for $2D$ problems seems difficult to say the least. The biggest problem with the adaptive spatial resolution method is that not only $\epsilon_r \partial_u f(u) \partial_v g(v)$ needs to be Fourier approximated, but also the functions $\epsilon_r \partial_u f(u) / \partial_v g(v)$ and $\epsilon_r \partial_v g(v) / \partial_u f(u)$ (not to mention the Fourier approximation of the functions $\partial_u f(u) \partial_v g(v)$,$\partial_u f(u) / \partial_v g(v)$ and $\partial_v g(v) / \partial_u f(u)$).



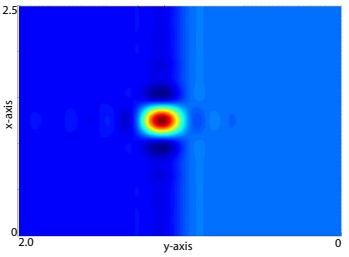
(a) ϵ_r in the xy basis.



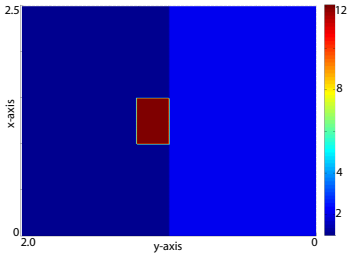
(b) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r\right]_y\right]_x$.



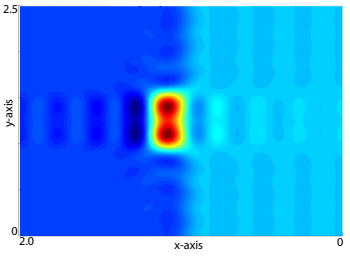
(c) ϵ_r in the yx basis.



(d) Fourier approximation based on the matrix elements of $\left[\left[\frac{1}{\epsilon_r}\right]_y^{-1}\right]_x$.

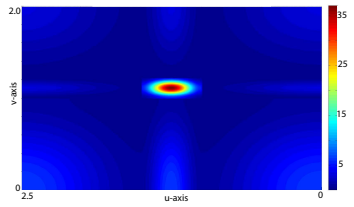


(e) ϵ_r in the yx basis.

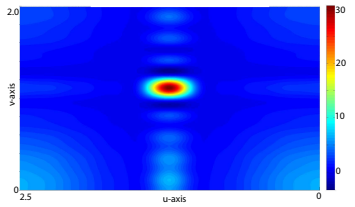


(f) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r\right]_y^{-1}\right]_x^{-1}$.

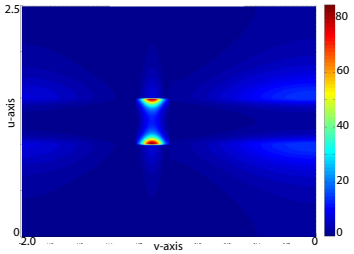
Figure 3.13: The function $epsilon_r$ and the three different Fourier approximations, $\eta = 0$, $M = N = 8$



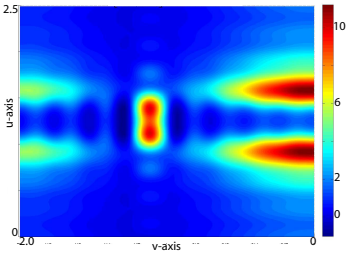
(a) $\epsilon_r \partial_u f(u) \partial_v g(v)$, uv basis, method 1.



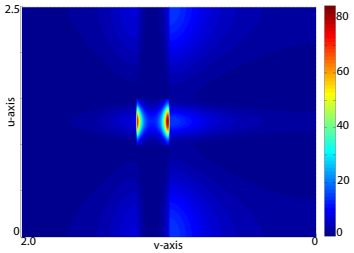
(b) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r \partial_u f(u) \partial_v g(v)\right]_v\right]_u$.



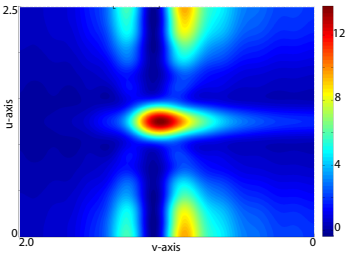
(c) $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$, uv basis, method 1.



(d) Fourier approximation based on the matrix elements of $\left[\left[\frac{1}{\epsilon_r} \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u$.



(e) $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$, uv basis, method 1.



(f) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u^{-1}$.

Figure 3.14: The functions $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$ and their respective Fourier approximations, $\eta = 0.75$, $M = N = 8$, method 1

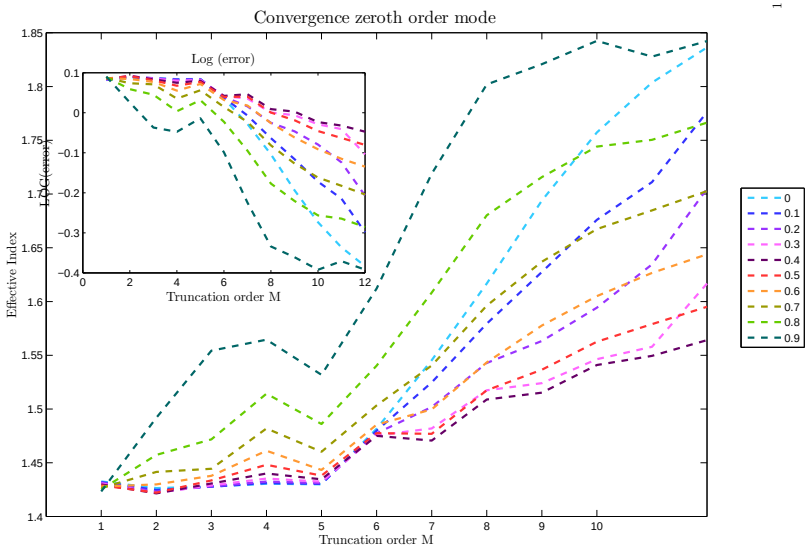
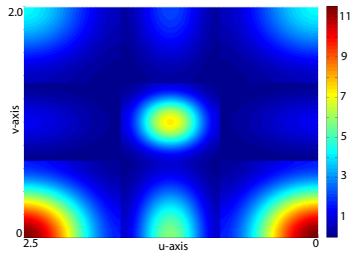


Figure 3.15: Convergence results adaptive spatial resolution algorithm with original discontinuities.

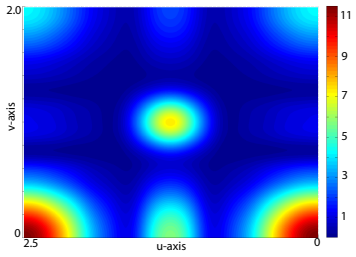
By making the numerical parameter η larger and larger we can make one of these functions substantially smoother, however, we can only do this for one of these functions. The two other functions usually become even more discontinuous than they already were. Convergence of the calculated eigenvalues does not improve noticeably for the three choices of discontinuities presented here. Although this is somewhat discouraging, we can note that we have in no case explored all the different possibilities here. Perhaps there is a specific stretching function that can smoothen all the required functions here, however, lack of time prevented us from looking into this in intricate detail.

3.4 Implementation

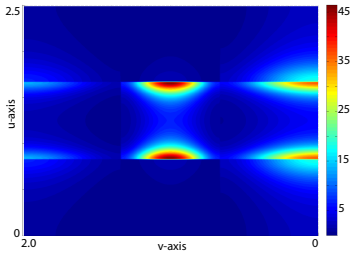
All described algorithms described here were directly implemented into CAMFR, which is freely available. The 2D version of the adaptive spatial algorithm however was implemented both in CAMFR and MATLAB, but where never made publicly available due to the convergence problems listed above. All simulation results in the rest of this work were obtained using CAMFR and the adaptive spatial resolution algorithm unless stated otherwise.



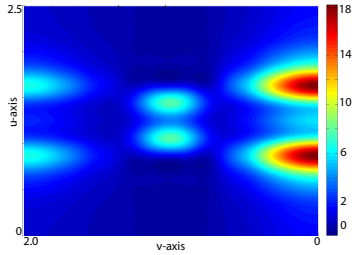
(a) $\epsilon_r \partial_u f(u) \partial_v g(v)$, uv basis, method 2.



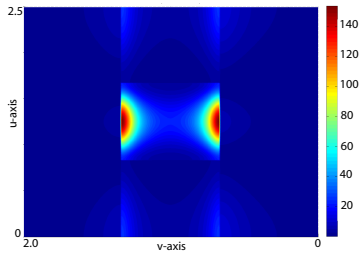
(b) Fourier Approximation based on the matrix elements of $\left[\left[\epsilon_r \partial_u f(u) \partial_v g(v)\right]_v\right]_u$.



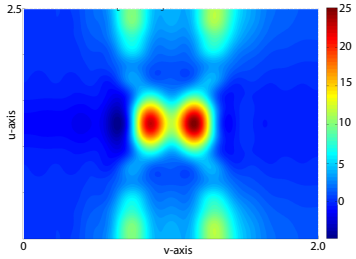
(c) $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$, uv basis, method 2.



(d) Fourier approximation based on the matrix elements of $\left[\left[\frac{1}{\epsilon_r} \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u$.



(e) $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$, uv basis.



(f) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u^{-1}$.

Figure 3.16: The functions $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$ and their respective Fourier approximations, $\eta = 0.75$, $M = N = 8$, method 2

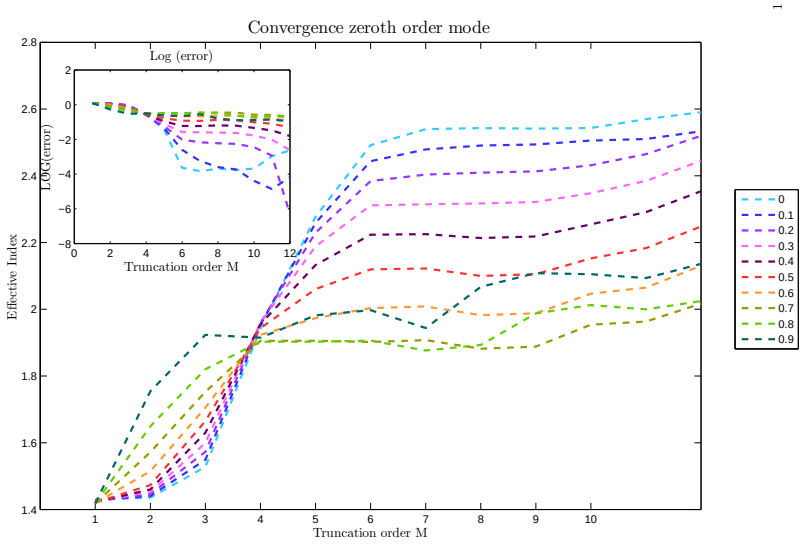
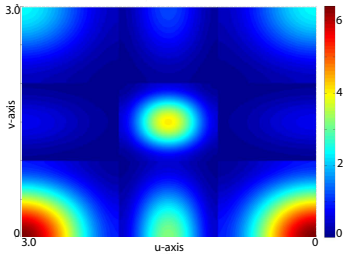
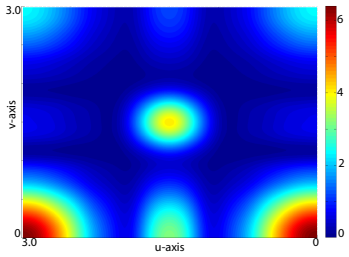


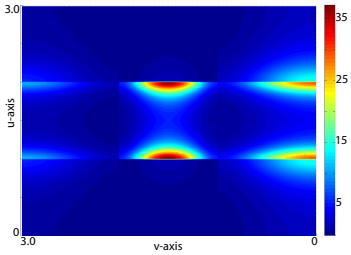
Figure 3.17: Convergence result adaptive spatial resolution algorithm, uniform discontinuities.



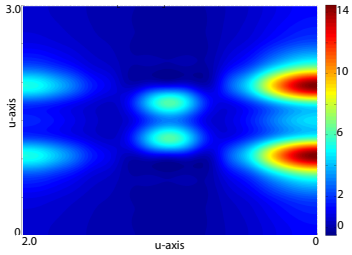
(a) $\epsilon_r \partial_u f(u) \partial_v g(v)$, uv basis, method 3.



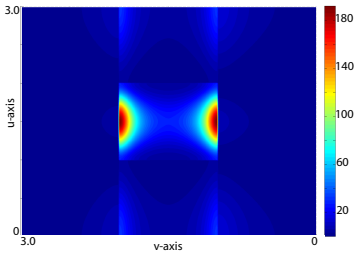
(b) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r \partial_u f(u) \partial_v g(v)\right]_v\right]_u$.



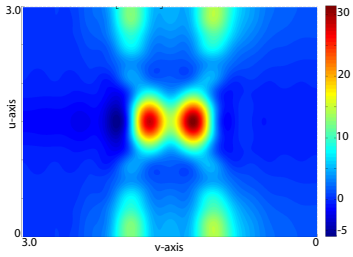
(c) $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$, uv basis, method 3.



(d) Fourier approximation based on the matrix elements of $\left[\left[\frac{1}{\epsilon_r} \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u$.



(e) $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$, uv basis, method 3.



(f) Fourier approximation based on the matrix elements of $\left[\left[\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}\right]_v^{-1}\right]_u^{-1}$.

Figure 3.18: The functions $\epsilon_r \partial_u f(u) \partial_v g(v)$, $\epsilon_r \frac{\partial_u f(u)}{\partial_v g(v)}$ and $\epsilon_r \frac{\partial_v g(v)}{\partial_u f(u)}$ and their respective Fourier approximations, $\eta = 0.75$, $M = N = 8$, method 3

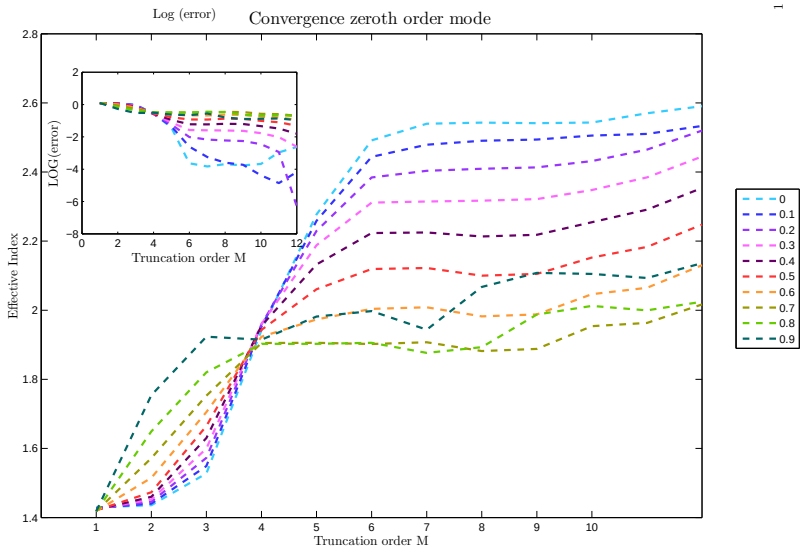


Figure 3.19: Convergence results adaptive resolution algorithm, uniform discontinuities, stretched structure.

References

[1] A. Taflove. *Computational Electrodynamics*. Artech House, Boston, 1995.

[2] W. H. P. Pernice, F. P. Payne, and D. F. G. Gallagher. *An FDTD method for the simulation of dispersive metallic structures*. Optical and Quantum Electronics, 38:843–856, 2006.

[3] W. L. Barnes, A. Dereux, and T. W. Ebbesen. *Surface plasmon subwavelength optics*. Nature, 424:824–830, 2003.

[4] Georgios Veronis and Shanhui Fan. *Bends and splitters in metal-dielectric-metal subwavelength plasmonic waveguides*. Applied Physics Letters, 87(13):131102–3, 2005.

[5] Kristy C. Vernon, Dmitri K. Gramotnev, and D. F. P. Pile. *Channel plasmon-polariton modes in V grooves filled with dielectric*. Journal of Applied Physics, 103(3):034304–6, 2008.

[6] M. L. Brongersma and P. G. Kik. *Surface Plasmon Nanophotonics*, volume 131 of *Springer Series in Optical Sciences*. Springer, 2007.

[7] P. Bienstman. *CAMFR* `http://sourceforge.org/projects/CAMFR`, 2008.

[8] P. Bienstman. *Two-stage mode finder for waveguides with a 2D cross-section*. Optical and Quantum Electronics, 36(1/3):5–14, 2004.

[9] P. Lalanne and E. Silberstein. *Fourier-modal methods applied to waveguide computational problems*. Optics Letters, 25(15):1092–1094, 2000.

[10] J P Hugonin and P. Lalanne. *Perfectly matched layers as nonlinear coordinate transforms: a generalized formulation*. J. Opt. Soc. Am. A, 22(9):1844–1849, 2005.

[11] P. Bienstman. *Rigorous and efficient modelling of wavelength scale photonic components*. PhD thesis, Ghent University, 2001.

[12] Philippe Lalanne and G Michael Morris. *Highly improved convergence of the coupled-wave method for TM polarization*. J. Opt. Soc. Am. A, 13(4):779 – 784, 1996.

[13] G Granet and B Guizal. *Efficient implementation of the coupled-wave method for metallic lamellar gratings in TM polarization*. J. Opt. Soc. Am. A, 13(5):1019–1023, 1996.

[14] M. G. Moharam and T. K. Gaylord. *Rigorous coupled-wave analysis of grating diffraction – E mode polarization and losses*. J. Opt. Soc. Am. A, 73:451 – 455, 1983.

[15] M G Moharam, Drew A Pommet, Eric B Grann, and T K Gaylord. *Stable Implementation of the Rigorous Coupled-Wave Analysis for Surface-relief Gratings: Enhanced Transmittance Matrix Approach*. J. Opt. Soc. Am. A, 12(5):1077–1086, 1995.

[16] Lifeng Li. *Use of Fourier series in the analysis of discontinuous periodic structures*. J. Opt. Soc. Am. A, 13(9):1870–1876, 1996.

[17] Lifeng Li. *Convergence of electromagnetic field components across discontinuous permittivity profiles : Comment*. JOSA Communications, 19(7):1443–1444, 2002.

[18] G Granet. *Reformulation of the lamellar grating problem through the concept of adaptive spatial resolution*. J. Opt. Soc. Am. A, 16(10):2510–2516, 1999.

[19] T Vallius and M Honkanen. *Reformulation of the Fourier Modal Method with adaptive spatial resolution: Application to multilevel profile*. Optics Express, 10(1):24–34, 2002.

[20] G Granet, J. Chandezon, J. P. Plumey, and K. Raniriharinosy. *Reformulation of the coordinate transformation method through the concept of adaptive spatial resolution. Application to trapezoidal gratings*. J. Opt. Soc. Am. A, 18(9):2102–2108, 2001.

[21] Eero Noponen and Jari Turunen. *Eigenmode method for electromagnetic synthesis of diffractive elements with three dimensional profiles*. J. Opt. Soc. Am. A, 11(9):2494–2502, 1994.

[22] Lifeng Li. *New formulation of the Fourier modal method for crossed surface-relief gratings*. J. Opt. Soc. Am. A, 14(10):2758–2767, 1997.

[23] J. P. Hugonin, P. Lalanne, I. Del Villar, and I. R. Matias. *Fourier modal methods for modeling optical dielectric waveguides*. Optical and Quantum Electronics, 37(1-3):107–119(13), 2005.

[24] E. Silberstein, P. Lalanne, J. P. Hugonin, and Q. Cao. *Use of grating theories in integrated optics*. Journal of the Optical Society of America a-Optics Image Science and Vision, 18(11):2865–2875, 2001.

[25] J. Tervo, M. Kuittinen, P. Vahimaa, J. Turunen, T. Aalto, P. Heimala, and M. Leppihalme. *Efficient Bragg waveguide-grating analysis by quasi-rigorous approach based on Redheffer's star product*. Optics Communications, 198(4-6):265–272, 2001.

[26] G Granet and J. P. Plumey. *Parametric formulation of the Fourier modal method for crossed surface-relief gratings*. J. Opt. A: Pure Appl. Opt., 4:S145–S149, 2002.

[27] Lifeng Li. *Reformulation of the Fourier modal method for surface relief gratings made with anisotropic materials*. Journal of Modern Optics, 45(7):1313–1334, 1998.

[28] Lifeng Li. *Justification of matrix truncation in the modal methods of diffraction gratings*. Journal of Optics A: Pure and Applied Optics, 1:531–536, 1999.

4

Biosensor Design

A biosensor must necessarily comprise four components: a microfluidic system for handling of the samples, an opto-electronic system for the excitation and detection of the optical wave, a biointerface functionalized with surface receptors and a data analysis module to extract the relevant information. In bringing these components together, a practical design issue must be considered: the trade-off between reusability and disposability.

Today, microfluidic components are relatively easy and cheap to produce, via fast replication technology. Novel microfluidic devices, as opposed to their early counterparts made of silicon and glass, are made of clear plastics with good optical and chemical stability. Materials such as polydimethylsiloxane (PDMS), polymethylmethacrylate (PMMA) and other thermosets are commonly used with fast-replication technologies (replica molding, hot embossing and injection molding) for low-cost and large-volume production of microfluidic chips. Interfacing the microfluidic chip to a silicon substrate is further facilitated by soft-plastic such as PDMS as it exhibits excellent reversible adhesion[1].

The biointerface performance is critical in any biosensor design. The functionalization of the metallic SPR surface endows the biosensor with its affinity to the targeted analytes. Much development has been carried out over

[1] PDMS can be bonded on an oxidized silicon chip reversibly without any treatment, and it is reported that this bond can withstand pressures up to 20 psi, exposure to oxygen plasma at 200W for 15 seconds can make the adhesion permanent [1].

many years, though not specifically for SPR, but for Au surfaces in general. The enrichment of the surface receptors and binding capacity of a surface is accomplished via a variety of surface chemical treatments. Self-assembled monolayer (SAM) and long polymer chains have been considered to create a stable, uniform probe surface or three-dimensional receptor matrices [2]. With this biochemical functionalization, the capacity of the surface to absorb the targeted analytes while rejecting non-specific binding directly determines the biosensor specificity. The surface chemistry used in this work will be described in chapter 5.

With any surface chemistry for an SPR sensor, the design starts with the selection of the surface plasmon supporting metal surface. Two metals are generally considered: silver and gold. Silver supports a longer propagation of SP's and thus is more sensitive to biomolecule adsorption. However, its rapid oxidation renders it unstable. Despite its shorter SP propagation length, gold is often preferred as it is stable and can be readily functionalized.

In addition to non-specific adsorption, several noise sources can limit the sensitivity of a biosensor. Various sophisticated data fitting methods have been described to enhance sensor performance. These will be described in the following chapter.

Of the four elements comprising a good biosensor, the main focus of this work is the opto-electronic transduction of the biomolecule adsorption to a detectable and quantifiable signal. This chapter describes the different design possibilities and limitations of a surface plasmon interferometer biosensor, and quantifies the impact of these design choices on the sensors most important specifications, bulk and adlayer sensitivity.

4.1 Sensing Principle

Using interference as a means to detect refractive index changes is not new, and many different integrated sensors using this principle have been realized in low-index [3] and high-index material systems [4]. For an extensive review of the different types of interferometers for label free optical detection the reader is refered to an excellent review paper by Fan *et al* [5].

Their working principle is based on the fact that refractive index variations induce a phase-shift in one of the arms of the Mach-Zehnder interferometer (MZI), this phase-shift results in an intensity variation in the output waveguide. General Mach-Zehnder interferometer design principles can be applied to a surface plasmon interferometer. A sketch of our device is depicted in Fig. 4.1. The interferometer consists of a gold layer (refractive index experimentally determined, see appendix A) embedded into the silicon membrane ($n = 3.45$) on top of a supporting silica layer ($n = 1.45$), all

dimensions and length scales are depicted in the figure. The high degree of asymmetry associated with the gold layer (top interface $n \propto 1.33$, bottom interface $n = 3.45$) assures that the surface plasmon modes associated with the upper and lower of the metal-dielectric interfaces will never be able to couple, as their phase velocity differs to much. So the gold layer possesses two distinct surface-plasmon modes which propagate through the structure without influencing each-other. Exciting these modes is done by end-fire coupling from a regular SOI waveguide with the transverse-magnetic ground mode [6, 7]. Two independently propagating surface plasmon modes are excited at the beginning of the gold layer. They propagate independently through the structure and the phase of the top surface plasmon mode is influenced by the refractive index of the environment, while the phase of the bottom surface plasmon mode is insensitive to any refractive index changes. At the end of the gold layer both surface plasmon modes excite the ground mode of the SOI waveguide. Depending on the relative phase of the surface plasmon modes their contributions to the ground mode will interfere constructively or destructively.

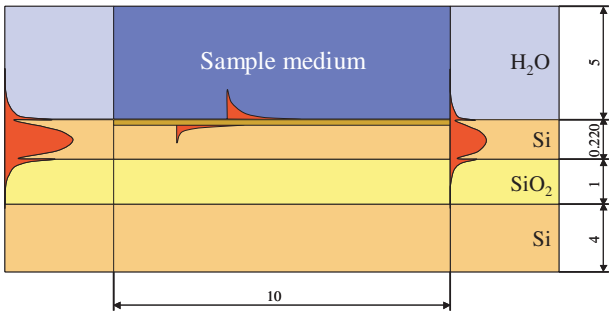


Figure 4.1: Schematical setup of the proposed structure, all dimensions in μm

Figure 4.2 illustrates the interferometric nature of our device. For a sensing section of length 10 μm, the transmitted intensity of the fundamental TM mode of the silicon slab waveguide is plotted as a function of refractive index of the sample medium. For this simulation (using the full-wave simulation technique as discussed in chapter 3), we have chosen a wavelength of 1.55 μm, which is in the near-infrared region and suitable for biosensing applications. The motivation for choosing this specific wavelength region is motivated mainly by the lower propagation losses of surface plasmon modes (propagation losses decrease with increasing wave-length), the use of SOI as a host substrate for the biosensor and the availability of measurement equipment and internal knowledge in our group [8]. When the upper and lower surface plasmon modes arrive in phase at the end of the sensing section, con-

structive interference leads to maximal transmission. However, for certain values of the sample refractive index, the phase difference between the two modes equals π, resulting into destructive interference and a minimum in the transmission curve.

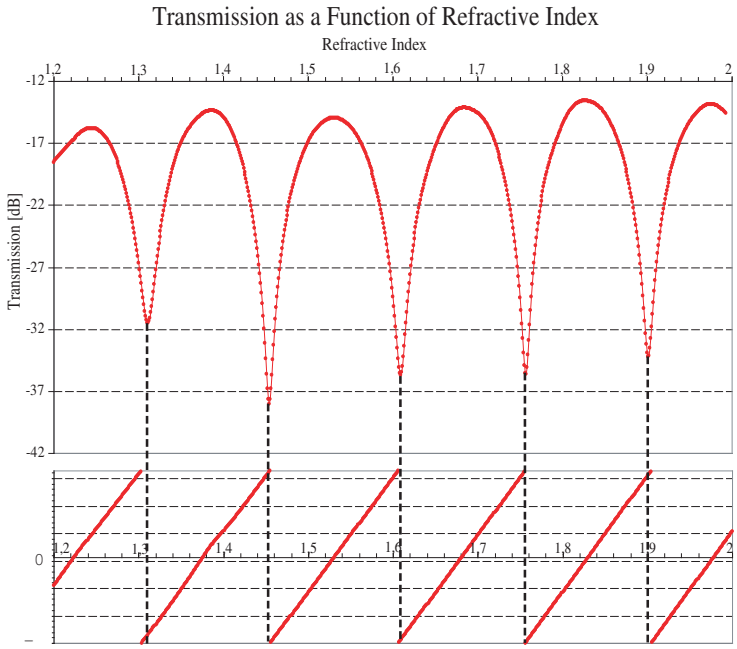


Figure 4.2: Transmission of the structure depicted in Fig. 4.1 as a function of the refractive index of the top cladding layer. The length of the Au layer is 10 μm, Au-thickness equals 50 nm, Si-thickness was 150 nm. Bottom graph depicts the relative phase difference the top and bottom surface plasmon mode acquire upon propagating through the device, destructive interference occurs if $\Phi = (2n+1)\pi$ with $n = 1, 2, 3, \ldots$

In order to exploit this principle as a (bio)sensor, two different measurement methods are commonly used. The first method is to use a monochromatic input mode and monitor the output power as a function of the refractive index of the sample. This detection approach is commonly called 'intensity interrogation mode' [9, 10]. The second mode of operation uses a broadband input mode and as a function of the refractive index of the sample medium we monitor the position of the spectral minima in the transmission curve. This approach has been called 'wavelength interrogation mode' in

the literature [9]

In Fig. 4.3 we have simulated the response of the structure shown in Fig. 4.1 to a broadband incoming waveguide mode. The refractive index of the sample medium is fixed at a value of 1.33. The obtained spectrum can also be explained by comparing the phase difference between the internal and the external surface plasmon modes as can be seen in the bottom of Fig. 4.3.

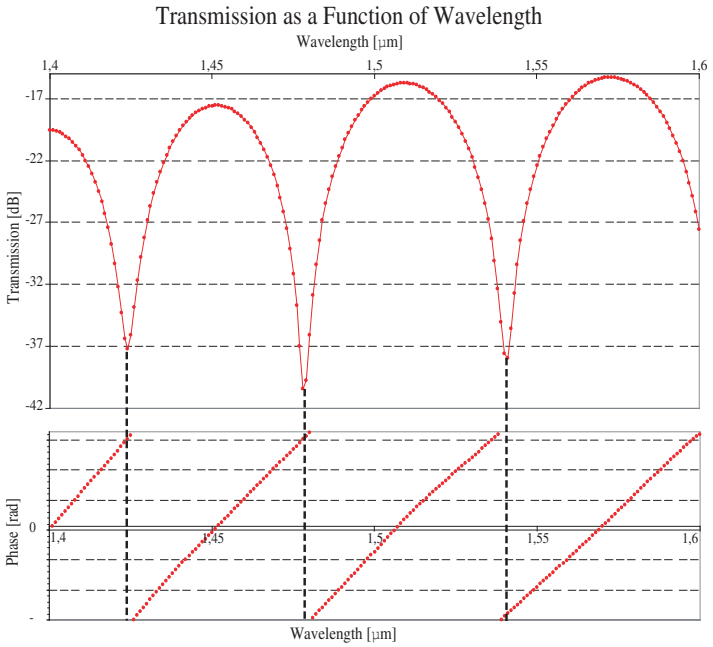


Figure 4.3: Transmission of the structure depicted in Fig. 4.1 as a function of wavelength. The length of the Au layer is 10 μm, Au-thickness equals 50 nm, Si-thickness was 150 nm, the refractive index of the top cladding is $n_{top} = 1.33$. Bottom graph depicts the relative phase difference the top and bottom surface plasmon mode acquire upon propagating through the device, destructive interference occurs if $\Phi = (2n + 1)\Pi$ with $n = 1, 2, 3, \ldots$

In both modes of operation a sensor length of approximately 10 μm should suffice, which is two orders of magnitudes smaller than current integrated surface plasmon sensors.

We will now take a look at the design requirements for such a device. Starting with the modal properties of the sensing region, then moving on to the coupling between the incoming TM ground mode and both surface plasmon modes, and finishing by tying all the different contributions together.

4.2 Factors affecting interference spectrum

In this section we present a rigorous model for the optical power transmission through a surface plasmon interferometer based on a step-index slab waveguide approximation, calculated according to the numerical techniques presented in the previous chapter. The general configuration of the device is presented in figure 4.4. Predicting the performance of such a device requires analysis of the waveguide modes supported by the metal-clad waveguide, of their excitation by an input waveguide, and of the power coupled to the output waveguide. The modes in the structure, their complex propagation constants yielding modal velocity and attenuation are numerically evaluated using CAMFR. The coupling of power from a mono-mode input waveguide across a step discontinuity into the modes of the metal-clad waveguide will be calculated and discussed. Finally, the total power coupled from these modes to the output waveguide will be calculated to yield the input/output characteristic of the sensor.

The model allows the determination of the modulation in output power of the device due to bulk refractive index changes or due to the adsorption of a thin organic layer to the sensor surface, which in turn leads to a measure of sensitivity. By explicitly calculating the dependence of the sensitivity on the composition of the composite waveguide, sensitive surface plasmon interferometers, optimized for an aqueous environment, can be designed.

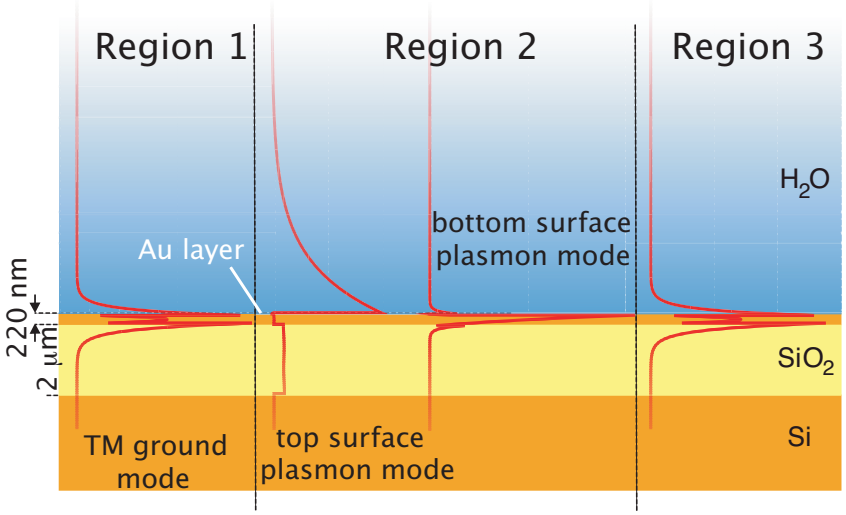


Figure 4.4: Simulation setup showing the different regions and waveguide modes to be analyzed.

The planar device shown in figure 4.4 must be broken down into two parts to be fully analyzed. Region 1 is considered to be a lossless single moded, single layer, input and output waveguide. This region is solely employed to couple light in and out of the multilayer region. Region 2, the metal-clad multilayer region acts as the interaction zone.

The first step in the sensor analysis is the analysis of the metal-clad region. The structure under study is shown in figure 4.4. It consists of a semi-infinite[2] gold film, supported by a semi infinite Si waveguide layer with a thickness of $150\ nm$, a semi infinite buried-oxide layer (thickness $2\mu m$) and a substrate layer composed of Si and covered by a semi-infinite homogeneous dielectric with the refractive index of water. The optical free-space wavelength of analysis is set to $\lambda_0 = 1.55\mu m$.

4.2.1 Modal properties and dispersion diagram

The modal solutions for a $50\ nm$ thick Au layer were found as a function of wavelength λ. The complex propagation constants of the lowest-order bound and leaky modes are plotted in figure 4.5. Several important trends can be discerned from these plots. First of all, in the wavelength region of 1500 to $1700\ nm$ we are quite far removed from the Au plasma frequency ω_p, therefore we are in the regime close to the lightline, and a featureless dispersion relation is to be expected. One can see that the wave-vector of the top surface plasmon mode is close to the H_2O-lightline, and agrees relatively well with the wave-vector of a surface plasmon mode propagating on an isolated $Au - H_2O$ surface. The ratio between field components E_x and E_y is approximately equal to $\epsilon_m'^{1/2}$, as predicted by theory in the long wavelength limit. The bottom surface plasmon mode however has a wavelength vector in between the Si and the SiO_2 lightline, characteristic of a dielectric guided mode. This means that the bottom surface plasmon mode is actually a hybrid surface plasmon-dielectric guided mode. Field components E_x and E_y are not related through single interface surface plasmon theory.

If the metal is thin compared to the penetration depth of the surface plasmon modes, the two surface plasmon modes can experience mutual coupling and therefore the modal properties will be more complex than the properties of the eigenmodes supported on the isolated surface. It is well known from literature [11, 12, 13, 14, 15] that thin metal layers in a symmetric environment support so called long-range surface plasmons

[2] semi-infinite refers to the fact that full-wave simulations are 2D in nature. Along the third dimension, perpendicular to the surface of the page, all layers are assumed to be infinitely extended.

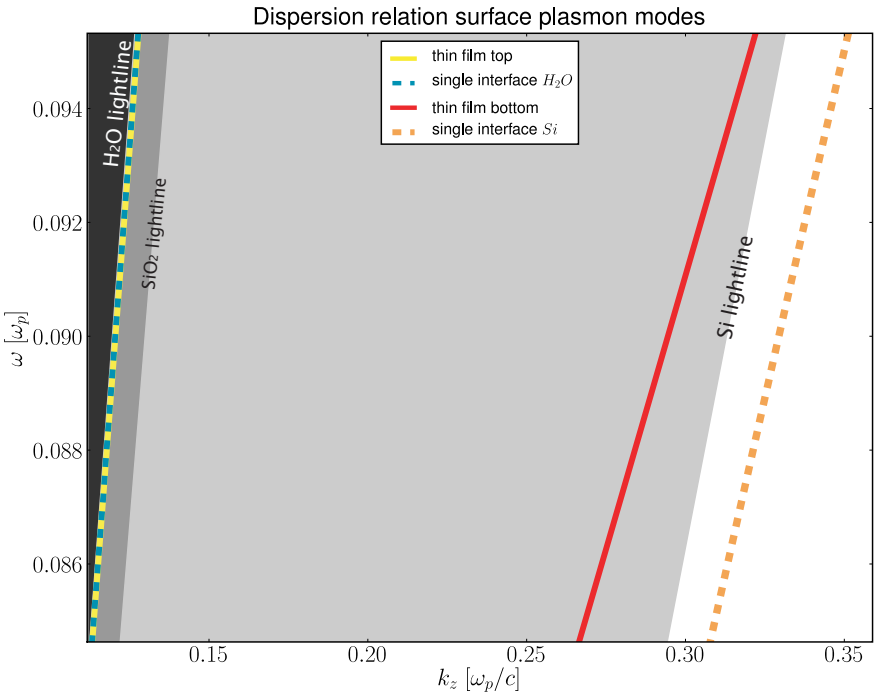
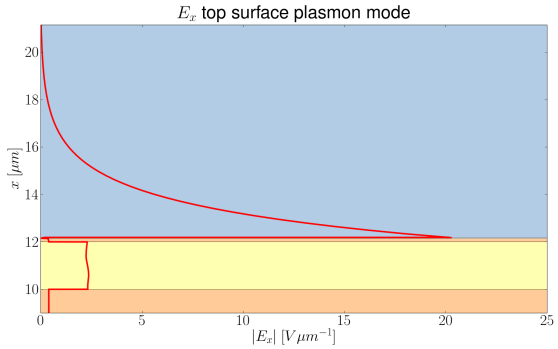


Figure 4.5: ω-k diagram for the bottom and top surface plasmon modes and comparison with modes propagating on isolated Au-H_2O and Au-Si surfaces. Numerical results obtained with CAMFR, Au refractive index experimentally determined (see appendix A).

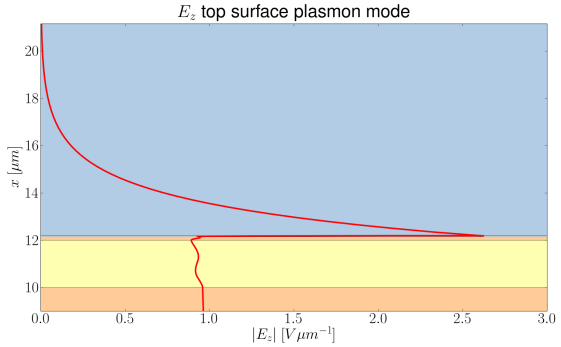
(or symmetric surface plasmons[3]) and short-range surface plasmons (or asymmetric surface plasmons). For different dielectrics this behavior is preserved until a certain asymmetry is reached, we will see that for our structure the symmetric-antisymmetric nomenclature will not be of much use. From hereon we will use the designation top surface plasmon mode and bottom surface mode.

In figures 4.6a, 4.7a, 4.6b, 4.7b, 4.6c and 4.7c one can see the transverse distribution of the three field components $|E_x|$, $|E_z|$ and $|H_y|$ for the modes at the upper and lower interfaces of the Au film. From these plots we can see that the top surface plasmon is indeed a purely bound surface plasmon mode, while the bottom mode is in fact a hybrid dielectric-surface plasmon mode exhibiting both features of a (leaky) surface plasmon mode at the lower surface and the TM zeroth order guided mode of the Si- waveguide.

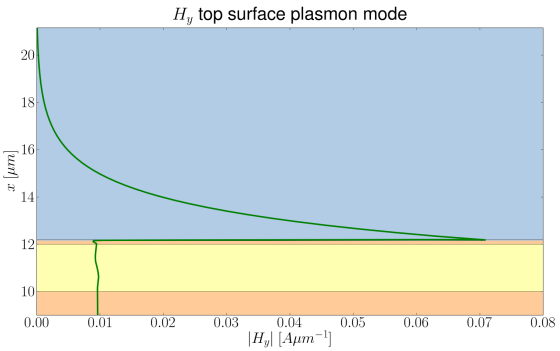
[3]The term symmetric/antisymmetric refers to the symmetry of the magnetic field H_y with respect to the interface



(a) $|E_x|$ field.



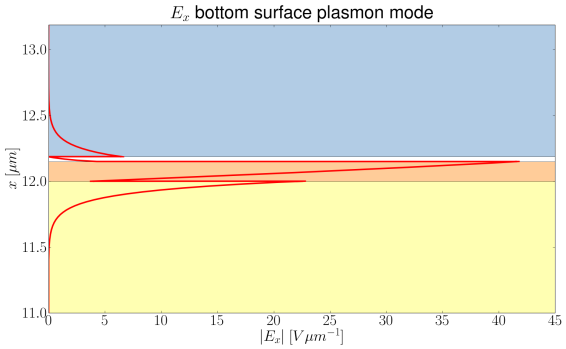
(b) $|E_z|$ field.



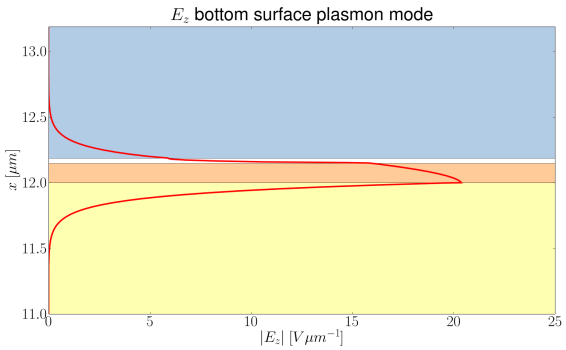
(c) $|H_y|$ field.

Figure 4.6: Field profiles of the top surface plasmon mode.

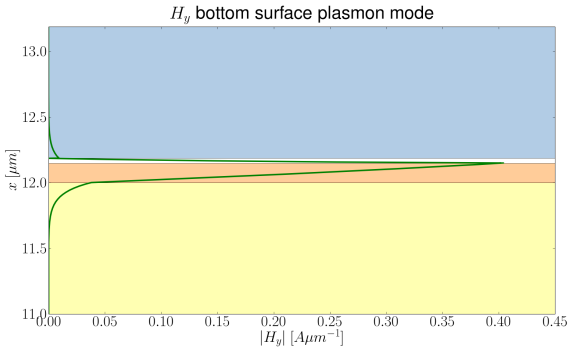
The energy flux of an electromagnetic field is given by its Poynting



(a) $|E_x|$ field

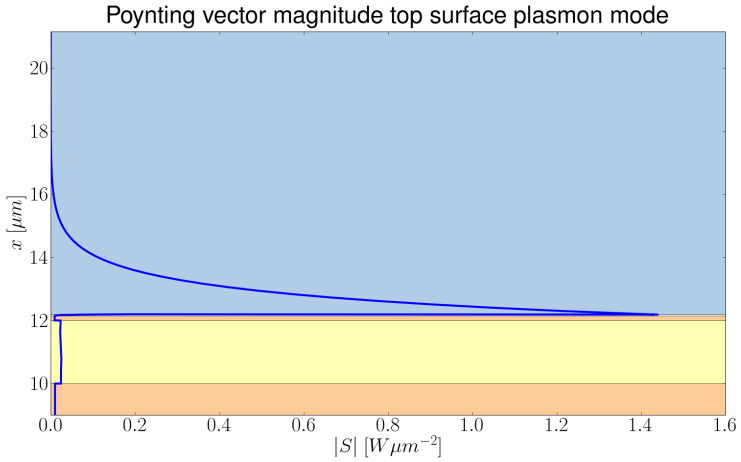


(b) $|E_z|$ field

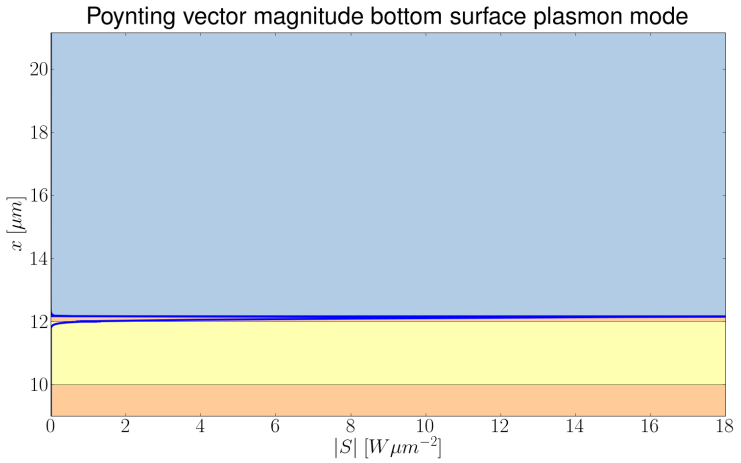


(c) $|H_y|$ field.

Figure 4.7: Field profiles of the bottom surface plasmon mode, notice the different vertical scale as compared to the plots of the top surface plasmon mode field profiles.



(a) Top surface plasmon mode.



(b) Bottom surface plasmon mode.

Figure 4.8: Magnitude of the Poynting vector of the top and bottom surface plasmon mode.

vector $\mathbf{P} = \mathbf{E} \times \mathbf{H}^*$, it is well known from literature that the Poynting vector of surface plasmon waves is oriented in the direction of the propagation direction with a slight inclination towards the metal surface, thus directing energy towards the lossy metal. The magnitude of the Poynting vector is a measure for the amount of energy flux, regardless of direction. This quantity is plotted in figures 4.8a and 4.8b. The energy flux of the top surface plasmon modes is located above the Au layer (although there is a very small contribution in the buried oxide layer), the bottom surface plasmon mode's energy flux is mostly confined to the bottom interface of the Au layer. This shows the level of separation between these two very distinct modes.

4.2.2 Influence of the Au film thickness

The behavior of these two modes as as function of gold thickness is of special interest, since varying the gold layer thickness is one of the parameters one can use to tune the sensor response. Over a broad domain of gold thicknesses (100-50 nm) the effective index remains essentially unchanged, below $50\ nm$ the effective index of the bottom surface plasmon mode increases, as does the loss. Confinement of that mode increases and for a Au thickness of $25\ nm$ the hybrid surface-plasmon dielectric mode evolves into a purely bound surface plasmon mode with an effective index larger than the refractive index of Si. The effective index of the top plasmon mode remains unchanged over the entire domain, the loss increases for small Au thicknesses. If the Au thickness becomes smaller than the surface plasmon penetration depth, radiation losses into the high index Si substrate increase substantially. Plots of the effective index and the losses of the top and surface plasmon mode can be seen in figures 4.9 and 4.10 respectively.

4.2.3 Influence of the Si slab thickness

The Si-slab thickness has an effect on both the effective index and the losses of the bottom hybrid surface plasmon mode, while the top surface plasmon mode's properties only show a very weak dependence on d_{Si}. In figure 4.11 the effective index and the losses of the top and bottom surface plasmon modes are plotted as a function of d_{Si} for several different Au-film thicknesses. As the thickness of the Si waveguide layer decreases, the fraction of the optical power of the bottom surface plasmon mode in the SiO_2 layer increase, leading to a decrease in the propagation constant. At first, losses increase substantially as the silicon thickness decreases since the plasmon mode gets more and more confined to the gold surface. At a certain point the Si thickness is too small to sustain a hybrid-surface

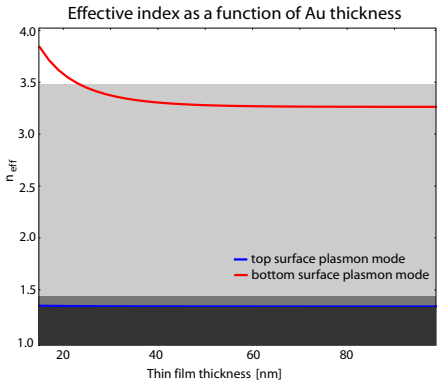


Figure 4.9: Effective index of the bottom and top surface modes as a function of Au layer thickness. The light shaded area is below the Si lightline, the darker shaded area is below the H_2O lightline, darkest areas is below the SiO_2 lightline

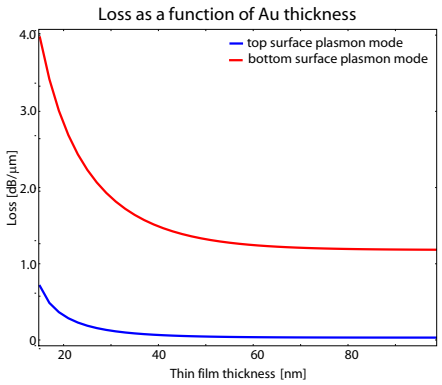


Figure 4.10: Loss as of the top and bottom surface plasmon mode as a function of Au layer thickness.

plasmon dielectric mode and the mode evolves to a surface plasmon mode on a Au-SiO_2 surface, with lower losses.

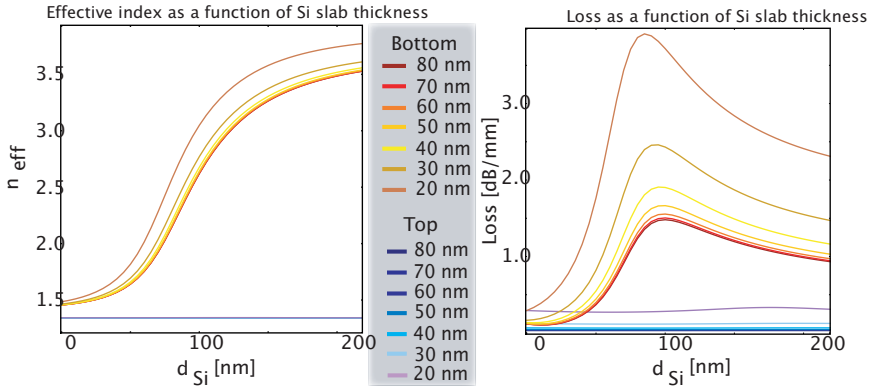


Figure 4.11: Effective index (left) and loss (right)of the bottom and top surface modes as a function of Si slab thickness, calculated for several Au layer thicknesses.

4.2.4 Single mode condition

In order to maximize visibility both interferometer arms (in casu the path along the top surface and the path along the bottom surface) should be made monomodal. For the upper path it is evident that the surface plasmon mode at the gold-dielectric (in this case H_2O) interface is the only guided mode that is supported. For the lower (dielectric) path this is less evident. However, one should take into account that the presence of the gold layer influences the cut-off properties of the silicon slab waveguide. Figure 4.12 plots the effective indices of all guided TM modes of a silicon membrane slab waveguide with and without a thin 50 nm gold layer on top. While the unclad waveguide has an index-guided mode for all thicknesses considered in this plot, the gold-clad waveguide has a cutoff thickness for index-guided modes of about $230\ nm$. For all thicknesses below this value, only the internal hybrid surface plasmon - dielectric mode remains. By carefully selecting the thickness of the silicon layer, in casu less than 230 nm, the only two modes that exist in the gold-clad waveguide section are the internal and the external surface plasmon wave, ensuring that both paths are monomode.

4.2.5 Coupling to and from surface plasmon modes

The second step in calculating the power transmission is to analyze modal coupling between regions 1 and 2. At the abrupt waveguide transition between regions 1 and 2 the incoming mode will be decomposed in all the guided and radiation modes of region 2, power distribution among these modes being dependent on the field profiles of all modes involved in the decomposition and thus on the physical structure of both regions. While a continuum of radiation modes will undoubtedly be excited too, these modes do not contribute to the final output as power coupled to these modes is essentially lost. At the end of the sensing region 2, the power of the aforementioned modes will be (partially) transferred to the guided mode of the output region. It is this output power of the device that carries the information on refractive thickness of the analyte and/or adlayer thickness.

Coupling losses were calculated by evaluating the numerically calculated modal transmission and reflection coefficients at the interface between regions 1 and 2. Results are plotted in Fig. 4.13 for different configurations of the sensing section. When the gold layer is not embedded but simply put on top of the silicon membrane, the coupling efficiency to the upper and lower surface plasmon modes differ strongly. However, by embedding the gold layer into the silicon membrane, the coupling efficiency can be adjusted over a very large range. Although it is possible to divide the input light equally over both branches of the interferometer, due to the difference

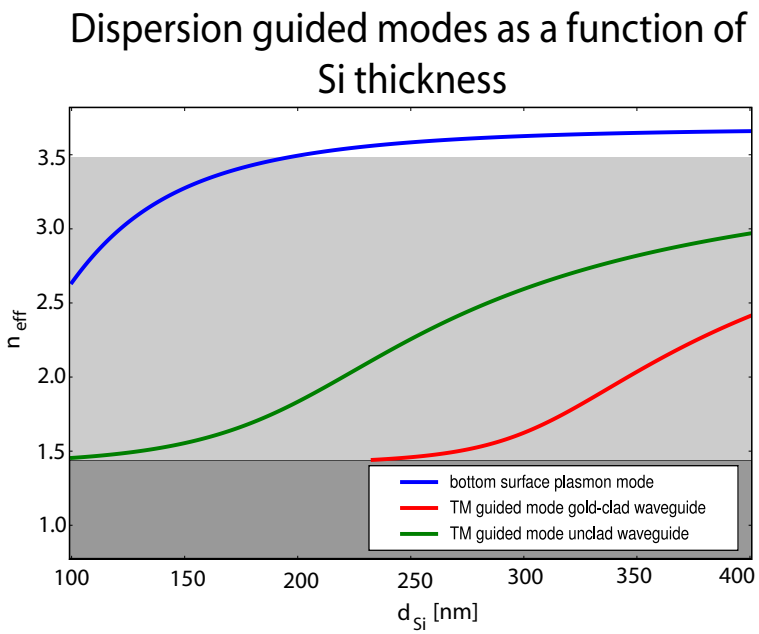


Figure 4.12: Dispersion of the effective indices of the guided modes as a function of the Si waveguide thickness. The presence of the gold layer on top of the waveguide shifts the effective index of the index-guided mode to lower values. For a waveguide thickness of 230 nm the effective index of this mode is smaller than the refractive index of SiO_2 so that this mode starts to radiate into the silica. For thicknesses below 230 nm the gold-clad Si waveguide has no index-guided modes.

in propagation loss of both surface plasmon modes this is not desirable.

4.2.6 Influence of finite stripe width

Up until now, the device was simulated as a 2D structure, with an infinite dimension in the direction perpendicular to the surface of the page (the structure was considered to be infinitely wide). This section describes what happens if the width of the structure is finite. A simple physical picture can be derived for the number of surface plasmon modes supported in a metal strip of finite width. Given that the propagation constant along the stripe asymptotically approaches the value along an extended film, it is reasonable to assume that the in-plane momentum for a surface plasmon k_{sp} is conserved for a surface plasmon on a finite width structure [16]. However, along the finite width stripe, this in-plane wave-vector can be separated in a component along the propagation direction (k_z) and a component along the

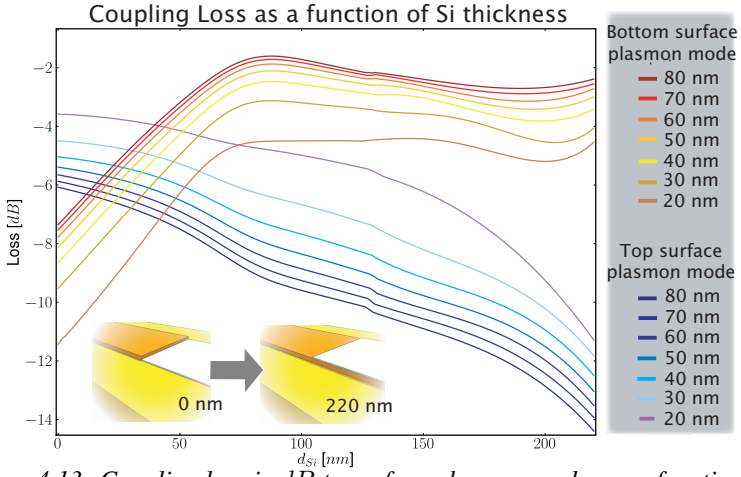


Figure 4.13: Coupling loss in dB to surface plasmon modes as a function of the waveguide thickness. The input waveguide has a thickness of 220 *nm, the thickness of the Si layer supporting the gold layer varies from* 220 *nm (Si layer is equally thick as the Si layer in the input waveguide) to* 0 *nm (Au layer is on top of the* SiO_2 *layer).Thickness of the Au layers varies from* 80 *to* 20 nm*, in steps of* 10 nm

lateral direction (k_y) such that

$$k_{sp}^2 = k_z^2 + k_y^2 \tag{4.1}$$

where neglecting the finite stripe thickness, the in-plane wave-vector can be approximated by the expression for a surface plasmon supported by a metal-dielectric interface (see equation 2.19). Or, less approximative, the numerically calculated value for a thin stripe surface plasmon can be used.

In order for wave-guiding to occur, there is a lower limit placed on upon the propagating component of the momentum. The guided surface plasmon should not be able to couple to the radiation modes of the surrounding dielectric material; therefore the propagating wave-vector should exceed the effective wave-vector in the dielectric medium (i.e. $|k_z| > k_0\sqrt{\epsilon_d}$). At this guiding limit we can derive the maximum value for the lateral wave vector

$$|k_x| < k_{sp} - k_0\sqrt{\epsilon_d} \tag{4.2}$$

This maximum value for k_x limits quantization along the lateral direction, and thus, we can approximate the number of allowed surface plasmon modes (N) by relating the width (W) of the interface to the lateral wave-vector. As the highest order supported mode would have N extrema in the lateral direction, the maximum wave-vector would be

$$k_{x,max} = \frac{N\pi}{W} \tag{4.3}$$

Simplifying equations 4.2 and 4.3, we can derive an approximate analytical expression for the number of supported modes [17]

$$N < \frac{W}{\pi}(k_{sp} - k_0\sqrt{\epsilon_d}) \tag{4.4}$$

The top Au interface will -due to the fact that the surface plasmon mode has a k-vector close to the lightline - always be monomodal, regardless of waveguide width. The bottom Au interface however, will become multimodal for a finite waveguide width. The exact number of modes propagating at the bottom interfaces depends on the thickness of the Si layer, for a thickness of $150\ nm$, a $10\ \mu m$ broad waveguide should posses up to 25 distinguishable surface plasmon modes, for a $3\ \mu m$ broad waveguide, this value drops to a mere 3. Simulations carried out in the rest of this work do not take this multimodal behavior into account.

4.3 Design Principles

After discussing the basic components comprising the surface plasmon interferometer in general, and calculating their dependence on waveguide slab thickness and gold layer thickness, we can now derive some general rules one should take into consideration when designing a surface plasmon interferometer for specific purposes. Multimode interferometer output depends on all earlier described parameters (d_{Si}, d_{Au}), this section ties all of those parameters together and calculates interferometer output. Power in the output of a Mach-Zehnder can be described by [18, 19],

$$P_{out} = P_{in}(|T_s|^4 e^{-2\alpha_s L} + |T_r|^4 e^{-2\alpha_r L})(1 + V\cos\Phi) \tag{4.5}$$

where the MZI visibility V is given by

$$V = \frac{2|T_s|^2|T_r|^2 e^{-\alpha_r L} e^{-\alpha_s L}}{|T_r|^4 e^{-2\alpha_r L} + |T_s|^4 e^{-2\alpha_s L}} \tag{4.6}$$

where Φ is the phase shift between guided modes in the sensing arm and the reference arm (including phase shifts due to the modal coupling), $|T_{r,s}|^2$ is the power transmitted from the input mode to the reference mode and the sensing mode respectively . The visibility factor (V) gives the contrast between the maximum and the minimum transmitted intensity and depends on the coupling factor of the splitter and the propagation losses of guided modes in the interferometer arms. Results obtained in this section will be calculated using the above approximative formulas, although all effective

indices of the contributing plasmonmode and the transmission coefficients used are calculated using the full-wave simulation.

The resonance wavelength or resonance refractive index of classic integrated SPR sensors can be tuned to some extent by using buffer layers between the dielectric waveguide layer and the gold layer or by choosing the appropriate metal [20, 21, 22, 23]. The fact that our device is based on interference rather than phase-matching allows one to tune the response of the device much more easily. There is no need for additional buffer layers, or a different metal to achieve the appropriate behavior, one simply has to modify the geometry of the device.

The first, and perhaps most important design parameter, which hasn't been discussed up until now, is the length of the sensing region L. By varying the length one can tune the visibility V, and the position of the minima in the transmission spectrum. However, as will be shown below, the length of the device has to satisfy two different, sometimes conflicting, equations, and as such can no longer be considered to be a parameter that can be freely chosen.

The wavelengths, or refractive indices of the sample-medium, for which the transmission is minimal are completely determined by the following equation

$$\begin{aligned} &|(\phi_{b0} + \phi_{prop,b} + \phi_{0b}) - \\ &(\phi_{t0} + \phi_{prop,t} + \phi_{0t})| = \pi, \end{aligned} \tag{4.7}$$

where ϕ_{b0} is the phase difference due to the coupling of the incoming silicon waveguide mode to the bottom surface plasmon mode, $\phi_{prop,b}$ is the phase due to the propagation of the bottom surface plasmon mode along the sensing region, and ϕ_{0b} is the phase difference due to the coupling of this bottom surface plasmon mode to the dielectric guided mode in the output section. Because of reciprocity $\phi_{b0} = \phi_{0b}$. Similar definitions hold for the top surface plasmon mode. Using this equation one can easily tune the position of minimal transmission to a desired wavelength or refractive index range.

The second equation involving the length of the sensing section deals with losses in the sensing structure, which obviously has a huge influence on the visibility V. To achieve total destructive interference the losses along each path should be the same, so that at the end both interacting modes carry the same amount of power. These conditions ensure that the visibility V will be maximal. Depending on the wavelength or refractive index for which one would want the transmission to reach a minimum, it is quite straightforward to prove, by taking the partial derivative of equation 4.6 with respect to the

sensor length L, that the optimal length is given by

$$L = \frac{1}{k_i^{top} - k_i^{bottom}} \frac{1}{\log e} \log\left(\frac{|T_{b0}|^2}{|T_{t0}|^2}\right), \tag{4.8}$$

where k_i^{bottom} and k_i^{top} are the imaginary (positive) parts of the k-vector of the bottom and top plasmonmode, T_{b0} and T_{t0} are the transmission coefficients of the incoming dielectric mode to the internal and the external plasmon modes at the chosen wavelength or refractive index for which the transmission should be minimal.

Equation 4.7 provides us with a multitude of possible lengths corresponding to the order of the interference effect, equation 4.8 only provides us with one single possible length. By varying only the length of the device it is impossible to satisfy both equations at the same time. This automatically brings us to the other two design parameters, which were discussed earlier, namely the thickness of the Si slab waveguide and the thickness of the gold layer.

For a given wavelength or refractive index, one can see from equation 4.8 that it is possible to change the length for which the power along both paths is the same. In order to do so we have to change the values of $|T_{t0}|^2$, and $|T_{b0}|^2$, which can be done by reducing the thickness of the silicon layer in the sensing region (as can be seen on Fig. 4.13). This provides us with an extra design parameter, which allows us to satisfy both equations more accurately.

Enhancing the coupling to the external plasmon mode, and in doing so minimizing the ratio $\frac{|T_{t0}|^2}{|T_{b0}|^2}$, results in other significant benefits, such as a smaller sensor length, an increased sensitivity of the sensor, and an overall higher transmission.

There is one fundamental limit however. Since the inner plasmon mode is propagating in a high index region its intrinsic loss will always be higher than the loss of the external plasmon mode (for a wavelength of 1.55 μm, and a refractive index of 1.2, the loss of the external plasmon mode is equal to 0.04 $dB/\mu m$ while the loss of the internal plasmon mode equals 1.79 $dB/\mu m$). Because of this the coupling to the inner mode should always be greater than the coupling to the outer mode.

To visualize the interferometer design rules, figure 4.14 and 4.15 show the visibility factor V and the total power transmitted through the structure P_{out} (for $P_{in} = 1$) as a function of the two most important design parameters, thickness of the Si slab (d_{Si}) and sensor length L. Since not all sensor lengths fulfill the phase condition 4.7, only a small number of lengths can be chosen for a given value of d_{Si}, these are plotted as gray lines in figure 4.14

and 4.15. The different lines correspond with different interference orders. The dotted line in figures 4.14 and 4.15 is the device length that satisfies equation 4.8, ensuring maximum visibility. The intersection points of the dotted line with the gray lines are ideal combinations of d_{Si} and L.

At first, starting from a 220 nm thick Si waveguide, both the reduction in coupling loss to the top surface plasmon mode and the increased propagation loss of the bottom surface plasmon mode lead to smaller ideal device lengths. At a certain thickness d_{Si} ($98nm$ for $d_{Au} = 50\ nm$), the propagation loss of the bottom surface plasmon mode decreases, leading to larger values of the ideal total device length L. As the difference between coupling loss to the top and bottom mode gets smaller and smaller, the ideal device length gets smaller again, ultimately becoming negative (and unphysical) when $|T_{t0}| > |T_{b0}|$.

Figures 4.14 and 4.15 shows overall transmission losses, based on formula 4.5. These graphs shows that the ideal working point for a device with a 50 nm thick gold layer is the intersection of the dotted line (power length) with the gray line (phase length) in the vicinity of $d_{Si} = 100\ nm$.

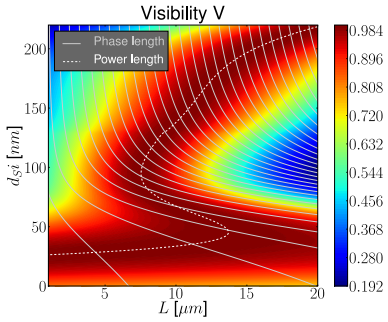


Figure 4.14: Visibility V (as defined in equation 4.6) as a function of d_{Si} and L. All values calculated for a wavelength of 1.55 $/mum$ and a Au layer thickness of 50 nm. Gray lines are lengths L for which the phase condition (eqn. 4.7) is fulfilled.

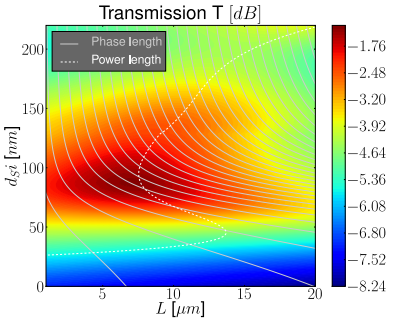


Figure 4.15: Transmission T (as defined in equation 4.5) as a function of d_{Si} and L. All values calculated for a wavelength of 1.55 μm and a Au layer thickness of 50 nm. Gray lines are lengths L for which the phase condition (eqn. 4.7) is fulfilled.

The correct design parameters, for a given operating wavelength and analyte refractive index can be calculated this way. One simply has to choose the Si-thickness and sensor length L, that maximize the visibility V. The ideal device length calculated using the visibility formula (the dotted line in figure 4.14), gives a number of possible device lengths, more specifically the intersection points with the phase length curves (gray lines). The best

candidate, having the fewest overall losses, can then be determined using figure 4.15.

An additional design parameter is the thickness of the gold layer. The effects of changing the thickness of the gold layer can be seen in figure 4.16 and 4.17, which show the visibility V and transmission T for a surface plasmon interferometer with a Au layer thickness of 30 nm. Reducing the gold layer thickness generally shifts ideal device length to smaller values. This can be attributed to a smaller difference in coupling loss to the top and bottom surface plasmon mode, and a larger difference in propagation loss between the top and bottom surface plasmon mode (as can be seen in figures 4.11 and 4.13). The decreasing propagation loss of the bottom mode for smaller d_{Si} values has a smaller effect on the power length curve (dotted line), since the smallest possible value of d_{Si} yielding physical values for the length satisfying equation 4.8 becomes larger ($\propto 50\ nm$ for $d_{Au} = 30\ nm$, compared to $\propto 26\ nm$ for $d_{Au} = 50\ nm$). This is due to the fact that the difference in coupling loss is smaller, hence these curves intersect for e values of d_{Si} if d_{Au} gets smaller. Therefore the peak in power length (dotted line in figure 4.14) is no longer present in figure 4.16.

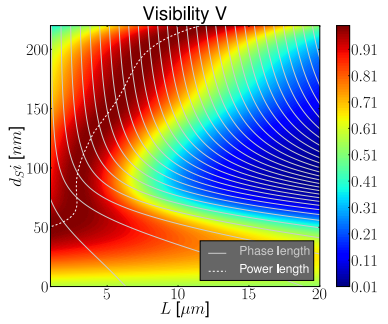


Figure 4.16: Visibility V (as defined in equation 4.6) as a function of d_{Si} and L. All values calculated for a wavelength of 1.55 $/mum$ and a Au layer thickness of 30 nm. Gray lines are lengths L for which the phase condition (eqn. 4.7) is fulfilled.

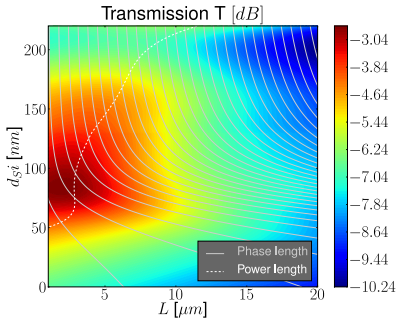


Figure 4.17: Transmission T (as defined in equation 4.5) as a function of d_{Si} and L. All values calculated for a wavelength of 1.55 $/mum$ and a Au layer thickness of 30 nm. Gray lines are lengths L for which the phase condition (eqn. 4.7) is fulfilled.

Even though the thickness of the Au layer is different in Figs. 4.15 and 4.17 a couple of general design rules for the length of the sensing region and the thickness of the silicon waveguide layer d_{Si} can be derived. Transmitted power severely drops for $d_{Si} < 50\ nm$ and for $d_{Si} > 160\ nm$, these values will serve as both a lower and upper boundary in subsequent

sensor design. As for the length of the sensing region L, device performance seems to be limited to $\approx 15\ /mum$ for a $50\ nm$ thick Au layer, and ≈ 10 μm for a $30\ nm$ gold layer. The maximum attainable length of the device seems to be inversely proportional to the thickness of the gold layer. These two observations will be used as rules-of-thumb while designing the final structure and calculating device sensitivity.

Using these design rules we have optimized the sensor depicted in Fig. 4.1 so that the transmission would be minimal for a refractive index of the sample medium of 1.33, and a wavelength of $1.55\ \mu m$. In the optimized design the thickness of the silicon membrane is equal to $101\ nm$, d_{Au} was $50\ nm$, and the length of the device is equal to $6.055\ \mu m$. Simulation results for this device are shown in Fig. 4.18.

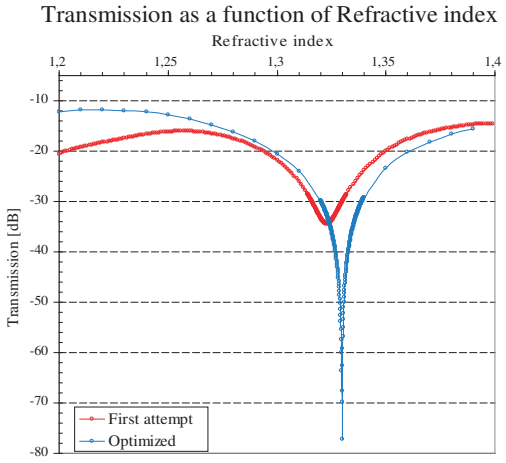


Figure 4.18: Comparison of the transmission as a function of wavelength before and after optimization of the design parameters d_{Si}, d_{Au} and L, full-wave simulation.

4.4 Sensitivity

4.4.1 Figures of Merit

In sensor development sensitivity S is an important parameter to evaluate sensor performance. Sensitivity is the magnitude of the sensor transduction signal change in response to the change in analyte. Fundamentally, it is determined by the strength of the light-matter interaction [19]. In most

evanescent based sensors this is determined by the fraction of the light in solution, or the light intensity at the sensor surface [24, 25]. The unit in which sensitivity is expressed obviously depends on the interrogation method (intensity vs. spectral), and the type of sensing performed (bulk vs. adlayer). Hence we can express sensitivity for a specific biosensor in units of $[nm/RIU]$ (bulk sensing, spectral interrogation), $[nm/nm]$ (adlayer sensing, spectral interrogation, analyte dependent), $[W/RIU]$ or $[dB/RIU]$ (bulk sensing, intensity interrogation) and $[W/nm]$ or $[dB/nm]$ (adlayer sensing, intensity interrogation, analyte dependent) [26]. As a means of comparison, demonstrated bulk sensing sensitivities for various refractometric sensors are given in table 4.1. In general, designs that enable high interaction between the optical mode (or power) and the sample have high sensitivity.

RI Sensor Configuration	Demonstrated RI sensitivity	Reference
SPR (prism coupled)	7120 nm/RIU	[27]
SPR (grating coupled)	3365 nm/RIU	[28]
Long period fiber grating	6000 nm/RIU	[29]
Capillary ring resonator	800 nm/RIU	[30]
2D Photonic Crystal	200 nm/RIU	[31]
Planar ring resonator	212 nm/RIU	[32]
Microsphere ring resonator	30 nm/RIU	[33]

Table 4.1: Sensitivity comparison of the different SPR approaches, table taken from [34].

It is relatively straightforward to convert a sensors bulk sensitivity S_{bulk} into a biomolecule sensitivity, because capture of biomolecules essentially changes the refractive index in the region probed by the sensor's evanescent fields. For ring resonator sensors, an analysis to convert bulk sensitivity to biomolecule sensitivity has been presented and experimentally verified [35], this analysis is generally applicable to all types of resonant RI sensors. The analysis leads to an expression for the spectral shift of the resonant mode, $\delta\lambda$, for biomolecule capture given a bulk sensitivity of S_{bulk}

$$\frac{\delta\lambda}{\lambda} = \sigma_p \alpha_{ex} \frac{2\pi\sqrt{n_m^2 - n_s^2}}{\epsilon_0 \lambda^2} \frac{n_m}{n_s^2} S_{bulk} \tag{4.9}$$

In this expression, σ_p is the surface density of the captured biomolecules, α_{ex} is the excess polarizability of the molecule, n_m the refractive index of the sensor material and n_s the refractive index of the sample buffer. One can now relate the wavelength shift $\delta\lambda$ with the amount of biomolecules

attached to the surface σ_p by the expression

$$S_{biomolecular} = \frac{\delta\lambda}{\sigma_p}, \tag{4.10}$$

thus obtaining a value for biomolecular sensitivity, which is highly dependent on the properties of the biomolecule itself.

Another important parameter is the detection limit. By taking the uncertainty σ in the transduction signal into account, the this value can be deduced , i.e. the minimal resolvable signal $DL = \sigma/S$ where S is the sensitivity.

Detection limit in the intensity interrogation approach is mostly limited by the signal-to-noise ratio. The detection limit for spectral interrogation factors in spectral resolution and system noise since it is dependent on the methodology for measuring the spectral shift. Spectral resolution and system noise detract from the precision with which the true center of the spectral attenuation feature can be located. The minimal error in determining the actual spectral position leads to finite sensor resolution, and subsequently to the need for advanced data-analysis methods. Similarly, the detection limit for sensors using intensity interrogation depends on the minimal measurable change in output intensity, and hence depends on several contributing noise factors.

For optical based evanescent refractive index sensors there are three ways to specify the detection limit. For refractometric sensing the DL reports the smallest sample refractive index change that can be accurately measured, a detection limit in units of refractive index units (RIU) is then naturally used to quantify sensor performance, which enables a rough comparison of the sensing capability among different optical technologies and structures. For biomolecular sensing the DL quantifies the minimum amount of analyte the sensor can measure. Surface mass density (or total mass) in units of pg/mm^2 (or pg) - which is what the sensor actually measures - can be used to quantify the sensing capabilities.

Another option is to use surface density ($[\#/mm^2]$). While experimentally surface mass density is difficult to determine accurately, it reflects the intrinsic detection capability of a sensor and can be used to evaluate sensor performance. Using this value one can compare sensors in terms of ability to detect a low density of molecules on the sensor surface. However, this does not include the sensors ability to capture molecules on the surface.

The third way is to use sample concentration (in units of ng/ml or molarity). Sample concentration is quite useful and easy to determine from an experimental point of view, as no detailed information regarding the mass density on the surface is needed. However, the DL defined in this manner depends on the target molecule and its affinity to the biorecognition molecule

on the surface, and therefore needs to be specified for each individual type of biomolecule. These three DL's are obviously correlated and the detailed relation between them needs to be studied for each biosensor.

The detection limits for the different surface plasmon resonance biosensors that were discussed in the introduction are mentioned in table 4.2

SPR Approach	DL(RIU)	Reference
Fiber and waveguide SPR		
-Micro-prism	8×10^{-4}	[36]
-Polarization-maintaining fiber	4×10^{-6}	[37]
-Multi-mode waveguide	4×10^{-3}	[38]
-DBR-waveguide	1.28×10^{-5}	[39]
-Dual-LED	5.2×10^{-4}	[40]
-Y-splitter and SU8 microfluidics	6×10^{-4}	[41]
Silicon SPR		
-SPR-Schottky	12×10^{-5}	[42]
Imaging SPR		
-Polarization contrast	2×10^{-6}	[43]
Sentata Spreeta	1×10^{-6}	[44]

Table 4.2: Detection limit comparison of the different SPR approaches, table taken from [45].

In the following sections we will discuss these sensitivity numbers for the biosensor that was studied in this work.

4.4.2 $\partial N_{eff}/\partial n_c$ and $\partial N_{eff}/\partial a$

The most important contribution to the bulk sensitivity S is undoubtedly $\partial N_{eff}/\partial n_c$ (where N_{eff} is the complex effective index of the top surface plasmon mode and n_c is the refractive index of the top cladding layer), this value breaks down into a real and an imaginary part yielding the effective bulk index sensitivity $\partial n_{eff}/\partial n_c$ and the normalized attenuation bulk sensitivity $\partial k_{eff}/\partial n_c$. Likewize, the most important contribution to surface sensitivity is $\partial N_{eff}/\partial a$, where a is the thickness of the adlayer. Again this value breaks down into a real and an imaginary part yielding the effective index surface sensitivity $\partial n_{eff}/\partial a$ and the normalized attenuation surface sensitivity $\partial k_{eff}/\partial a$.

These sensitivities can be computed directly by approximating the differentials via $O(h^2)$ central finite-difference formulas:

$$\frac{\partial N_{eff}}{\partial n_c} = \frac{N_{eff}(n_c + h_c - jk_c) - N_{eff}(n_c - h_c - jk_c)}{2h_c} \quad (4.11)$$

$$\frac{\partial N_{eff}}{\partial a} = \frac{N_{eff}(a + h_a) - N_{eff}(a - h_a)}{2h_a} \quad (4.12)$$

To within numerical accuracy, these approximation improve in reliability as $h_a, h_c \rightarrow 0$[4]. Surface sensitivity depends on a (and bulk sensitivity depends on n_c), but only weakly so if a is small compared to the mode width.

Another expression can be derived for the bulk and surface index sensitivity, based on perturbation theory [46]. It is possible to determine the change in the phase constant of a mode due to a longitudinally invariant perturbation in the permittivity distribution of the structure. From this phase change we can determine the effective index change. Neglecting losses we can write the perturbed and unperturbed fields as:

$$\mathbf{E}_u = \mathbf{E}(x,y)e^{-jk_u z} \quad (4.13)$$

$$\mathbf{H}_u = \mathbf{E}(x,y)e^{-jk_u z} \quad (4.14)$$

$$\mathbf{E}_p = \mathbf{E}(x,y)e^{-jk_u z} \quad (4.15)$$

$$\mathbf{H}_p = \mathbf{E}(x,y)e^{-jk_u z} \quad (4.16)$$

The perturbed permittivity distribution is related to the original one by

$$\epsilon_p(x,y) = \epsilon_u(x,y) + \Delta\epsilon(x,y) \quad (4.17)$$

From [46] the change in the phase constant due to the perturbation is given by

$$k_p - k_u = \omega \frac{\int\int_{A_\infty} \Delta\epsilon(x,y)\mathbf{E}_p\mathbf{E}_u^* dA}{\int\int_{A_\infty} (\mathbf{E}_u^* \times \mathbf{H}_p + \mathbf{E}_p \times \mathbf{H}_u^*)\mathbf{z} dA} \quad (4.18)$$

where $\mathbf{z}$ is the unit vector in the z-direction and A_∞ is the infinite transversal cross-sectional area. In the case of small perturbations $\Delta\epsilon$ the perturbed fields can be taken as the unperturbed ones. Under this assumption, and writing $\Delta\epsilon$ in terms of refractive indices,

$$\Delta\epsilon(x,y) = \epsilon_0(n_p^2(x,y) - n_u^2(x,y)) \quad (4.19)$$

one can obtain the following expression for the change in effective index:

$$n_{eff,p} - n_{eff,u} = \sqrt{\frac{\epsilon_0}{\mu_0}} \frac{1}{4P_{ave}} \int\int_{A_\infty} (n_p^2 - n_u^2)|\mathbf{E}_u|^2 dA, \quad (4.20)$$

[4] In these equations we have knowingly neglected perturbations to k_c which are always present due to the Kramers-Kronig relations

where P_{ave} is the time-averaged real power carried by the mode

$$P_{ave} = \frac{1}{2}\int\int_{A_\infty} Re(\mathbf{E}_u \times \mathbf{H}_u^*).\mathbf{z}dA \tag{4.21}$$

For the case when a perturbation h_c, with $h_c << n_c$, is applied to the refractive index of the carrier fluid occupying area A_c

$$n_p^2(x,y) - n_u^2(x,y) = \begin{cases} (n_c + h_c)^2 - n_c^2 & \text{for } x,y \text{ in } A_c, \\ 0 & \text{elsewhere} \end{cases} \tag{4.22}$$

equation 4.20 simplifies to

$$n_{eff,p} - n_{eff,u} = \sqrt{\frac{\epsilon_0}{\mu_0}}\frac{(n_a^2 - n_u^2)}{4P_{average}}\int\int_{A_a} |\mathbf{E}_u|^2 dA, \tag{4.23}$$

If the perturbation corresponds to the appearance with an adlayer of transversal cross-sectional area A_a and index n_a

$$n_p^2(x,y) - n_u^2(x,y) = \begin{cases} n_a^2 - n_u^2 & \text{for } x,y \text{ in } A_a, \\ 0 & \text{elsewhere} \end{cases} \tag{4.24}$$

equation 4.20 simplifies to

$$n_{eff,p} - n_{eff,u} = \sqrt{\frac{\epsilon_0}{\mu_0}}\frac{n_c h_c}{4P_{average}}\int\int_{A_a} |\mathbf{E}_u|^2 dA, \tag{4.25}$$

From the above it is observed that $\Delta n_{eff} > 0$ if $n_a^2 > n_u^2$, which is generally the case for biosensors. Maximizing Δn_{eff} clearly means maximizing the overlap between the adlayer area A_a and the modal area, similar considerations apply to the bulk sensitivity.

The rather cumbersome equations 4.23 and 4.25 could be used to determine the surface and bulk sensitivities, respectively, instead of the finite-difference approximations given by equations 4.11 and 4.12, we will use the latter because they are much easier to compute.

Figure 4.19 shows the dependence of the effective index of both surface plasmon modes on the effective index of the cladding. In line with previous results, the top surface plasmon mode displays a linear dependence on top cladding index refractive index (at $1.55\ \mu m$, $\partial n_{eff}/\partial n_c \approx 1.032$, $\partial k_{eff}/\partial n_c \approx 7.49 \times 10^{-9}$), while the bottom mode is hardly affected by top cladding index variations (at $1.55\ \mu m$, $\partial n_{eff}/\partial n_c \approx 0.011$, $\partial k_{eff}/\partial n_c \approx -0.0026$). Quantitatively speaking, the top surface plasmon mode is approximately 1000 times as sensitive to refractive index changes than the bottom surface plasmon mode.

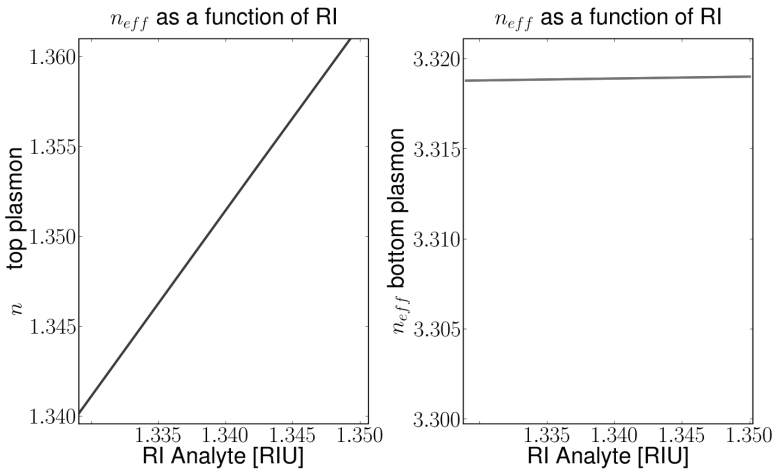


Figure 4.19: Dependence of n_{eff} of both the top and bottom surface plasmon mode on top cladding refractive index.

If a very thin layer is adsorbed at the surface the picture becomes more complicated. While the variation of the effective index of the top surface plasmon mode is still larger than the change in effective index of the bottom surface mode, the ratio between both is smaller and is dependent on the thickness of the adsorbed layer. For thin layers (e.g. $5\ nm$) the top surface plasmon mode is about 20 times more sensitive, the thicker the layer gets, the more the ratio between both sensitivities evolves to the bulk ratio of 1000. This is due to the fact that the bottom surface plasmon mode does have field components on the outside of the Au layer, although these components are limited to a very small region near the surface. The imaginary part of the effective index shows a similar behavior. Both effects get worse as the thickness of the Au layer decreases (Fig. 4.20).

4.4.3 MZI phase sensitivity

The prototypical transducer element considered in this section is the Mach-Zehnder interferometer. Interferometric sensors use a differential phase-based detection scheme for sensing, meaning that they require a reference wave (the bottom surface plasmon mode) to mix with the signal (the top surface plasmon mode) in order to translate phase information into detectable intensity information. The two most common ways of providing a reference signal is the dual-arm Mach-Zehnder configuration on the one hand, and multi mode sensors, where reference and signal propagate in the same

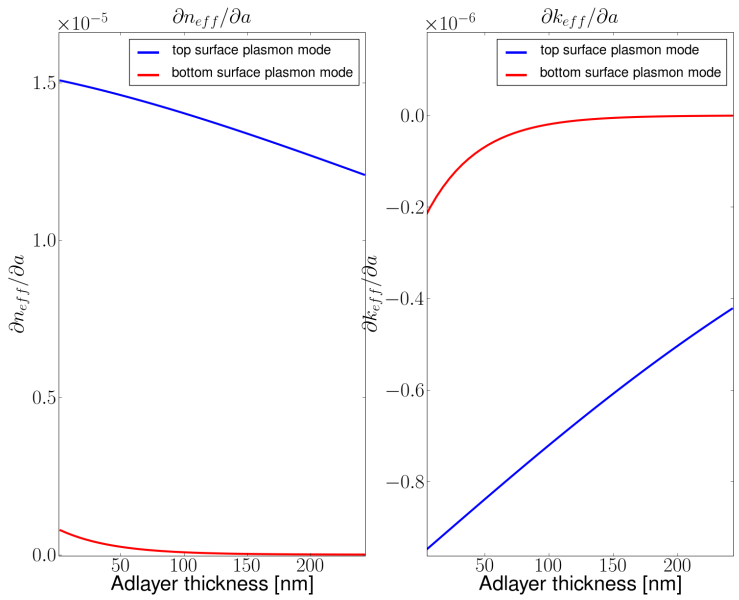


Figure 4.20: Derivatives $\partial n_{eff}/\partial a$ and $\partial k_{eff}/\partial a$ for the top and bottom surface plasmon mode. Simulation wavelength was $1.55\ \mu m$, the refractive index of the adlayer was 1.34, and the thickness of the Au layer equals $50\ nm$.

physical channel but in different modes, on the other hand. The device described in this thesis clearly is a planar waveguide implementation of the latter kind.

Various detection schemes have been reported in literature ranging from intensity, phase, coupling angle to spectral interrogation. While we will briefly discuss the intensity mode, we will mainly focus on the spectral technique, where a broad bandwidth light source or tunable laser is used and variations in the sensed material are detected by measuring changes in the wavelength dependence of the spectral power transfer function.

The relationships derived in this section will hold for a generic MZI implemented with attenuating waveguides, including dielectric waveguides that have losses, and variable power coupling ratios.

4.4.3.1 Intensity Interrogation

The output power of a single-output MZI, in the most general case where the reference and sensing waveguides have different attenuations α_r and α_s, and different power coupling factors $|T_r|^2$ and $|T_s|^2$ respectively, is given by [5]

$$P_{out} = P_{in}(|T_s|^4 e^{-2\alpha_s L} + |T_r|^4 e^{-2\alpha_r L})(1 + V\cos\Phi) \tag{4.26}$$

where the MZI visibility V is given by

$$V = \frac{(2|T_s|^2|T_r|^2)e^{-\alpha_r L}e^{-\alpha_s L}}{|T_r|^2 e^{-2\alpha_r L} + |T_s|^2 e^{-2\alpha_s L}} \tag{4.27}$$

and the phase Φ can be written as

$$\Phi(\lambda, n_c) = 2\pi L \frac{n_{eff,s}(\lambda, n_c) - n_{eff,r}(\lambda, n_c)}{\lambda} \tag{4.28}$$

Since P_{out} is the only measurable in this structure, we can define the MZI phase sensitivity as $\partial P_{out}/\partial\Phi$, which works out to

$$\frac{\partial P_{out}}{\partial \Phi} = -P_{in}(|T_s|^4 e^{-2\alpha_s L} + |T_s|^4 e^{-\alpha_r L})V\sin\Phi \tag{4.29}$$

As is evident from the above (and well known in literature [19, 47]), maximum sensitivity occurs for $\Phi = p\pi/2$ with $p = \pm 1, \pm 3, \pm 5, \ldots$ and sensitivity vanishes for $\phi = q\pi$ with $q = \pm 0, \pm 1, \pm 2, \pm 3, \ldots$. From equation 4.29, it is immediately apparent that the waveguide losses introduce a *power penalty* relative to lossless waveguides in the sense that P_{in} will need to be increased in order to recover the lossless MZI phase sensitivity.

The total sensitivity of a Mach-Zehnder interferometer operating in intensity interrogation mode can then be written as

$$\begin{aligned} S_{bulk} &= \frac{\partial P_{out}}{\partial n_c} = \frac{\partial P_{out}}{\partial \Phi}\frac{\partial \Phi}{\partial n_c} \\ &= -P_{in}(|T_s|^4 e^{-\alpha_s L} + |T_s|^4 e^{-\alpha_s L}) \\ &\quad V\sin\Phi\frac{2\pi L}{\lambda}\left(\frac{\partial n_{eff,s}(\lambda, n_c)}{\partial n_c} - \frac{\partial n_{eff,r}(\lambda, n_c)}{\partial n_c}\right) \end{aligned} \tag{4.30}$$

[5] For notational convenience, and to keep the following description as general as possible, we will use α_s and α_r instead of k_i^t and k_i^b, T_{0b} and $T_{0t,0b}$ are replaced by T_r and T_s.

for bulk sensing, in units $[W/RIU]$ and

$$\begin{aligned} S_{surface} &= \frac{\partial P_{out}}{\partial a} = \frac{\partial P_{out}}{\partial \Phi}\frac{\partial \Phi}{\partial a} \\ &= -P_{in}(|T_s|^4 e^{-\alpha_s L} + |T_r|^4 e^{-\alpha_r L}) \\ &\quad V \sin\Phi \frac{2\pi L}{\lambda}\left(\frac{\partial n_{eff,s}(\lambda, a)}{\partial a} - \frac{\partial n_{eff,r}(\lambda, a)}{\partial a}\right) \end{aligned} \tag{4.31}$$

for surface sensitivity, in units $[W/nm]$

After extensive device optimization, we can calculate the the sensitivity of the surface plasmon interferometer setup can be as high as 10000 dB/RIU. In conjunction with an optoelectronic system which can measure changes in the optical power of 0.01 dB, the detection limit (DL) could be as small as 10^{-6}, well within range of the detection limits of the various SPR approaches listed in table 4.2. Although this value is quite satisfactory, the intensity interrogation approach does not lend itself easily to advanced data-analysis methods since only one single datapoint is available per measurement. Tackling noise issues is therefore quite difficult using this approach.

4.4.3.2 Spectral Interrogation

In spectral interrogation sensors the output power is detected for different values of the source wavelength λ in a specific, given range. Following small changes in the surrounding parameters that change the modal propagation (refractive index cladding layer, temperature or pressure variations, ...), spectral shifts are detected. Specifically for refractive index variations, the sensitivity can be defined in units of nm/RIU by

$$S_{bulk} = \frac{\Delta\lambda}{\Delta n_c}, \tag{4.32}$$

where Δn_c is change of bulk refractive index of the sensed material and $\Delta\lambda$ represents the corresponding spectral shift at constant output power. Surface sensitivity can be defined as

$$S_{surface} = \frac{\Delta\lambda}{\Delta a}, \tag{4.33}$$

For an interferometric sensor where the light propagates in two different paths (albeit physically separated or modally separated), and assuming that the amplitude of both modes is not affected by changes in wavelength in the vicinity of the working point, the peaks (both minima and maxima) in the

spectral power transfer function are again given by equation 4.28. Using implicit function differentiation and assuming that $\partial\Phi(\lambda, n_c) \neq 0$ at the working wavelength λ, the sensitivity can be expressed as a function of $n_{eff,s,r}$, λ and n_c, under the constraint of a constant phase difference as

$$S_{bulk} = \frac{\partial\lambda}{\partial n_c} = \frac{\partial\Phi(\lambda, n_c)/\partial n_c}{\partial\Phi(\lambda, n_c)/\partial\lambda} \tag{4.34}$$

for bulk sensitivity

$$S_{surface} = \frac{\partial\lambda}{\partial a} = \frac{\partial\Phi(\lambda, a)/\partial a}{\partial\Phi(\lambda, n_c)/\partial\lambda} \tag{4.35}$$

for surface sensitivity. Examining both expressions shows the possibility of a divergence in sensitivity; i.e. the sensitivity can become infinite when the denominator of eq. 4.34 approaches zero [48]. Explicitly, this condition will be reached if the group velocity v_g of both contributing modes is equal for the working point wavelength λ and a given n_c. This can be seen in the following derivation

$$\begin{aligned}
\frac{\partial\Phi(\lambda, n_c)}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\left[\frac{2\pi L}{\lambda}(n_{eff,s} - n_{eff,r})\right] \\
&= \frac{2\pi L}{\lambda^2}\left[\left(n_{eff,s} - \lambda\frac{\partial n_{eff,s}}{\partial\lambda}\right) - \left(n_{eff,r} - \lambda\frac{\partial n_{eff,r}}{\partial\lambda}\right)\right] \\
&= \frac{2\pi L}{\lambda}\left[n_{g,s} - n_{g,r}\right] \\
&= \frac{2\pi Lc}{\lambda^2}\left(\frac{1}{v_{g,s}} - \frac{1}{v_{g,r}}\right)
\end{aligned} \tag{4.36}$$

In order to maximize spectral interrogation bulk and surface sensitivity, the two modes need to have quasi similar group velocities. For a Mach-Zehnder interferometer with similar arms, this condition can easily be satisfied. In a surface plasmon interferometer setup like the one proposed in this work, the difference in physical origin between both interfering modes (a purely bound surface plasmon mode versus as hybrid dielectric-leaky surface plasmon mode), makes this a quasi impossible condition to fulfill. In this section we will take a look at sensitivity as a function of the thickness of the Si slab waveguide, and determine ideal design parameters.

Bulk sensitivity S_{bulk} as a function of the thickness of the Si slab waveguide was calculated for several thicknesses of the Au layer. Results are shown in figure 4.21. While the group index of the top surface plasmon mode does not change as the Si slab thickness gets smaller, the bottom

surface plasmon mode's group index n_g does. At first, moving from a 220 nm Si slab waveguide towards thinner waveguides, n_g increases as the mode gets more dispersive (as can be seen in Fig. 4.22), accordingly, sensitivity drops, later however, as the mode slowly evolves into a purely bound surface plasmon mode largely confined to the SiO_2 cladding layer, the group index of the bottom surface plasmon mode becomes smaller. Since the difference between the two group indices becomes smaller, the sensitivity becomes substantially larger.

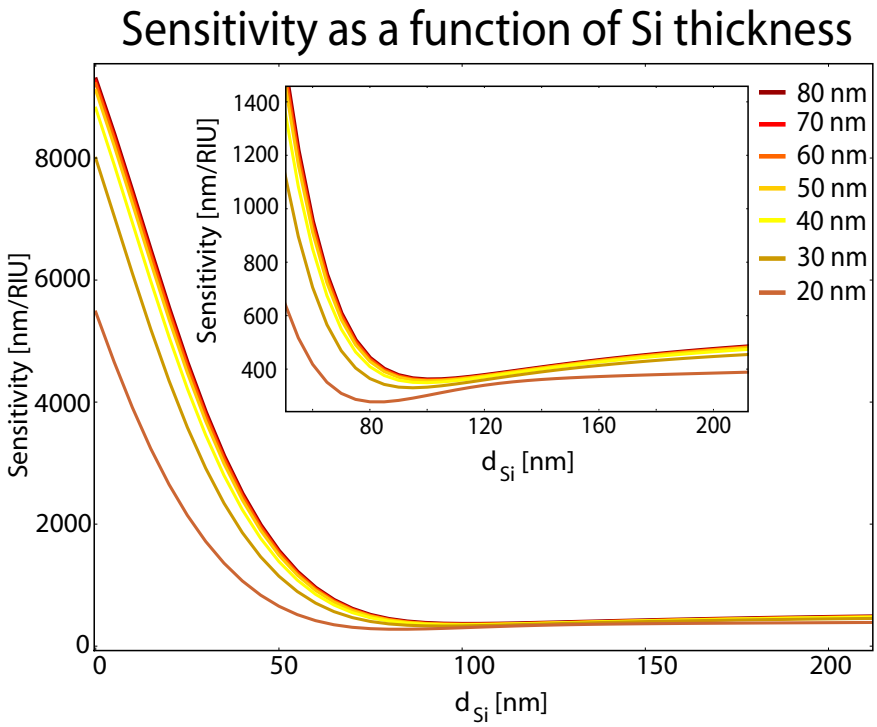


Figure 4.21: Bulk refractive index sensitivity S_{bulk} as a function of Si slab waveguide thickness and Au layer thickness. Inset shows the region between 60 and 220 nm in more detail.

Theoretically, bulk refractive index sensitivity and Si waveguide thickness are inversely proportional in the region $0 < d_{Si} < 50nm$, however from Fig. 4.15, one can see that the overall transmission through a device with such a small Si slab waveguide thickness is noticeably smaller than for devices with a thicker Si slab waveguide thickness. Finding the right balance between bulk sensitivity, interference visibility and overall power

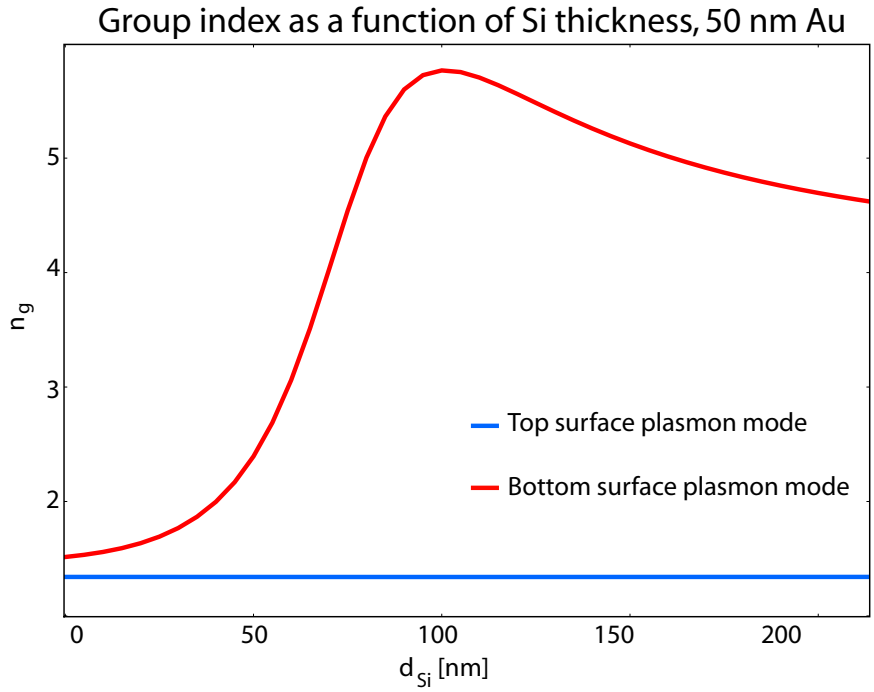


Figure 4.22: Calculated group index n_g of the two surface plasmon modes as a function of Si thickness. As the thickness of the Si layer decreases the difference between n_g decreases, leading to an increase in sensitivity S_{bulk}

transmission is a difficult design exercise. In this work we opted to maximize interference visibility V, in order to easily detect wavelength shifts. The dependence of the bulk refractive index sensitivity on the thickness of the Au layer is negligible, as long as d_{Au} is larger than 30 nm, the effect of the length of the sensing section is also quite small compared to the dependence on d_{Si}.

The device described in this work, while a compromise between sensitivity and output extinction ratio, still has a sensitivity value approximately be equal to 400 nm/RIU. To attain a theoretical detection limit of 10^{-6} RIU, the minimal detectable wavelength shift should be around 400 fm, which is obviously a technological challenge. This is already an indication of the need of specialized data fitting algorithms that take the entire spectrum into account instead of relying on one single measurement.

Surface sensitivity $S_{surface}$ in pm/nm is plotted in Fig. 4.23 as a function of d_{Si} and d_{Au} in Figs. 4.23 and 4.24. The same trend that was

apparent for the bulk sensitivity S_{bulk} is manifested in these graphs, the underlying reason is again the group index of both surface plasmon modes. For very thin gold layers (less than 20 nm), $S_{surface}$ can become negative, due to the fact that the bottom surface plasmon mode is no longer confined to the bottom surface of the Au layer and hence its effective index n_{eff} will be influenced by the adsorbed layer. Thicker Au layers however, will not display this effect.

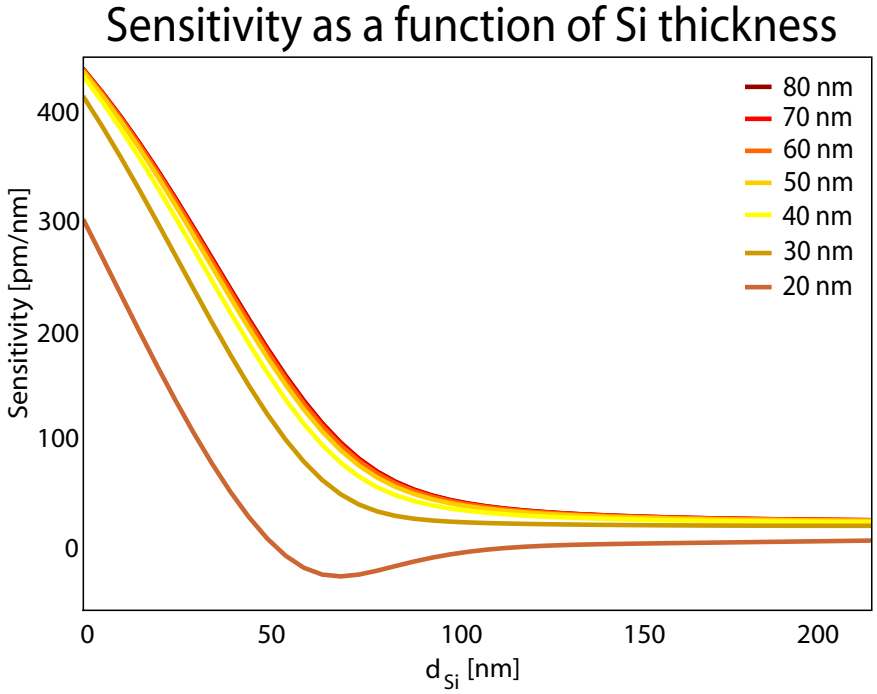


Figure 4.23: Surface sensitivity $S_{surface}$ as a function of Si slab waveguide thickness and Au layer thickness. The refractive index of the adlayer in this calculation is 1.34.

To summarize, while the bulk detection limit can be as high as 10^{-6} RIU (sensitivity value of 10000 dB/RIU), the approach is very sensitive to noise sources. A similar detection limit using spectral interrogation would require the detection of sub-pm wavelength shifts (sensitivity value of 400 nm/RIU), but the entire spectrum can be used to approximate this wavelength shift, potentially limiting detrimental effects of noise sources.

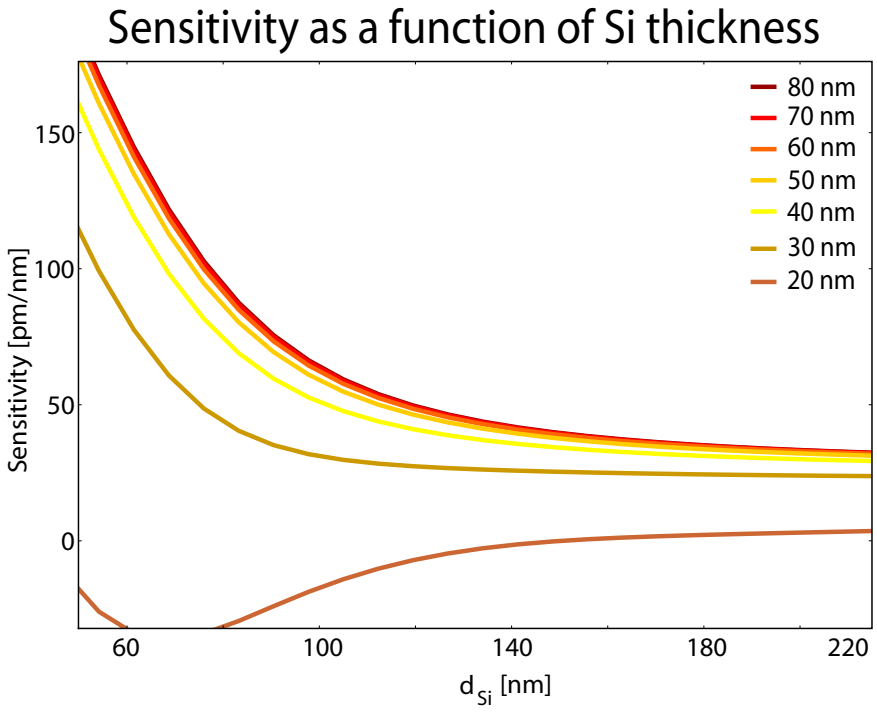


Figure 4.24: Surface sensitivity $S_{surface}$ as a function of Si slab waveguide thickness and Au layer thickness (zoom of the region between 50 and 220 nm). The refractive index of the adlayer in this calculation is 1.34.

References

[1] W. Chang, D. Akin, M. Sedlak, M. R. Ladisch, and R. Bashir. *Poly(dimethylsiloxane) (PDMS) and silicon hybrid biochip for bacterial culture*. Biomedical Microdevices, 5(4):281–290, 2003.

[2] M. A. Cooper. *Optical biosensors in drug discovery*. Nature Reviews Drug Discovery, 1(7):515–528, 2002.

[3] B. Luff, J. Wilkinson, J. Piehler, U. Hollenbach, J. Ingenhoff, and N. Fabricus. *Integrated optical Mach-Zehnder biosensor*. Journal of Lightwave Technology, 16:583–592, 1998.

[4] F Prieta, Sepúlveda, A. Calle, A. Llobera, C. Dominguez, A. Abad, A. Montoya, and L. M. Lechuga. *An integrated optical interferometric nanodevice based on silicon technology for biosensor applications*. Nanotechnology, 14:907–912, 2003.

[5] X. Fan, I. M. White, S. I. Shopova, H. Zhu, J. D. Suter, and Y. Sun. *Sensitive optical biosensors for unlabeled targets: a review*. Anal. Chem. Acta, 620(8):8–26, 2008.

[6] T. Nikolajsen, K. Leosson, I. Salakhutdinov, and S. Bozhevolnyi. *Polymer-based surface-plasmon-polariton stripe waveguides at telecommunication wavelengths*. Applied Physics Letters, 82(5):1510–1518, 2003.

[7] M. Hochberg, T. Baehr-Jones, C. Walker, and A. Scherer. *Integrated plasmon and dielectric waveguides*. Optics Express, 12:5481–5486, 2002.

[8] W. Bogaerts. *Nanophotonic Waveguides and Photonic Crystals in Silicon-on-Insulator*. PhD thesis, Ghent University, 2004.

[9] J. Homola, S. Y. Sinclair, and G Gauglitz. *Surface plasmon resonance sensors: review*. Sensors and Actuators B -Chemical, 54:3–15, 1999.

[10] J. Homola, I. Koudela, and S. Y. Sinclair. *Surface plasmon resonance sensors based on diffraction gratings and prism couplers: sensitivity comparison*. Sensors and Actuators B -Chemical, 54:16–24, 1999.

[11] P. Berini. *Plasmon-polariton modes guided by a metal film of finite width*. Optics Letters, 24(15):1011–1013, 1999.

[12] P. Berini. *Plasmon-polariton modes guided by a metal film of finite width bounded by different dielectrics*. Optics Express, 7(10):329–335, 2000.

[13] P. Berini. *Plasmon-polariton waves guided by thin lossy metal films of finite width: Bound modes of asymmetric structures*. Physical Review B, 63(12):art. no.–125417, 2001.

[14] P. Berini, R. Charbonneau, N. Lahoud, and G. A. Mattiussi. *Characterization of long-range surface-plasmon polariton waveguides*. Journal of Applied Physics, 98:043109–043109–12, 2005.

[15] P. Berini. *Figures of merit for surface plasmon waveguides*. Optics Express, 14(26):13030–13042, 2006.

[16] Rashid Zia, Jon A. Schuller, Anu Chandran, and Mark L. Brongersma. *Plasmonics: the next chip scale technology*. Materials Today, 9(7-8):20–27, 2006.

[17] R. Zia, M. D. Selker, and M. L. Brongersma. *Leaky and bound modes of surface plasmon waveguides*. Physical Review B, 71(16):165431, 2005.

[18] X. Daxhelet, J. Bures, and R. Macejko. *Temperature-independent all-fiber modal interferometer*. Optical Fiber Technology, 1:373–376, 1995.

[19] P. Berini. *Bulk and surface sensitivities of surface plasmon waveguides*. New Journal of Physics, 10:105010, 2008.

[20] J. Čtyrocký, J Homola, P.V. Lambeck, S. Musa, H.J.W.M. Hoekstra, R.D. Harris, J.S. Wilkinson, and N.M. Usievich, B.and Lyndin. *Theory and Modelling of Optical Waveguide Sensors Utilising Surface Plasmon Resonance*. Sensors and Actuators B -Chemical, 54:66–73, 1999.

[21] J. Homola, J. Čtyrocký, M. Skalský, J. Hradilová, and P. Kolářová. *A Surface Plasmon Resonance Based Integrated Optical Sensor*. Sensors and Actuators B -Chemical, 38-39:286–290, 1997.

[22] J Dostálek, J Čtyrocký, J Homola, E Brynda, M Skalský, P Nekvindová, J Špiriková, J Škvor, and J Schröfel. *Surface plasmon resonance biosensor based on integrated optical waveguide*. Sensors and Actuators B -Chemical, 76:8–12, 2001.

[23] F.J. Bueno, O. Esteban, N. Díaz-Herrara, M.C. Navarrete, and A. González-Cano. *Sensing properties of asymmetric double-layer-covered tapered fibers*. Applied Optics, 43:1615–1620, 2004.

[24] O. Parriaux. *Normalized analysis for the sensitivity optimization of integrated optical evanescent-wave sensors*. Journal of Lightwave Technology, 16(4):573–582, 1998.

[25] J. Homola. *On the sensitivity of surface plasmon resonance sensors with spectral interrogation*. Sensors and Actuators B -Chemical, 41(3):207–211, 1997.

[26] R. Ince and R. Narayanaswamy. *Analysis of the performance of interferometry, surface plasmon resonance and luminescence as biosensors and chemosensors*. Analytica Chimica Acta, 569(2):1–20, 2006.

[27] P. Pfeifer, U. Aldinger, G. Schwotzer, S. Diekmann, and P. Steinrucke. *Real time sensing of specific molecular binding using surface plasmon resonance spectroscopy*. Sensors and Actuators B -Chemical, 54(1-2):166–175, 1999.

[28] C. A. Barrios, K.B. Gylfason, B. Sánchez, A. Griol, H. Sohlström, and M. Heolgado. *Slot waveguide biochemical sensor*. Optics Letters, 32:3080–3082, 2007.

[29] L. Rindorf, J. B. Jensen, M. Dufva, L. H. Pedersen, P. E. Hoiby, and O. Bang. *Photonic crystal fiber long-period gratings for biochemical sensing*. Optics Express, 14(18):8224–8231, 2006.

[30] M. Sumetsky, R. S. Windeler, Y. Dulashko, and X. Fan. *Optical liquid ring resonator sensor*. Optics Express, 15:14376–14381, 2007.

[31] E. Chow, A. Grot, L. W. Mirkarimi, M. Sigalas, and G. Girolami. *Ultracompact biochemical sensor built with two-dimensional photonic crystal microcavity*. Optics Letters, 29(10):1093–1095, 2004.

[32] T. Allsop, R. Neal, S. Rehman, D. J. Webb, D. Mapps, and I. Bennion. *Generation off infared surface plasmon resonances with high refractive index sensitivity utilizing titled fiber Bragg grafings*. Applied Optics, 46(22):5456–5460, 2007.

[33] N. M. Hanumegowda, Stica C. J., Patel B. C., I. M. White, and X. Fan. *Refractometric sensors based on microsphere resonators*. Applied Physics Letters, 87:201107, 2005.

[34] I. M. White and X. Fan. *On the performance quantification of resonant refractive index sensors*. Optics Express, 16(2):1020–1028, 2008.

[35] H. Zhu, I. M. White, J. D. Suter, P. S. Dale, and X. Fan. *Analysis of biomolecule detection with optofluidic ring resonator sensors*. Optics Express, 15(15):9139–9146, 2007.

[36] K. Kurihara, H. Ohkawa, Y. Iwasaki, O. Niwa, T. Tobita, and K. Suzuki. *Fiber-optic conical microsensors for surface plasmon resonance using chemically etched single-mode fiber*. Analytica Chimica Acta, 523(2):165–170, 2004.

[37] M. Piliarik, J. Homola, Z. Manikova, and J. Ctyroky. *Surface plasmon resonance sensor based on a single-mode polarization-maintaining optical fiber*. Sensors and Actuators B -Chemical, 90(1-3):236–242, 2003.

[38] P. Stöcker, B. Menges, U. Langbein, and S. Mittler. *Multimode waveguide mode surface plasmon coupling: a sensitivity and device realizability study*. Sensors and Actuators A -Physical, 116(2):224–231, 2004.

[39] C. W. Lin, K. P. Chen, C. N. Hsiao, S. M. Lin, and C. K. Lee. *Design and fabrication of an alternating dielectric multi-layer device for surface plasmon resonance sensor*. Sensors and Actuators B -Chemical, 113(1):169–176, 2006.

[40] A. Suzuki, J. Kondoh, Y. Matsui, S. Shiokawa, and K. Sozuki. *Development of novel optical waveguide surface plasmon resonance (SPR) sensor with dual Light Emitting Diodes*. Sensors and Actuators B -Chemical, 106:383–387, 2005.

[41] P. I. Nikitin, A. A. Beloglazov, V. E. Kochergin, M. V. Valeiko, and T. I. Ksenevich. *Surface plasmon resonance interferometry for biological and chemical sensing*. Sensors and Actuators B -Chemical, 54(1-2):43–50, 1999.

[42] P. I. Nikitin and A. A. Beloglazov. *A multipurpose sensor based on surface plasmon polariton resonance in a Schottky structure*. Sensors and Actuators A -Physical, 42(1-3):547–552, 1994.

[43] M. Piliarik, H. Vaisocherova, and J. Homola. *Towards parallelized surface plasmon resonance sensor platform for sensitive detection of oligonucleotides*. Sensors and Actuators B -Chemical, 121(1):187–193, 2007.

[44] *http://www.sensata.com/contact/spreeta-analytical-sensor.htm*.

[45] X. D. Hoa, A. G. Kirk, and M. Tabrizian. *Towards integrated and sensitive surface plasmon resonance biosensors: A review of recent progress*. Biosensors & Bioelectronics, 23(2):151–160, 2007.

[46] R. F. Harrington. *Time-harmonic electromagnetic fields*. IEEE Press Series on Electromagnetic Wave Theory. Wiley-IEEE Press, New York, 2001.

[47] R. Levy, A. Peled, and S. Ruschin. *Waveguided SPR sensor using a Mach-Zehnder interferometer with variable power splitting ratio*. Sensors and Actuators B -Chemical, 119:20–26, 2006.

[48] Ronen Levy and Shlomo Ruschin. *Critical sensitivity in hetero-modal interferometric sensor using spectral interrogation*. Optics Express, 16(25):20516–20522, 2008.

5 Fabrication

THE following sections give an overview of fabrication processes used in this work, and motivates specific choices that were made during the development cycle. Figure 5.1 shows the entire fabrication process. The Si-etch process is a basically a two step process. The first step is generally a shallow etch step (RIE-ICP etch), meaning that 70 nm of the Si layer is etched away (*shallow etch*). After removing the resist pattern, a second step consists in applying a resist layer for the second time, defining structures in this resist layer lithographically and an additional etch step, this time all the way through the Si layer (*deep etch*). For this work, the sequence of shallow and deep etch was reversed, and the shallow etch resist layer was not removed from the wafer, as it can be used as the masking layer for Au evaporation. After dicing and O_2 plasma cleaning, a thin Au layer (37 nm) is deposited using Joule evaporation. The protective Benzocyclobutene (BCB) layer is then spin-coated on the surface of the sample. A small rectangular region above the surface plasmon interferometer structures is exposed and subsequently developed and cured to form a stable interlinked layer.. After processing, the left and right part of the samples are polished to obtain nice facets for end-fire coupling. Then the devices are cleaned with an oxygen plasma cleaning step. Standard cleaning methods for removing organic residues such as RCA-1[1] and RCA-2 [1] cleaning steps strip the

[1]The RCA-1 clean (sometimes called *standard clean-1*, SC-1), developed by Werner Kern at RCA (Radio Corporation of America) laboratories in the late 1960's, is a procedure

Au layers from the Si since Au-Si adhesion is relatively weak. After this cleaning step devices are either ready to be measured in a bulk refractive index sensing experiment, or, for biosensing purposes, are ready to be chemically modified. Each of these steps will, in order of appearance, be discussed in this chapter.

5.1 Silicon-on-Insulator Deep UV Lithography

The technology for Silicon-on-Insulator nanophotonic integrated circuits was developed in the PhD work of W. Bogaerts [2] and is carried out at IMEC, Leuven. The whole process flow uses standard CMOS fabrication methods on SOI wafers having a 220 nm top guiding Silicon layer and a 2 μm buried oxide layer. Deep UV projection lithography at 193 nm, and high numerical aperture optics, are used to define the patterns. The patterns are etched into the Silicon layer with the resist pattern as a mask using low-pressure, high density $Cl_2/O_2/He/HBr$ ICP-RIE (see 5.2.2). A two-step etch process has been developed in order to obtain shallow etched parts (70 nm) and deeply etched parts (220 nm, full silicon top layer etch). The alignment accuracy between both lithography steps is of the order of 50 nm.

Silicon wafers with a diameter of $200mm$ were used as the substrate to build an $a-Si:H$[2] photonic integrated circuit. A 2 μm silicon dioxide layer (BOX) was deposited using a high density PECVD process followed by a chemical mechanical polish (CMP). Roughness after polishing was measured using atomic force microscopy (AFM), and was found to be smaller than 0.1 nm, this can be translated to bottom surface roughness for the $a-Si$ thin film. A 220 nm layer of $a-Si:H$ was deposited over the BOX by PECVD in a commercial parallel plate reactor. Silane (SiH_4) was used as the precursor gas, combined with Helium (He) for dilution. To obtain a good quality a-Si film, the He/SiH_4 ratio should equal 9 and a power of 180 W was used, while the deposition temperature was fixed at 300 degrees. More information with regards to a-Si:H process development

for removing organic residue and films from silicon wafers. The decontamination works based on sequential oxidative desorption and complexing with $H_2O_2 - NH_4OH - H_2O$ (RCA1). A second RCA-2 clean (SC-2) using $H_2O_2 - HCl - H_2O$ is often used to further clean the surface.

[2]When depositing a-Si using a PECVD process, dangling bonds will be present in the resulting a-Si layer. These dangling bonds cause mid-bandgap absorption in the infra red wavelength range, however, resulting in considerable propagation losses. Incorporating hydrogen during the deposition process can saturate these bonds, the resulting material is called hydrogenated a-Si (a-Si:H) [3].

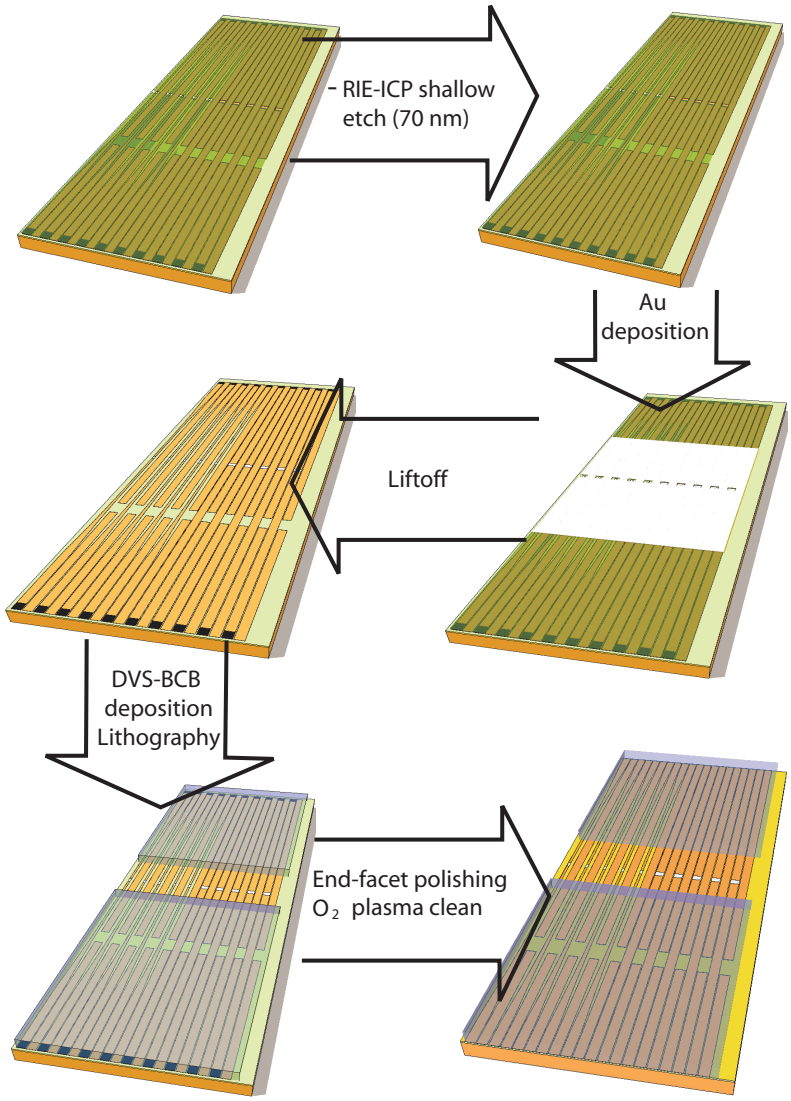


Figure 5.1: Overview of all the steps in the fabrication cycle of a surface plasmon interferometer sensor.

can be found in [3].

5.2 Etching Process

Following lithography, the pattern needs to be defined in the semiconductor material using an etching technique. Since the devices developed in this work required vertical interfaces at both waveguide discontinuities, wet etching methods cannot be used since they lead to isotropic or crystallographical etch anisotropy regardless of crystal orientation and extremely accurate dimensional control due to the combination of both physical and chemical etch components [4]. We will now discuss the underlying physics of dry etching methods, starting with the most basic setup, Reactive Ion Etching (RIE) and then moving on to a more sophisticated, advanced method, Inductive Coupled Plasma Reactive Ion Etching. The latter process was used in this work.

5.2.1 Reactive Ion Etching

A typical Reactive Ion Etching (RIE) configuration is shown in figure 5.2 The vacuum chamber is filled with gases and a plasma is created by RF-excitation of two parallel electrodes. Ions are accelerated to the lower electrode (where the sample is located) and strike it at near-vertical incidence, consequently etching the sample. High process pressure and low voltages results in predominant chemical etching of the substrate. Low pressure and high voltages results in physical ion sputtering of the substrate.

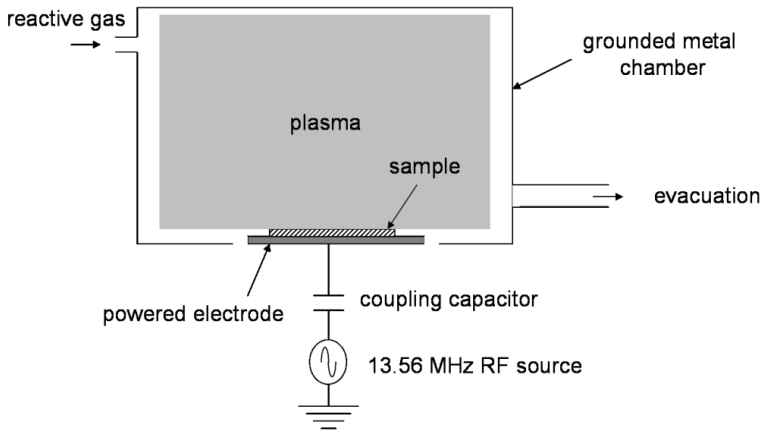


Figure 5.2: Reactive Ion Etching (RIE), from [4]

5.2.2 Inductively Coupled Plasma RIE Etching

Inductively Coupled Plasma RIE, schematically shown n figure 5.3 is a high-density plasma etching technique. The plasma is generated in a chamber that is encircled by an inductive coil, and diffuses plasma from the generating region towards the sample. The plasma has a very high density and low energy, as a result, etching can take place with low damage while still obtaining high etch rates.

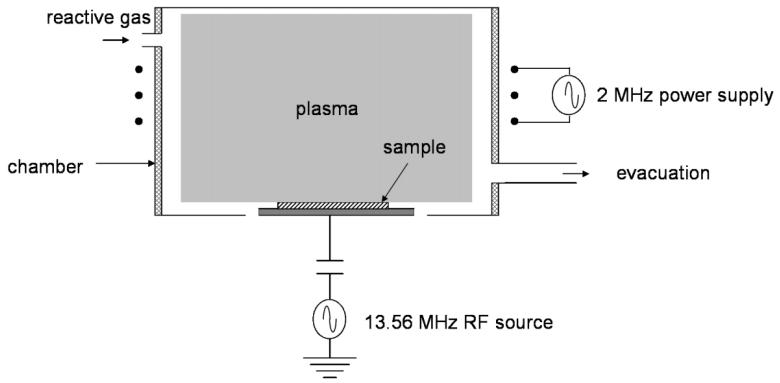


Figure 5.3: Inductively Coupled Reactive Ion Etching (ICP-RIE), from [4]

Since surface plasmon modes propagate at the interface of the Au-Si/H_2O-Au layer, their complex propagation constant is highly dependent on surface roughness. Minimizing surface roughness is key towards low-loss surface plasmon modes. Moreover, surface roughness also determines the quality of the Au film deposited after etching, as shown in figures 5.4 and 5.5. On the left the roughness of the Si surface before Au-deposition is shown, on the right the cross-section of the deposited Au-film is shown.

Using an Oxford ICP Plasmalab System 100 , we have optimized (or attempted to optimize) the etching parameters for Si-etching. Although [5] suggests a $HBr/NF_3/O_2/SF_6$ gas mixture for deep etching Si, and $HBr/Cl_2/O_2$ gas mixture for shallow etching, some of the contributing components in this mixture were not available in our cleanroom at the time.

First etch test were carried out using only SF_6 as etchant, $5\ W$ of ICP power and a variation of the gas pressure in the process chamber and power of the RIE etching process. The SF_6 gas pressure determines the density of the plasma in the chamber, this in turn determines the number of ions that will impact the Si surface and hence the etch rate. Usually, when etching Si, low pressure in the range of $10 - 20\ mTorr$ is preferred since it allows the silicon etch byproducts to easily exit from the increasing trench etch

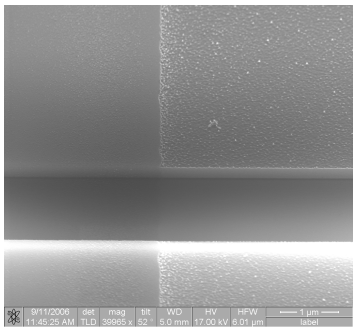


Figure 5.4: 35° SEM view of the surface roughness of an RIE-ICP etched Si surface.

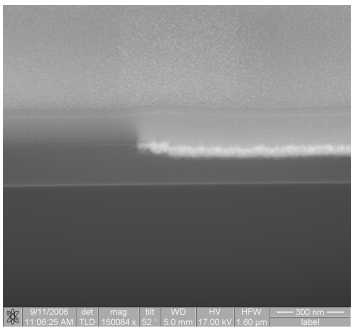


Figure 5.5: SEM picture of an FIB cross-section of the same sample after Au deposition. The Au film quality is deplorable, a direct consequence of the original surface roughness of the surface on which the Au film was deposited.

depth by minimizing the mean free path collisions between exiting etch byproducts and incoming reactive etch ions and maintains a straight sidewall trench profile by controlling undercutting of the mask. The power setting determines how much power the RF supply outputs to the chamber. Etch rate increases with power, and hypothetically, so does surface roughness. RMS-roughness, defined as

$$R_{RMS} = \sqrt{(\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(z(x_m, y_n) - \langle z \rangle))^2} \tag{5.1}$$

was characterized using AFM (Agilent 5500 AFM/SPM in tapping contact mode). A topographical picture of the etched Si-surface is shown in figure 5.6, the RMS roughness as a function of chamber pressure and RIE power is shown in figure 5.7 From figure 5.7 it is clear that surface roughness increases as a function of RIE power and chamber pressure. The smallest value for RMS roughness is $1.53\ nm$, for an RIE power of $10\ W$, and a chamber pressure of $20\ mTorr$. Etch rate at these parameters was found to be too slow for practical purposes. Higher powers lead to more acceptable etch rates, at the cost of more (and unacceptable) surface roughness.

Therefore, for a second run of experiments, we tried varying a different parameter, namely the oxygen content. Fluorine radicals etch silicon spontaneously, the addition of oxygen to sulfurhexafluoride plasma forms a passivating inorganic film (siliconoxyfluoride) on this silicon surface. Therefore, etching directionality is determined by the desorption capabilities of the

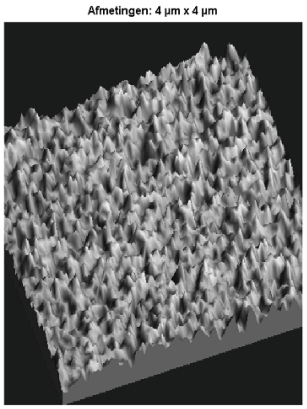


Figure 5.6: Surface Topography (image acquired using an AFM in tapping mode) of etched Si surface, RIE power $20\ W$, chamber pressure $10\ mTorr$.

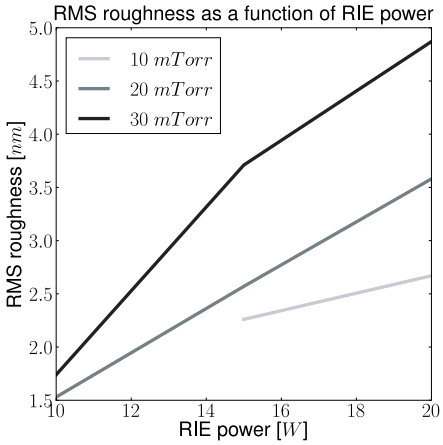


Figure 5.7: RMS roughness as a function of RIE power and chamber pressure.

etching passivation film and the etching product by the energetic ions. The etch rate is therefore determined by the thickness of this passivation layer and by the density of the fluoride radicals, which as the O_2 concentration decreases drops due to dilution [6]. The thickness of the O_2 film is determined by the O-atom density, the ion impact and local temperature. The F atom density is function of the SF_6 flow. Therefore at higher O_2 content the silicon etching rate decreases due to more passivated silicon surface and less available etching species [6]. Characterization of surface roughness was the least of our concerns for this second run, as the introduction of O_2 in the gas mixture severely degrades the step discontinuity at the beginning and end of the sensor region. All attempts at finding a good etch recipe resulted in gradual slopes instead of sharp discontinuities, results can be seen in figure 5.8 and 5.9.

Etching with SF_6 leads to considerable surface roughness, etching with SF_6 and O_2 to gradual discontinuities instead of sharp discontinuities. There was no trade of between surface roughness and discontinuity profile since all samples measured with O_2 showed this gradual profile. Etching with larger RIE powers could potentially solve the discontinuity profile problem, but then the etch rate would become too large to accurately control etch depth.

The fact that the Oxford ICP Plasmalab machine in our cleanroom is not a dedicated Si etching tool, but is most often used for Indium Phosphide

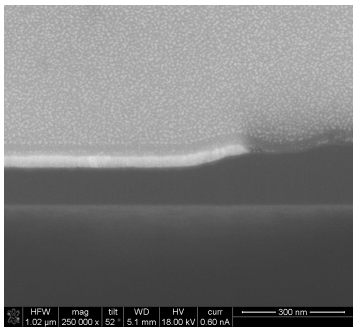


Figure 5.8: SEM picture of the discontinuity at the beginning of the sensor region, RIE/ICP power $14W/5W$*, chamber pressure* $30\ mTorr$*,* SF_6 20 $sccm$*,* O_2 10 $sccm$*.*

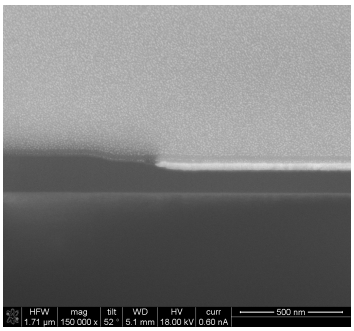


Figure 5.9: SEM picture of the discontinuity at the end of the sensor region, RIE/ICP power $14W/5W$*, chamber pressure* $30\ mTorr$*,* SF_6 20 $sccm$*,* O_2 10 $sccm$*.*

(InP) and Benzoxyclobutene (BCB) etching also had a detrimental impact on the reproducibility of the etching process. All of the above mentioned issues motivated the choice made in this work to etch the sensor region using the Silicon RIE-ICP etching equipment at IMEC. During the shallow etch, the sensor region was etched as well. This resulted in very smooth Si surfaces, with a nice vertical discontinuity profile at both the start and end section of the sensing region. The biggest disadvantage of this approach is the fact that the etch depth is kept fixed at $70\ nm$ (optimal etch depth for the grating couplers). For proof-of-principle measurements however, this is an acceptable compromise since these devices will still have acceptable specifications.

5.3 Thin Film Deposition

In this work thin films (the Au-layers) were deposited using Joule evaporation (Leybold Univex). The method is based on Joule heating of a high resistance material in a vacuum chamber. A current is sent through a crucible containing the metal to be evaporated, the metal melts and the chamber is filled with metal vapor. The metal is then deposited onto the chips which are mounted upside down above the crucible. A quartz crystal, capable of measuring thickness of less than a single atomic layer with 0.5% accuracy, is used to monitor the deposition rate[3].

[3]The quartz crystal oscillates at a resonance frequency that is dependent on the thickness and mass of the film deposited onto it. When a voltage is applied across the faces of a

5.4 BCB Deposition

In this work, the samples were covered with the a photo definable BCB layer, a thin stripe above the SPI sensors was exposed and consequently developed. This will protect the SOI waveguides from damage that could originate from pressing the flowcell onto the chip, furthermore the planarizing capabilities of the BCB layer will ensure a watertight seal at the flowcell-BCB interface.

Divinylsiloxane-bis-benzocyclobutene (DVS-BCB) is a symmetrical molecule consisting of a silicon backbone terminated by two benzocyclobutene rings. Partial curing transforms this monomer into an oligomer, which can be used in a solution by adding mesitylene. The DVS-BCB layer thickness attained by spin coating is determined by the amount of mesitylene added and the degree of polymerization of the oligomer in the solution. Upon curing, the benzocyclobutene ring thermally opens to form o-quinodimethane. This is a very reactive intermediate product which readily undergoes a Diels-Alder reaction with an available vinylsiloxane group, to form a three-dimensional network [7, 8, 9]. During polymerization no by-products are created.

Several adhesion promoter solutions were developed to improve adhesion of the polymer to a wide range of surfaces. Adhesion promoter molecules for use with silicon wafers have the generalized chemical formula $G - Si(OR)_3$ [7], where R can be hydrogen, methyl, ethyl or other more complex groups. The resulting silanol group is reacts with free hydroxyl groups on an oxidized surface. The G-group consists of a moiety with favorable interaction with the polymer.

The most advantageous features of DVS-BCB are the low optical loss at telecommunication wavelengths, the low shrinkage upon cure, its high glass transition temperature (allowing a large post-bonding thermal budget) and its excellent planarization properties. The major drawback is the low thermal conductivity.

In this work the DVS-BCB layer was soft-cured, which leads to a 75% polymerization of the DVS-BCB. A maximum temperature of $180°C$ for 40 minutes was used for soft-curing, under an atmosphere containing less than 100 ppm oxygen, to prevent oxidation of the DVS-BCB. To achieve this, nitrogen was purged through the curing chamber.

DVS-BCB solutions are formulated and commercialized by Dow Chemicals as the CYCLOTENE product series, these have the disadvantage that they are photo-unsensitive, although they can be etched using a gas mixture

properly shaped piezoelectric crystal, the crystal is distorted and changes shape in proportion to the applied voltage. At certain discrete frequencies of applied voltage, a condition of very sharp and repeatable electromechanical resonance is seen.

of SF_6 and O_2. Recently photo-sensitive BCB compounds have become available (commercially known as CYCLOTENE 4000). These compounds can be exposed to UV light, after which the exposed area can be developed using a puddle-development process. The advantage here is that this eliminates the etching step and therefore minimizes potential surface damage that could occur during this etch step.

The refractive index of the BCB layer is $n_{BCB} = 1.628$ [10], this means that the deposition of BCB layers on top of the adiabatic tapers required to couple from broad input waveguides to narrower waveguides will substantially improve the efficiency of these tapers. Low taper efficiency for strip waveguides with asymmetric cladding layers is due to an unwanted anti-crossing behavior of the quasi TM and TE modes in the taper region. Reducing the asymmetry by depositing higher index cladding layers (partly) removes this effect (for a symmetric cladding situation, the problem is non-existing). Further details can be found in appendix B.

5.5 Coupling light in-and out of the device

The coupling problem is an important problem, not just for sensing applications, but for nanophotonics in general. Efficient and broadband coupling of light into or out of photonic chips is not straightforward due to the mismatch between the mode size of an optical fiber and the typical mode size in a photonic chip waveguide. Efficiency is very important for surface plasmon applications due to the high propagation losses in the devices. Also broadband coupling is of importance since the minima in the transmission spectrum are relatively broad. The choice here is between end-fire coupling in plane and out-of-plane coupling with grating couplers optimized for TM polarization.

5.5.1 TM Grating Couplers

Grating couplers are designed to couple light between optical fibers and waveguides efficiently. Coupling efficiency is optimal when the overlap between the field profile of the mode supported by the single mode fiber and the field profile of the field diffracted by the grating is maximal. A very instructive analysis of grating couplers and simple design rules are given by Tamir in [11] and [12].

The grating coupler design was carried out by simulating a 2D cross-section of the grating. For 1D gratings this can be done since the width of the waveguide ($10 - 12\mu m$) is much larger than the height and the wavelength.

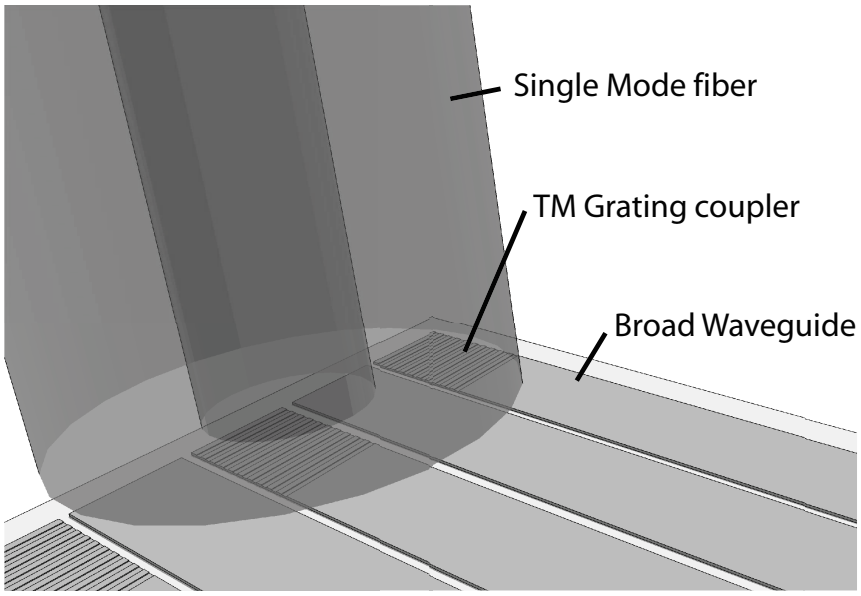


Figure 5.10: Schematical representation of a TM grating coupler.

The method is extensively described in the PhD of Dirk Taillaert [13], we will only summarize the important aspects here.

In the CAMFR simulations the grating is described as an output coupler. This is in no respect limiting, since, if coupling between two modes is considered, the structure can considered to be reciprocal. The fundamental mode of the waveguide is excited and the electric and magnetic field distributions are calculated. Although the fiber itself is not included in the simulation window, coupling efficiency η to fiber is calculated by multiplying the ratio between P_{up} and P_{in} (the power flux radiated upwards and the input power carried by the waveguide mode) by an overlap integral between the upward radiated field profile (at a chosen distance above the grating, typically a few μm) and a Gaussian fiber mode with beam diameter $2w_0 = 10.4\mu m$:

$$\eta = \gamma \times \frac{P_{up}}{P_{in}} \tag{5.2}$$

with

$$\gamma = \left| \frac{\int\int \mathbf{E} \times \mathbf{H}^*_{fiber}}{\int\int \mathbf{E} \times \mathbf{H}^* \int\int \mathbf{E}_{fiber} \times \mathbf{H}^*_{fiber}} \right| \tag{5.3}$$

For a pure exponentially decaying field profile the overlap integral with a Gaussian beam is at most 0.801. Formula 5.2 gives the coupling efficiency of

the simulated 2D problem, which is a good approximation of the 3D coupling problem. By multiplying this value by a correction factor (preferably close to 1.0), the lateral dimension of the waveguide can be taken into account. For a $12\mu m$ wide SOI waveguide, this correction factor is can be calculated to be 0.907 [13]. The coupling efficiency values quoted in the rest of this chapter are the efficiency values for the 2D simulation problem.

5.5.1.1 Coupling efficiency

The coupling efficiency was calculated numerically as a function of grating period and fill factor. While the grating period should fulfill the Bragg-condition, the fill-factor determines the coupling strengths, and thus the bandwidth of the grating coupler. The groove etch depth was chosen to be $70\ nm$, which is the standard etch depth in the shallow-etch CMOS process, the fiber tilt was $10°$. Efficiency plots are shown in figure 5.11.

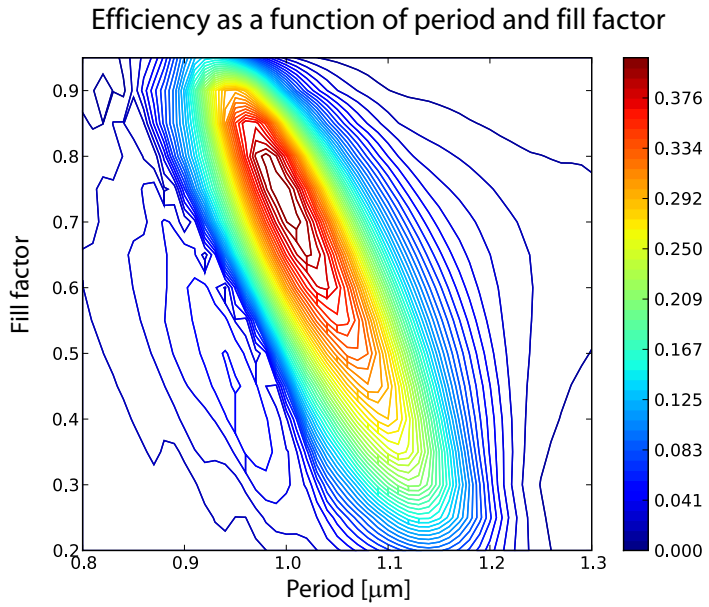


Figure 5.11: Coupling efficiency η as a function of period and fill factor of the grating, groove etch depth $70\ nm$, fiber tilt $10°$, 14 periods

Maximal coupling efficiency is obtained for a period of $990\ nm$ and a fill factor of 0.75. In the here adopted conventions for fill factor this means that the etched grooves form 25% of the total grating coupler surface. The grating periods of 1040 and $1080\ nm$ are mentioned in table 5.1 for comparison with experimental data.

Period	Fill Factor	Efficiency
990	0.75	40.968
1040	0.5	29.455
1080	0.5	32.637

Table 5.1: Maximum coupling efficiencies η, numerically calculated

The fact that the coupling efficiency can be increased by reducing the fill factor indicates that TM polarized light requires a weaker grating than TE polarized light does. To confirm this theory, the coupling efficiency was calculated as a function of period and etch depth for a grating consisting of 12 periods, a fill factor of 0.5 and a fiber tilt of 10 degrees, results are shown in figure 5.12

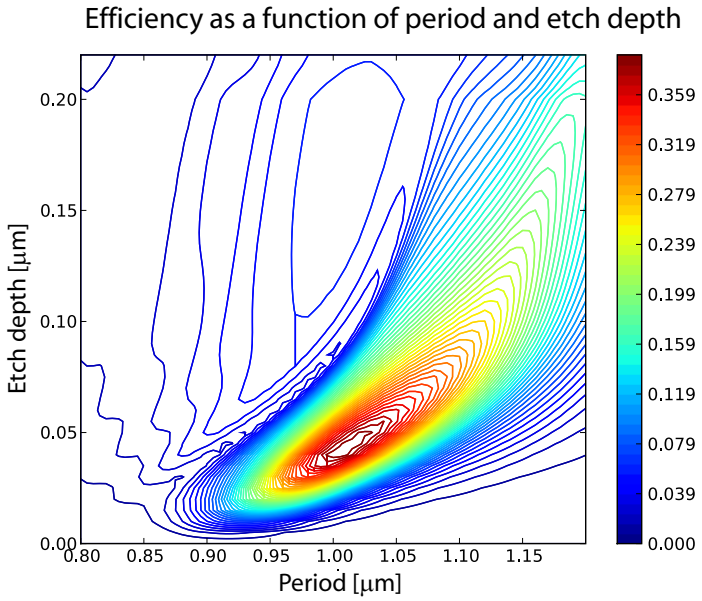


Figure 5.12: Efficiency as a function of period and etch depth of the grating region, fill factor 0.5, fiber tilt 10°, number of periods 14

Period	Etch depth	Efficiency
1005	0.4	39.1487

The fact that a smaller etch depth increases coupling efficiency again

points to the fact that for TM polarization weaker gratings are required. Field plots of TM grating couplers vs TE grating couplers point in the same direction. For a TM grating most of the scattering occurs at the first few periods while the rest of the grating contributes substantially less.

5.5.1.2 Experimental results

Experimental results were obtained by positioning a tunable laser source over the grating at a distance of approximately $10\ \mu m$. At the output, a fiber connected with a photodetector is positioned over the output grating. In most cases, both fibers are tilted 10 degrees from vertical position in order to avoid second order reflections. The transition to single-mode photonic wires is omitted in these test structures. Therefore, waveguide losses can be neglected since the waveguides are $3 - 10\ \mu m$ wide. By assuming that input grating and output grating are identical, the coupling efficiency of a single grating coupler can be calculated from :

$$\eta = \sqrt{P_{out}/P_{in}} \tag{5.4}$$

Grating couplers were fabricated with a period of $1080\ nm$ and different fill factors. Comparison of measurement results and simulation results is shown in figure 5.13, 5.14, 5.15 and 5.16. Measurement results show poor performance compared to simulation results, both in terms of absolute efficiency as in terms of $3dB$ bandwidth.

In a second fabrication run grating couplers with a period of $1040\ nm$ were fabricated. The groove width is $520\ nm$ and the number of periods is 18. Figure 5.17 shows the simulated efficiency versus the measurement results. While the maximum efficiency is comparable with the previous run, the overall efficiency is substantially less noisy than the previous samples. However, the shape of the curves does not match simulation results. Additional samples to determine the physical reason for this discrepancy have been designed, however, the samples did not arrive on time to be characterized.

Measurement	Max Wavelength (nm)	Efficiency
1	1555	24.02
2	1558.5	27.47

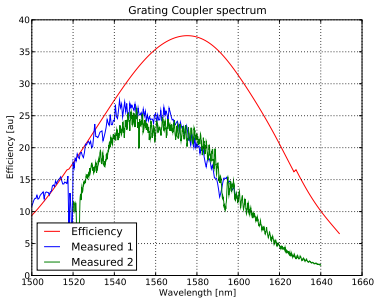


Figure 5.13: Efficiency η grating coupler, period $1080\ nm$*, groove width* $500\ nm$*, 14 periods*

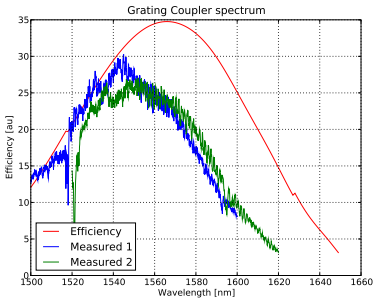


Figure 5.15: Efficiency η grating coupler, period $1080\ nm$*, groove width* $540\ nm$*, 16 periods*

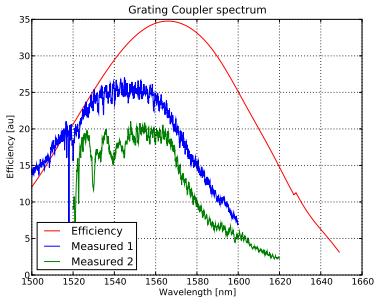


Figure 5.14: Efficiency η grating coupler , period $1080\ nm$*, groove width* $540\ nm$*, 14 periods*

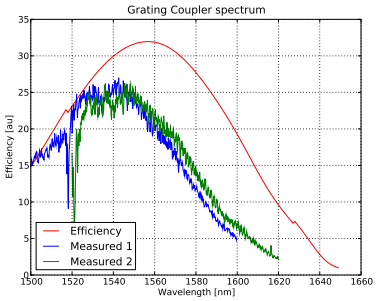


Figure 5.16: Efficiency η grating coupler, period $1080\ nm$*, groove width* $580\ nm$*, 14 periods*

5.5.2 End-fire Coupling

End fire coupling involves launching light into an optical waveguide through a tapered fiber and collecting the transmitted light at the opposite side of the sample with an objective.

Though conceptually very simple and broadband, in practice end-fire coupling has some considerable technical disadvantages over out-of plane coupling. Coupling efficiency from the lensed fiber to a thin SOI waveguide can be estimated to be of the order of $-10dB$, outcoupling losses from the sample facet to the detector adds approximately $-8dB$ loss. Moreover, coupling efficiency is very sensitive to sample alignment. In practice, even careful alignment cannot prevent coupling efficiency to vary from one waveguide to the next, or from one sample to another.

Facet quality determines coupling efficiency. To produce clean facets,

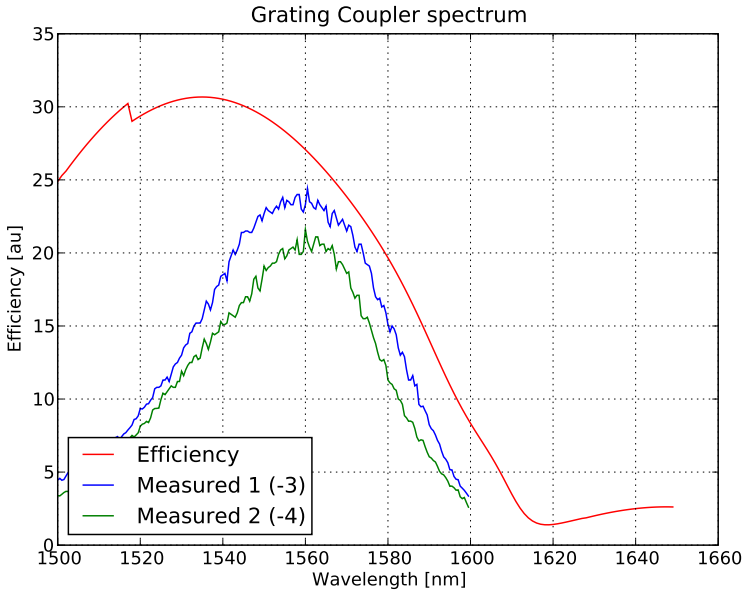


Figure 5.17: Efficiency η grating coupler, period $1040\ nm$, groove width $520\ nm$, 18 periods. The labels -1 and -3 indicate th column on the wafer and thus the exposure dose during deep-etch lithography.

extensive post-processing is required. Since SOI structures are supported by a $725\ \mu m$ thick substrate, and Silicon is very resilient to stress, cleaving a nice facet in an SOI wafer is difficult to say the least. Experimentally determined cleaving quality of thick SOI samples was deplorably low. Two solutions were considered and pursued, the easiest procedure is backlapping the Silicon substrate (using an LM15 pad polishing machine). This procedure is illustrated in figure 5.18.

Before backlapping, to protect the structures on the die, a layer of photoresist is deposited and cured. The die is then glued onto a glass plate with bee's wax heated to $130°C$. After cooling, the glass plate is mounted on the backlapping machine using a vacuum chuck. Depending on the size of the die and the number of dies attached to the glass plate, backlapping with a steel plate and a suspension of alumina powder (Al_2O_3) can take up anywhere in between 1 to 4 hours. The result is that the original $725\ \mu m$ substrate is thinned down to about $250\ \mu m$. Removal of the now thin sample from the plate can be done by reheating the bee's wax, dissolving the wax residue on the sample in TCE (C_2HCl_3), and consequently removing the

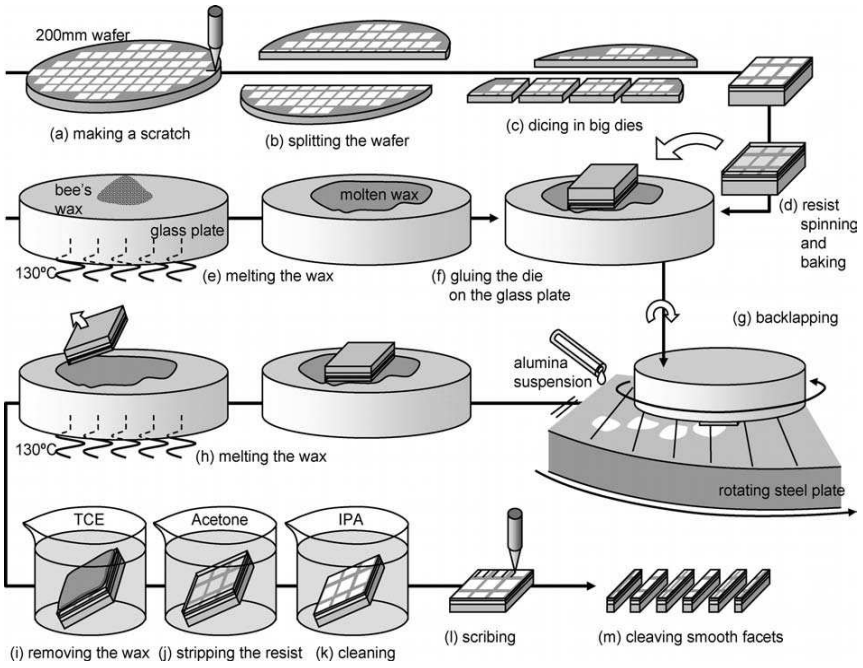


Figure 5.18: Postprocessing to make incoupling facets.The dies are coated with resist to protect the structures and glued upside-down on a glass plate with hot bees wax. Then, the $725\mu m$ thick substrate is backlapped down to about $250\mu m$. The thin sample is removed from the plate by melting the wax again, and the residual wax is dissolved in TCE. Then,the resist coating is stripped with acetone and the sample is cleaned with isopropanol (IPA) and DI water. With a diamond needle, scratches are made where the incoupling facets should be and the sample is cleaved by applying pressure on the bottom. Figure taken from [2]

TCE and the photoresist layer with acetone, isopropanol (IPA) and DI water. Cleaving is carried out by mechanically introducing a small crack at the edge of the sample, and subsequently applying pressure to the sample so that it cracks along a primary crystal plane. This technique is difficult to master, and is limited to use in a research laboratory, since the very nature of the technique renders it unsuitable for commercial application. The results can be very good, but tend to be variable in SOI because the introduction of buried oxide layer means the surface and substrate layers of silicon are not directly connected. Furthermore, when working with $a - Si$ top layers, there is no well defined crystallographic orientation present in this layer.

The most common method of preparing waveguide facets is probably polishing. The sample endface is polished by lapping with abrasive mate-

rials with sequentially decreasing grit sizes. Usually a prescribed recipe is followed which can result in an excellent surface finish, although rounding of the endface is a common problem [14, 15].

For this work samples were polished at the Technical University of Berlin, member of the photonics packaging platform ePIXpack [16].

5.6 Surface Chemistry

As discussed in previous chapters gold surfaces are commonly used in SPR as they are good substrates for the generation of surface plasmons and for their stability to oxidation. The secret of a successful affinity biosensor partially hides in the chemical interface layer between the transducer system and the biological receptor molecules, since a molecule at low concentration needs to be detected in the presence of a much larger concentration of non-specific molecules. This provides the device with a selectivity for specific biomolecules. The chemical modification of gold substrates is typically obtained by formation of self assembly monolayers (SAMs) from alkanethiols ($HS(CH_2)_nX$), which can carry useful functional groups [17].

Alkanethiols chemisorb on gold surfaces, typically from solution or vapor, by forming a gold sulphur bond, presumably gold(I)thiolate, which is fairly strong ($H = 28Kcal/mol$). The SAM has a highly ordered, almost crystalline structure, with chains densely packed and tilted 30° from the vertical to the surface, in order to minimize free volume and to maximize van der Waal's interactions between the oligomethylene chains. The chemical nature of the head group does not seem to affect much the structure of the SAM, if the size of the head group is not larger than that of a methyl group. These surfaces are stable and highly reproducible. Moreover, the thickness of only few nanometers allows obtaining a better signal compared to other matrices commonly used (carboxymethylated dextran [18]), as the binding events take place closer to the metal surface where the evanescent wave is more intense.

Mixed SAMS consist of two different alkanethiols, deposited simultaneously on the gold surface, creating a self-assembled monolayer structure incorporating various functional groups. While the first thiol molecule caries a functional group that is capable of covalently binding the biological receptors (e.g. a carboxylic endgroup), the second molecule incorporates a functional group (e.g., hydroxyl, EO) that shows excellent properties to reduce the nonspecific absorption of undesired entities [19, 20, 21, 22, 23, 24]. Incorporating both alkenethiols in a mixed SAM enables full control over the number of functional groups, the steric hindrance, and the nonspecific adsorption [25].

For example, antifouling properties[4] of PEO (polyethylene oxide) are very well reported, due to the tightly bound layer of water on the EO chains that prevent protein adsorption on the surface [26]. The hindrance property of protein rejection is intrinsically related to the self-repulsion of PEO in water. The protein resistance has been shown to increase with longer chains and higher surface density of the SAM [27]. Therefore, mixed SAMs containing carboxylic functional groups and EO oligomers allow the attachment of receptor molecules and reduce non specific interactions, enhancing the performances of the system.

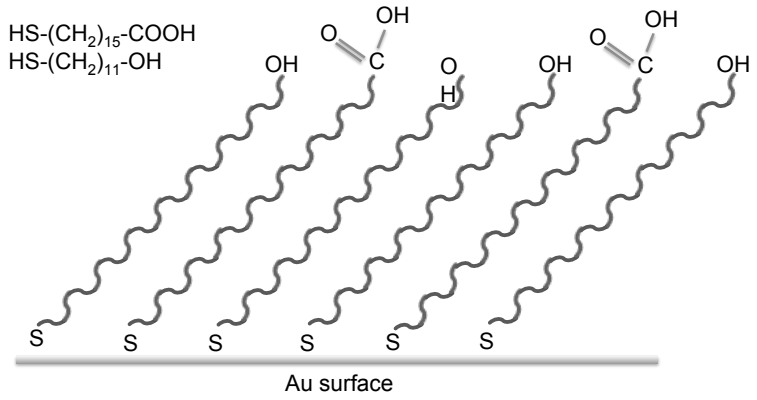


Figure 5.19: Preparation of mixed SAMs on Au surfaces, $16-MHA/11-MuOH$ mixed self-assembled monolayer.

Often used self assembled monolayers are the combination of 16-mercapto-hexadecanoic acid ($16-MHA$) carrying the functional carboxyl groups, and 11-mercapto-1-undecanol ($11-MuOH$) or ethylene oxide containing alkyl thiols to provide the antifouling functionality. The non-specific adsorption of some proteins on mixed SAM from $16-MHA/11-MuOH$ and from $16-MHA/SH-C_6EO_3-OH$ was reported to be comparable [28]. The former process is depicted in figure 5.19

Another well established SAM is a combination of alkanethiols that present oligo(ethyleneglycol) ($EG_{n=3,6}$) groups, these molecules resist the adsorption of proteins, and alkanethiols that present $(EG)_6OR$ groups, where R is the ligand of interest. This engineered surface has protein resistance properties and at the same time the protruding functional groups allow easier interaction with the analyte [21].

[4] Antifouling refers to the property of a surface to prevent itself from being covered by any biomolecule.

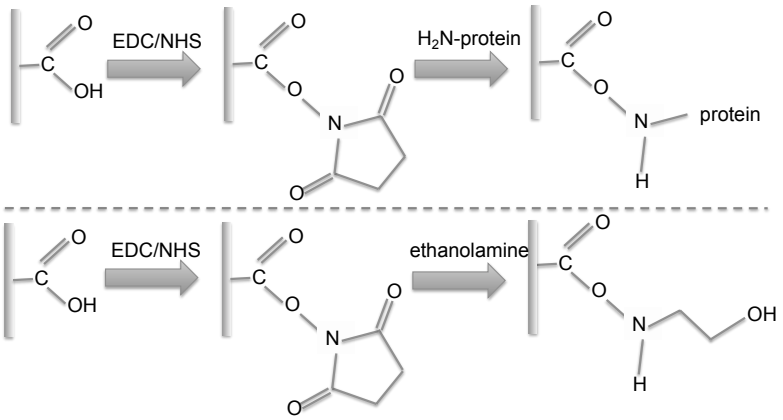


Figure 5.20: Covalent coupling of amino groups to a carboxylic-terminated surface and the blocking of unreacted N-hydroxysuccinimide esters via ethanolamine

After carboxyl-group activation and covalent molecular receptor coupling, the remaining active groups should be blocked to avoid covalent bonding of the analyte to the surface. This is usually achieved with ethanolamine (as shown in Fig. 5.20). Another approach to block remaining carboxyl groups is the use of PEO containing molecules as blocking agents (e.g. triethylene glycol monoamine, $H_2N-(CH_2-CH_2-O)_3-H$) after the activation and the coupling of the receptor, and a reduction of the non specific adsorption on the surface was also observed [29].

5.6.1 This Work

Since the required chemical expertise was not available in-house, surface modification was done in collaboration with the Polymer Chemistry and Biomaterials Group (PBM-UG) at Ghent University. Self-assembly monolayers were prepared by immersing for 24 hours the cleaned gold substrates in $1mM$ ethanolic solutions of 16-Mercaptohexadecanoic acid ($16-MHA$) and 11-Mercaptoundecanol ($11-MuOH$) in volume ratio $90:10(11-MuOH:16-MHA)$ (see Fig. 5.19). For more information, we refer the reader to [28]. Samples were placed in the orbital shaker for $24h$ at $150rpm$, after which they were rinsed with ethanol and dry blown with argon.

Biotin was coupled to the carboxylic functionalities of the surface after activation with EDC/NHS [5] [30] in water ($0.4M/0.1M$). Ethanolamine

[5] EDC=1-ethyl-3-(3-dimethylamino-propyl)carbodiimide / NHS=n-hydroxysucchinimide

hydrochloride $1M$ pH 8.5 was added to deactivate the unreacted carboxylic groups and to remove electrostatically bound molecules. A $2ml$ of a $1mg/ml$ solution of 5-(biotinamido)pentylamine in PBS (pH 7.4) was used to couple biotin to the surface (biotine was dissolved in 0.25ml dry DMF and 1.75 ml PBS is added to achieve 2ml of a 1 mg/ml solution). Immobilization and recognition experiments using the SPR apparatus, non-specific adsorption and competitive interaction experiments have been performed on these sample, and have been thoroughly described in [28]

5.7 Final devices

To summarize, the final devices that were experimentally characterized consisted of a series of waveguides, available in three different widths ($450nm$, $3\mu m$, 10 μm), having shallowly etched sensing regions (70 nm etch) with lengths ranging from 4 to 10 μm, in $1\mu m$ steps. (Part of) the mask design for these devices can be seen in figure 5.21, a top view of the fabricated device can be seen in figure 5.22. The first step in the fabrication process shown in figure 5.1 was performed in the IMEC cleanrooms, the following steps (except the polishing step) were carried out in our cleanroom facilities. In figure 5.23, the FIB cross-section of the final fabricated device is shown.

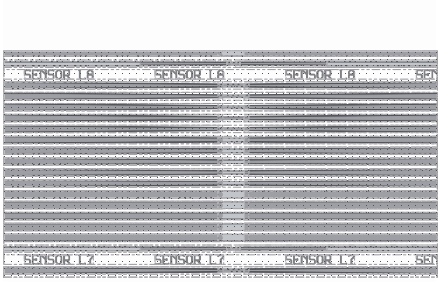


Figure 5.21: Detail of the final DUV mask design

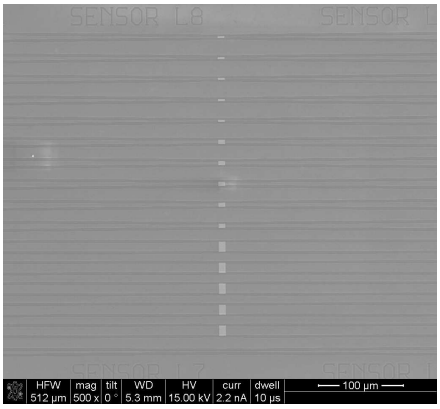


Figure 5.22: SEM top view of the final fabricated samples after Au deposition and liftoff

Figure 5.23: FIB cross-section of a fabricated sample (the sample depicted here is the sample that was measured in measurement run 1, see chapter 7 for experimental results).

References

[1] W. Kern. *Handbook of semiconductor cleaning technology*. Noyes Publishing, Park Ridge, NJ, 1993.

[2] W. Bogaerts. *Nanophotonic Waveguides and Photonic Crystals in Silicon-on-Insulator*. PhD thesis, Ghent University, 2004.

[3] S. Selvaraja, E. Sleeckx, M. Schaekers, W. Bogaerts, D. Van Thourhout, P. Dumon, and R. Baets. *Low-loss amorphous silicon-on-insulator technology for photonic integrated circuitry*. Optics Communications, 182(9):1767–1770, 2009.

[4] R. J. Shul and S. J. Peraton. *Handbook of advanced plasma processing techniques*. Springer-Verlag, Berlin Heidelberg, 2000.

[5] S. M. Sze. *Semiconductor devices: physics and technology*. John Wiley and Sons Inc, New York, 1986.

[6] D. Vrtacnik, D. Resnik, U. Aljancic, M. Mozek, and S. Amon. *Silicon dry etching profile control by RIE at room temperature for MEMS applications*. In Conference on Device and Process Technologies for Microelectronics, MEMS and Photonics IV, volume 6037, Brisbane, 2005. SPIE.

[7] G. Roelkens. *Heterogeneous III-V/silicon photonics: bonding technology and integrated devices*. PhD thesis, Ghent University, 2007.

[8] P. E. Garrou, R. H. Heistand, M. G. Dibbs, T. A. Mainal, C. E. Mohler, T. M. Stokich, P. H. Townsend, G. M. Adema, M. J. Berry, and Turlik. I. *Rapid thermal curing of BCB dielectric*. IEEE Transactions on Components Hybrids and Manufacturing Technology, 16(1):46–52, 1993.

[9] Y. H. So, P. E. Garrou, J. H. Im, and K. Ohba. *Benzocyclobutene-based polymers for microelectronic applications*. Polymers for Microelectronics and Nanoelectronics, 874:279–293, 2004.

[10] DOW. *Processing procedures for CYCLOTENE 4000 series photo-BCB resins (puddle develop process)*, 2009.

[11] T. Tamir and S.T. Peng. *Analysis and Design of Grating Couplers*. Applied Physics, 14(3):235–254, 1977.

[12] T. Tamir. *Topics in Applied Physics: Integrated Optics*. Springer-Verlag, New York Heidelberg Berlin, 1979.

[13] D. Taillaert. *Grating couplers as interface between optical fibers and nanophotonics waveguides*. PhD thesis, Ghent University, 2004.

[14] G. T. Reed and A. P. Knights. *Silicon Photonics: an Introduction*. John Wiley and Sons, 2004.

[15] L. Zimmermann, M. Schnarrenberger, T. Mitze, J. Bruns, and K. Petermann. *Study of reactive ion etched (RIE) facets of silicon-on-insulator (SOI) rip waveguides*. In S. RivaSanseverino and M. Artiglia, editors, 4th Workshop on Fibres and Optical Passive Components, pages 260–264, Mondello, ITALY, 2005. IEEE.

[16] `http://www.epixpack.eu/`.

[17] J. Homola, H. Vaisocherova, J. Dostalek, and M. Piliarik. *Multi-analyte surface plasmon resonance biosensing*. Methods, 37(1):26–36, 2005.

[18] B. Johnson, S. Löfas, G. Lindquist, R.S. M. Hilgren, and H. Hansson. *Comparison of methods for immobilisation to carboxymethyl dextran surfaces by analysis of the specific activity of monoclonal antibodies*. Journal of Molecular Recognition, 8(1-2):125–131, 2004.

[19] P. Kingshott and H. J. Griesser. *Surfaces that resist bioadhesion*. Curr. Opin. Solid State Mater. Sci., 4:403–412, 1999.

[20] K. L. Prime and G. M. Whitesides. *Self assembled organic monolayer model systems for studying adsorption of proteins at surfaces*. Science, 252:1164–1167, 1991.

[21] E. Ostuni, L. Yan, and G. M. Whitesides. *The interaction of proteins and cells with self-assembled monolayers of alkanethiolates on gold and silver*. Colloids and Surface B: Biointerfaces, 15(1):3–30, 1999.

[22] R. L. C. Wang, H. J. Kreuzer, and M. Grunze. *Molecular conformation and solvation of oligo(ethylene glycol)-terminated self assembled monolayers and the resistance to protein adsorption*. J. Phys. Chem. B, 101:9767–9773, 1997.

[23] M. Lindblad, M. Lestelius, A. Johansson, P. Tengvall, and P. Thomson. *Cell and soft tissue interactions with methyl- and hydroxyl-terminated alkane thiols on gold surfaces*. Biomaterials, 18(15):1059–1068, 1997.

[24] M. A. Cooper. *Optical biosensors in drug discovery*. Nature Reviews Drug Discovery, 1(7):515–528, 2002.

[25] K. Jans, K. Bonroy, R. De Palma, G. Reekmans, H. Jans, W. Laureyn, M. Smet, G. Borghs, and G. Maes. *Stability of mixed PEO - Thiol SAMSs for biosensing applications*. Langmuir, 24:3949–3954, 2008.

[26] M. Silván, A. Valsesia, D. Gilliand, G. Ceccone, and F. Rossi. *Ion-beam treatment of PEO: towards a physically stabilized anti-fouling film*. Surface and Interface Analysis, 36(8):733–736, 2004.

[27] L. Li, S. Chen, J. Zheng, B. D. Ratner, and S. Jiang. *Protein adsorption on oligo(ethylene glycol)-terminated alkanethiolate self-assembled monolayers: The molecular basis for nonfouling behavior*. J. Phys. Chem. B, 109(7):2934–2941, 2005.

[28] Jordi Girones. *GOA Final report*. PhD thesis, Polymer Chemistry and Biomaterials Group, Ghent University, 2008.

[29] F. Frederix, K. Bonroy, G. Reekmans, W. Laureyn, A. Campitelli, M. A. Abramov, W. Dehaen, and G. Maes. *Reduced nonspecific adsorption on covalently immobilized protein surfaces using poly(ethylene oxide) containing blocking agents*. J. Biochem. Biophys. Methods, 58:67–74, 2004.

[30] T. Cass and F.S. Ligler. *Immobilized biomolecules in analysis: a practical approach*. New York: Oxford Univ. Press, 1998.

6 Measurement Software and Data Analysis

SPECIFIC software has been written to facilitate the measurement process for bulk sensing and biosensing measurements. This chapter describes all of the necessary software components and basically falls apart into two distinct sections. The first section deals with the software that was written specifically to acquire data during measurements. The second part is concerned with the optimization of the data-analysis method used in the bulk sensing and biosensing experiments. The central theme in this section is to determine which data analysis method will give the best estimate of the refractive index of the analyte, or, during surface sensing experiments, provide the best estimate of the thickness of the adsorbed layer.

6.1 Data Acquisition

Programming measurement instruments can be a real pain. There are many different protocols, sent over many different interfaces and bus systems (GPIB, RS232, USB). For every programming language you want to use, you have to find libraries that support both your device and its bus system.

In order to ease this unfortunate situation, the VISA [1] specification was defined in the middle of the 90ies. Today VISA is implemented on all

[1] Virtual Instrument Software Architecture, supported by National Instruments.

significant operating systems. A couple of vendors offer VISA libraries, partly with free download. These libraries work together with arbitrary peripheral devices, although they may be limited to certain interface devices, such as the vendor's GPIB card.

The VISA specification has explicit bindings to Visual Basic, C, and G (LabVIEW's graphical language). However, you can use VISA with any language capable of calling functions in a DLL. Python is such a language. The benefits of using Python to write dedicated measurement programs are:

- Python is an easy-to-learn scripting language with very short development cycles
- The online Python community is very active and generally has answers for most programming issues.
- It represents a high abstraction level, which perfectly blends with the abstraction level of measurement programs.
- It has a very rich set of native libraries, including numerical and plotting modules for data analysis and visualization.
- For complex measurement programs, a text-based programming language is preferable over a graphical programming language.

Since all data analysis methods described in the following section were implemented in Python, for this work, all data acquisition was done using the pyVISA library. Specific, and specialized measurement procedures were written to enable bulk automated bulk refractive index sensing and biosensing measurements.

6.2 Data Analysis

6.2.1 Introduction

Data processing is the process of transforming raw data recorded from a sensor into quantities of interest. This is a final step in a measurement and its quality determines the result of the whole measurement process. Except really simple experiments raw data typically has the form of a set of pairs (x_i, y_i); $i = 1, n$, where x_i corresponds to a position (spatial, angular, wavelength) and y_i is a value measured at this point (power, current, voltage, charge). The raw y_i data may be preprocessed prior to analysis. This includes averaging over several spectra to reduce shot noise, zero offset (subtraction of a dark current of photodiode or CCD) and normalization

(division by a reference spectrum). Obtained spectra may be viewed as a random vector with distribution depending on parameters of the system studied, including our quantities of interest. Ideally, data processing is an estimation of quantities of interest with following features:

- No bias: expected value is equal to the estimated quantity, in other words, results are not systematically offset from their real values.
- Minimal noise: variance of results is minimal. The method with minimal variance is chosen from all unbiased methods.

In order to optimize the detection limit the correct choice of data-analysis algorithm is quite important. A number of data analysis methods have been reported in literature. Most methods rely on a number of weighed data points and trace the shift of the transmission minima, which is then returned as the output of the method. The described methods range from simple data processing methods such as polynomial curve fitting, Lorentzian curve fitting and centroid methods [1], to more extensive locally weighted parametric regression methods [2], optimal linear methods [3] and asymmetric curve fitting methods [4].

Especially for sensing applications, since the detection limit depends on measurement noise, careful analysis of the acquired spectrum is of the utmost importance. In this work we will describe and compare polynomial curve fitting algorithms, the Lorentzian fitting algorithm and two distinct centroid methods.

6.2.2 Polynomial Fitting Methods

One particular choice of dip finding algorithm is to fit a polynomial to the region of the transmission spectrum surrounding the minimum. The location of the minimum of the polynomial is then determined analytically (or numerically using a bounded minimization function such as Brent's rule). The method is illustrated in figure 6.1(a). The polynomial can be determined via a least-squares fit, which has the benefit of being linear and the minimum of the polynomial can be easily localized iteratively. A proper threshold has to be set and only part of the spectrum under the curve will be used. The disadvantage is that the variance of the determined position of the minimum is higher than in other methods.

6.2.3 Lorentzian Fitting Methods

A second method uses a Lorentzian curve (or a product of the Lorentzian curve(s) and a simple envelope function) as an appropriate model function

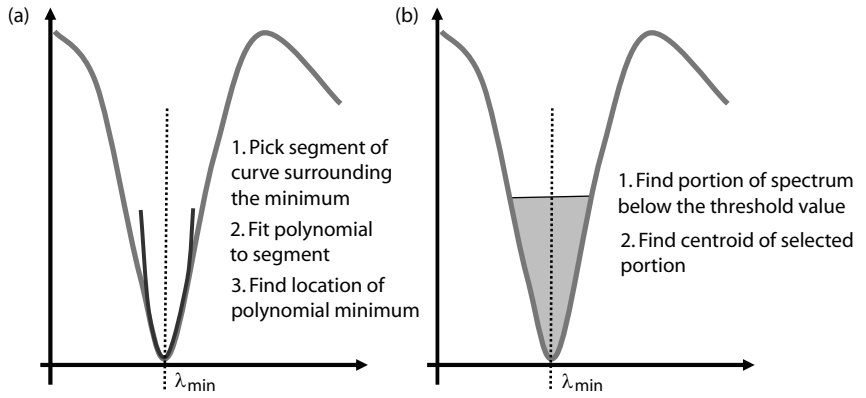


Figure 6.1: Illustration of the algorithms behind polynomial (or) Lorentzian curve fitting and (weighted) centroid minima finding.

for the spectrum.

$$T = 10log\left(\frac{r(FWHM/2)^2 + (\lambda - \lambda_{min})^2}{(FWHM/2)^2 + (\lambda - \lambda_{min})^2}\right) \tag{6.1}$$

where, r is a fitting parameter, $FWHM$ is the Full-Width at Half-Maximum of the dip and λ_{res} is the spectral minimum. The fit is non-linear but with a reasonable initial estimate converges in a few iterations. There is no need for an artificial threshold (although one can be used), it is usually more robust against overall light fluctuations and it has better signal-to-noise ratio than the least-squares polynomial methods.

6.2.4 Centroid Method

This method is based on the notion that if the attenuation feature in the spectrum is symmetric, the minimum of the spectrum will be identical to the centroid of the shaded region in Fig. 6.1. Therefore, the spectrum may be characterized by setting a threshold value $T_{threshold}$ and finding the centroid of the portion of the spectrum which lies below $T_{threshold}$. The spectral minimum is given by

$$\lambda_{min} = \frac{\sum_i^{T<T_{threshold}}(\lambda_i(T_i - T_{threshold}))}{\sum_i^{T<T_{threshold}}(T_i - T_{threshold})} \tag{6.2}$$

Although the method is very simple, it is nevertheless very useful. Moreover, the algorithm leads to a better linearity and increases the dynamic range, because only the important part of the transmission curve is used, i.e. the

dip. The threshold is an important parameter which affects the noise level of the centroid output. A higher threshold value will generally reduce the centroid noise by including more data points in the calculation of the resonance position. Another noise-reduction principle that can be applied is linear interpolation between data points in order to create a piecewise-linear spectrum from the discrete normalized spectrum. Doing so more accurately determines the exact spectral position where $T_i < T_{threshold}$, and hence leads to a better estimate of the spectral minimum by including more relevant data point. This also enables one to measure with a limited resolution, but obtain well defined centroid values with small errors. If the dip is asymmetrical the centroid will not be at the position of the minimum and hence the method will no longer provide unbiased values. However, for biosensing purposes, this deviation is of no importance since only relative movements are of interest. If the shape of the dip is constant, the difference between the centroid wavelength and the minimum wavelength is constant, so the offset will cancel upon subtracting the new centroid position from the old centroid position. If the shape of the dip changes upon movement, the difference between actual dip and centroid value will change ever so slightly, however, this can easily be solved by calibrating the device with well-known samples. The biggest problem of the centroid algorithm is that

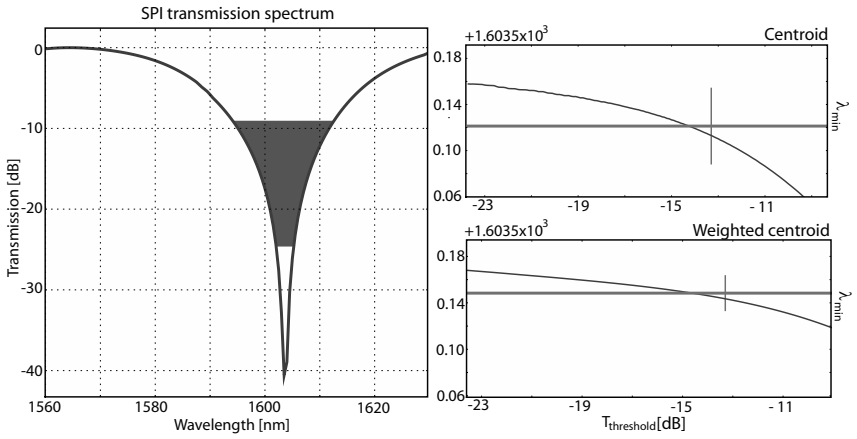


Figure 6.2: Comparison linearity error unweighted and weighted centroid algorithm. Left: Transmission spectrum typical SPI sensor, H_2O cladding, shaded regions indicate the consecutive threshold values, step size is $0.1\ dB$. Right: linearity error unweighted centroid algorithms as a function of $T_{threshold}$ (top) and linearity error of the weighted centroid algorithm as a function of $T_{threshold}$ (bottom).

it initiates a linearity error due to the (arbitrary) cut-off limit. The origin is

that transmission near cutoff will change their values in a nonlinear fashion. Reducing this linearity error can be done by multiplying the value at each wavelength by a weight factor. In this case values closer to cutoff are given a smaller contribution to the response. By weighing each transmission value with it's distance from $T_{threshold}$, a very simple but utterly effective weighing procedure can be implemented. This method is denoted as the weighted centroid algorithm. The response is then calculated as

$$\lambda_{res} = \frac{\sum_i^{T<T_{threshold}}(w_i\lambda_i(T_i - T_{threshold}))}{\sum_i^{T<T_{threshold}}(T_i - T_{threshold})} \quad \text{with} \quad w_i = T_{threshold} - T_i \tag{6.3}$$

Although the introduced linearity errors cannot be avoided, using the weighted centroid algorithm reduces this value by a factor of 5, this is shown in figure 6.2. In this graph the value of the threshold used $T_{threshold}$ varies from the minimal transmission T_{min} to the dynamically calculated maximum threshold value $T_{max,dyn}$ which is defined by

$$\frac{\int^{T>T_{max,dyn}}(T - T_{max,dyn})d\lambda}{\int^{T<T_{max,dyn}}(T_{max,dyn} - T)d\lambda} = c \tag{6.4}$$

So the ratio of the area of the transmission spectrum below and above $T_{max,dyn}$ is the same for all spectra (being the constant c). This removes the susceptibility to correlated light-source noise, such as an overall change in the intensity of the light source. Such changes cause the spectrum to be raised or lowered relative to the threshold, including more or fewer data points in the centroid calculation,which can easily be countered by dynamically adjusting the threshold value of the baseline [5]. The difference between two consecutive threshold values was $0.1\ dB$. Each different threshold value gives rise to a different value for λ_{min}.

Other, more advanced, centroid methods have been developed and are currently being used, modifications typically try to minimize the effect of experimental noise on the calculated minimal wavelength. The most prominent technique here is the interpolated tracking centroid algorithm [1]. This modified algorithm makes the assumption that the width of the resonant dip does not change over the course of the experiment and consists of two interrelated portions: a tracking algorithm and an interpolation routine.

6.2.5 Limitations

Although the former data-analysis techniques have build an impressive track record when it comes to surface plasmon resonance sensing, and can easily

be used to analyze surface plasmon interferometer spectra, they do display some shortcomings:

- they provide no real information about what sensor resolution can be expected from a particular sensor and data analysis algorithm.
- since each method only uses a portion of the spectrum and reduces that portion to one single parameter, useful sensor data is thrown away.

Nevertheless, above mentioned data analysis methods are widely used and for (bio)sensing purposes are sufficiently accurate, as these shortcomings are only minor issues [3].

6.3 Example

Figure 6.3 shows the typical spectrum of a surface plasmon interferometer (etch depth 70 nm, Au-thickness 37 nm, Au length 8 μm). Using this spectrum as input we have evaluated 6 different algorithms: fitted second order polynomial, fitted third order polynomial, fitted fourth order polynomial, Lorentzian fit, centroid method and the weighted centroid method.

We have applied a two-point moving average filter to the spectral data, and normalized each spectrum to 0 dB. Figure 6.4 shows the evolution of the dip position for six different dip tracking algorithms. All dip finding algorithms use a similar linear interpolated data set. Of these calculated dip positions the mean and the standard deviation (representing sensitivity to $T_{threshold}$) were calculated, these values are presented in table 6.1 and in figure 6.5.

$NaCl$ conc. $[m\%]$		0	2	4	6	9	10
M1	$\lambda_{min}[nm]$	1603.637	1600.717	1599.366	1598.035	1596.076	1595.462
	$\delta\lambda_{min}[nm]$	0.014	0.005	0.015	0.023	0.031	0.032
M2	$\lambda_{min}[nm]$	1603.640	1600.720	1599.369	1598.036	1596.073	1595.455
	$\delta\lambda_{min}[nm]$	0.011	0.006	0.016	0.022	0.028	0.027
M3	$\lambda_{min}[nm]$	1603.642	1600.717	1599.358	1598.020	1596.049	1595.432
	$\delta\lambda_{min}[nm]$	0.009	0.004	0.007	0.011	0.014	0.015
M4	$\lambda_{min}[nm]$	1603.665	1600.713	1599.352	1598.015	1596.049	1595.435
	$\delta\lambda_{min}[nm]$	0.002	0.001	0.004	0.007	0.010	0.011
M5	$\lambda_{min}[nm]$	1603.648	1600.715	1599.361	1598.028	1596.066	1595.452
	$\delta\lambda_{min}[nm]$	0.009	0.003	0.010	0.016	0.021	0.022
M6	$\lambda_{min}[nm]$	1603.661	1600.712	1599.352	1598.015	1596.048	1595.434
	$\delta\lambda_{min}[nm]$	0.006	0.002	0.005	0.008	0.011	0.012

Table 6.1: Numerical comparison of the different minima finding algorithms.

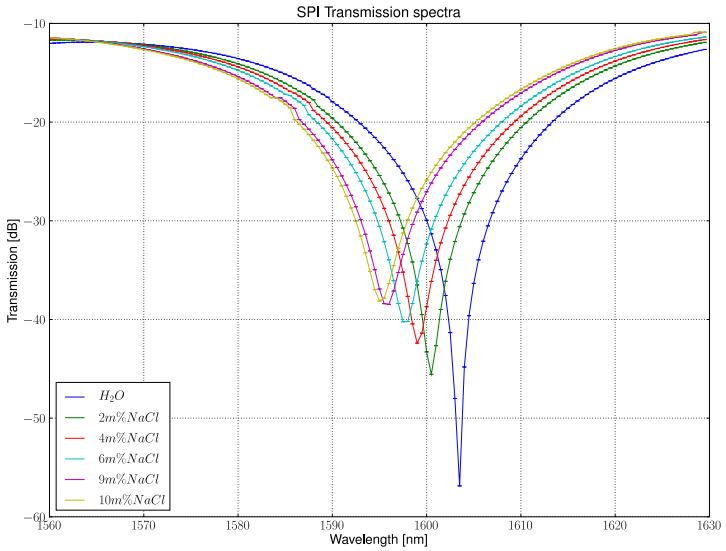


Figure 6.3: Simulated transmission spectrum for a surface plasmon interferometer (etch depth 70 nm, Au-thickness 37 nm, Au length 8 μm)

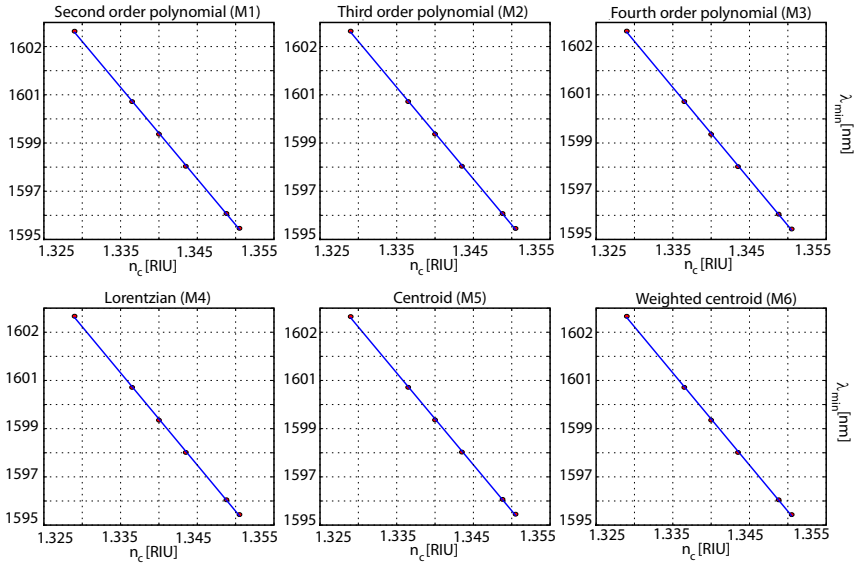


Figure 6.4: Comparison of the different dip finding algorithms

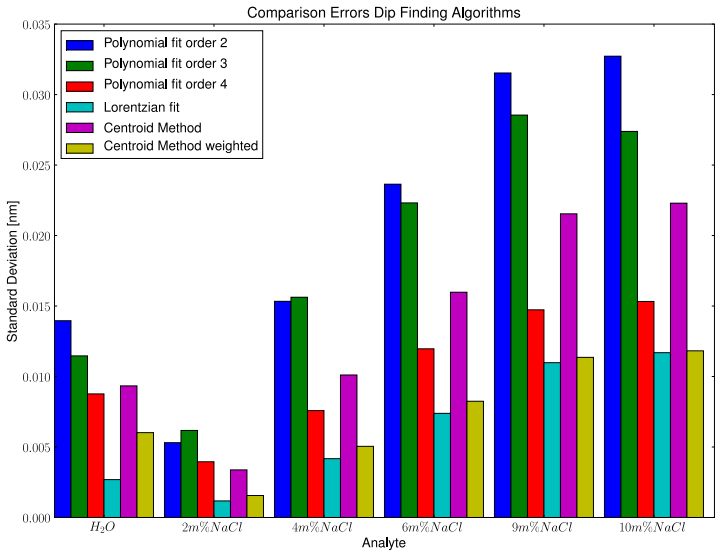


Figure 6.5: Comparison of the different minima finding algorithms: standard deviation errors

From this table we can deduce that among the curve fitting algorithms the Lorentz fitting algorithm clearly is superior to the fitted second order, third order and fourth order algorithms, in the sense that λ_{min} displays the smallest dependence on the value of $T_{threshold}$ With regards to the centroid algorithms it is quite obvious that the weighted centroid algorithm outperforms the original centroid algorithm. The Lorentzian fit algorithm and the weighted centroid algorithm have a comparable performance.

From the calculated resonance wavelengths one can easily deduce the sensitivity $S_{bulk}[nm/RIU]$ by fitting a straight line to these values and calculating the slope. The variation of dip position versus refractive index of the analyte seems to be linear to a good approximation. Sensitivity values are given in table 6.2.

Using formula 4.34 we can determine the theoretical sensitivity for this device as being $391nm/RIU$, which, given the approximative nature of formula 4.34, can be considered adequately accurate. The value obtained here is more accurate, since this value is obtained by full-wave simulation and subsequent analysis of the spectra. Table 6.2 also gives the fitting error of the linear approximations, which is indicative of the linearity of the sensor response. One can see that all six fitting methods show a linear dependency

Method	$S_{bulk}[nm/RIU]$	$\delta S_{bulk}[nm/RIU]$
Fitted second order polynomial	380.168	0.037
Fitted third order polynomial	380.550	0.036
Fitted fourth order polynomial	381.871	0.035
Lorentzian fit	382.597	0.043
Centroid method	381.072	0.040
Weighted centroid method	382.536	0.042

Table 6.2: Numerical comparison of the calculated sensitivity value for the different minima finding algorithms.

Method	$DL[RIU]$	$\delta DL[RIU]$
Fitted second order polynomial	5.370×10^{-5}	2.835×10^{-5}
Fitted third order polynomial	4.888×10^{-5}	2.364×10^{-5}
Fitted fourth order polynomial	2.732×10^{-5}	1.162×10^{-5}
Lorentzian fit	1.670×10^{-5}	1.152×10^{-5}
Centroid method	3.622×10^{-5}	1.964×10^{-5}
Weighted centroid method	1.931×10^{-5}	1.035×10^{-5}

Table 6.3: Numerical comparison of the calculated detection limit for the different minima finding algorithms.

between the refractive index of the analyte and the position of the dip.

Since the refractive index uncertainty σ_{RIU} is related to the uncertainty of the minimal dip position $\delta\lambda_{min}$ by [6]

$$\sigma_{RIU} = \frac{\delta\lambda_{min}}{S_{bulk}}. \tag{6.5}$$

one can calculate the detection limits (DL) for this device (since the detection limit equals the uncertainty σ_{RIU} of the determined refractive index), these calculated values can be found in table 6.3

As to be expected, since the Lorentzian curve fitting method and the weighted centroid methods give the most reliable estimates for the position of the minima, these two dip finding algorithms also have the smallest detection limits.

6.4 Simulated Noise

With added amplitude noise, the located spectral minimum is not likely to be at the exact position where it should be theoretically. The amplitude noise processes thus result in a random spectral deviation of the measured spectral position of the minimum.

It is important to analyze the process of this deviation and its dependence on the magnitude of the noise. Although the noise processes under consideration are well characterized by standard statistical distributions, the extremum operation prevents a direct analytical solution for the statistical distribution of the resulting spectral variation of the minimum. However, the system can be numerically studied by using a Monte Carlo simulation. In the simulation, transmission spectra were calculated using the full wave

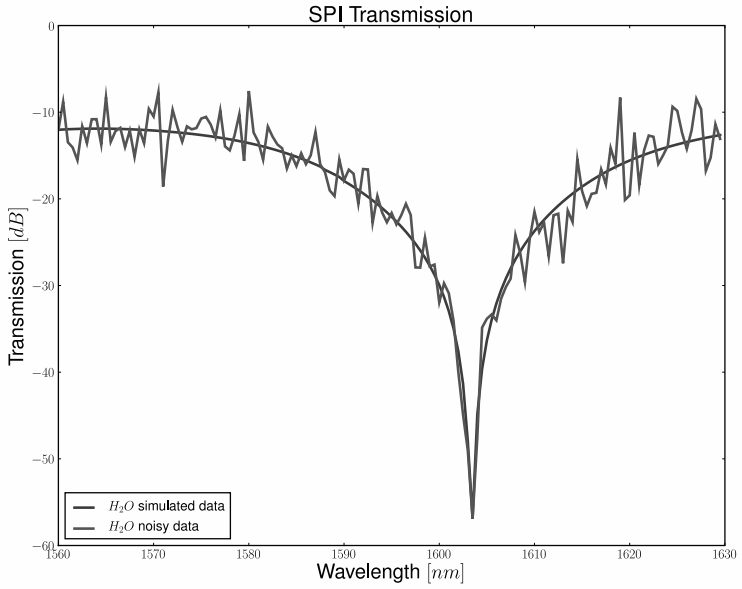


Figure 6.6: Illustration of the simulated spectrum with and without added noise, the magnitude of the standard deviation of the Gaussian white noise distribution equals $5\ dB$.

electromagnetic simulation techniques discussed earlier. These calculated transmission spectra are then discretized into points with an even spacing of $\delta\lambda$, corresponding to the minimum step size when the tunable laser is performing a wavelength scan. Randomly generated white Gaussian-type intensity noise is superimposed onto the spectral data points (although in

Method	$\lambda_{min}\ [nm]$	$\delta\lambda_{min}\ [pm]$	$\delta\lambda_{offset}\ [pm]$
Weighted centroid			
- $T_{threshold} = T_{ave}$	1603.647	8.923	30.282
- $T_{threshold} = T_{dyn,max}$	1603.613	12.7364	74.793
Lorentzian			
- $T_{threshold} = T_{ave}$	1603.695	5.348	4.305
- $T_{threshold} = T_{dyn,maxl}$	1603.693	5.075	4.048
$\lambda_{noiseless}$: 1603.694 nm			

Table 6.4: Comparison of the robustness to noise of the weighted centroid and the Lorentzian fit algorithm.

reality shot-noise is produced by a Poisson process), this can be seen in figure 6.6. The magnitude of the noise is expressed in dB, this is the standard deviation of the resulting noise distribution. Finally, a Lorentzian peak fitting and a weighted centroid peak fitting is applied to the now noisy data points in order to extract the spectral minimum λ_{min} from the data in 1000 Monte Carlo simulations. The minimum resolvable wavelength shift $\delta\lambda_{min}$ is then statistically defined as the standard deviation of the calculated minima of each iteration. $\delta\lambda_{offset} = |\lambda_{min} - \lambda_{min,noiseless}|$ is the deviation of the calculated spectral minimum from $\lambda_{noiseless}$. As we have seen in the previous section, the calculated position of the dip also is dependent on the exact value of the threshold. To study this effect, both the average value of all calculated minima, and the final value at $T_{dyn,max}$ will be used in the calculation of the minimal resolvable wavelength shift. The former will be indicated with $T_{threshold} = T_{ave}$, the latter with $T_{threshold} = T_{dyn,max}$. The probability density function (PDF) for the Monte-Carlo simulation of $T_{threshold} = T_{ave}$ and $T_{threshold} = T_{dyn,max}$, for a noise distribution with variance $0.25\ dB$ is plotted in Fig. 6.7. It's straightforward to see that the resulting variance of the calculated minimal wavelength λ_{min} is noticeably smaller when averaging over different values for the threshold. Exact numerical values are given in table 6.4, both for the weighted centroid algorithm as for the Lorentzian fitting algorithm. The standard deviation errors of the calculated minimal wavelengths of the Lorentzian fitted data are smaller than both weighted centroid methods (although the difference with is quite small). Furthermore, adding noise seems to bias the weighted centroid algorithms, therefore $\delta\lambda_{offset}$ is substantially larger than for the Lorentzian methods. As was mentioned earlier, this needn't be a huge problem, since we are only interested in relative wavelength shift, rather than absolute values.

Figure 6.9 and 6.10 plot the minimum resolvable wavelength $\delta\lambda_{min}$

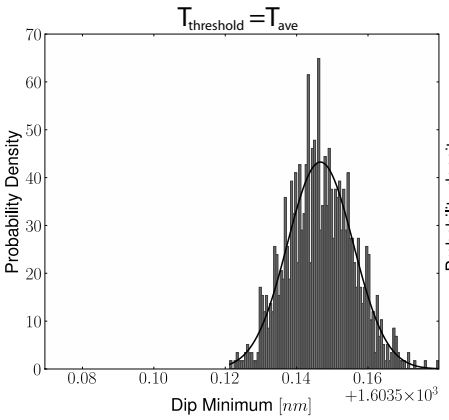


Figure 6.7: Histogram and fitted Gaussian of the calculated spectral position of λ_{min}, 1000 Monte-Carlo iterations, λ_{ave} shown.

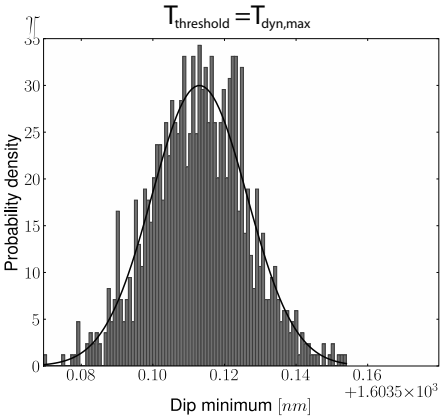


Figure 6.8: Histogram and fitted Gaussian of the calculated spectral position of λ_{min}, 1000 Monte-Carlo iterations, λ_{final} shown.

and the wavelength offset $\delta\lambda_{offset}$ as a function of the standard deviation of the noise distribution. With increasing standard deviation of the noise distribution, the standard deviation of the calculated value of λ_{min} becomes larger for both minima finding algorithms. Averaging λ_{min} over several threshold values clearly improves performance for the weighted centroid algorithm, for the Lorentzian fitting algorithm there is little difference between the two curves. $\delta\lambda_{min}$ for both Lorentzian fitting methods are smaller than the values for both weighted centroid methods. Only for very large noise values is the performance of the weighted centroid algorithm with $T_{threshold} = T_{ave}$ comparable to the Lorentzian fitting algorithms. $\delta\lambda_{offset}$ increases monotonically for all four minima finding algorithms, again, the better performing algorithms are both Lorentzian fitting methods.

From the calculated values of λ_{res}, theoretical detection limits in the presence of noise can be derived, these values are shown in figure 6.11. The Lorentzian λ_{ave} fitting algorithms is superior for all values of δ_{noise}. This will be the motivation for choosing this algorithm in the actual experiments.

6.5 Conclusion

Data-analysis methods are very important for biosensing and sensing applications in general, since it is the resolution with which one can determine the exact position of the (spectral) minimum that ultimately defines the sensor's

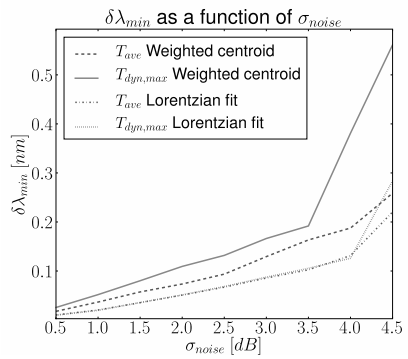


Figure 6.9: Evolution of the minimal resolvable wavelength shift $\delta\lambda_{min}$ as a function of noise, $\delta\lambda_{min}$ determined using the weighted centroid algorithm and the Lorentzian fitting algorithm. The number of Monte Carlo simulations was 1000*.*

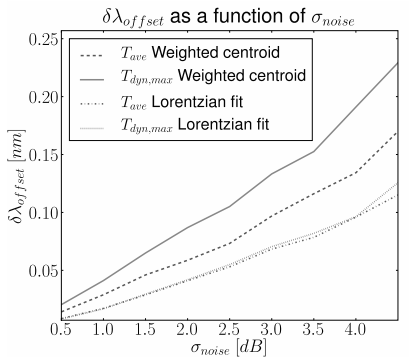


Figure 6.10: Evolution of the wavelength offset $\delta\lambda_{offset}$ as a function of intensity noise, $\delta\lambda_{offset}$ determined using the weighted centroid algorithm and the Lorentzian fitting algorithm. The number of Monte Carlo simulations was 1000*.*

detection limit. Although noise sources should be avoided as much as possible, inherently every measurement will be noisy to some extent. How this noise propagates through the data-analysis method employed determines the detection limit. Six different minima finding algorithms were discussed, their sensitivity to variation in the threshold value were analyzed. Of these six methods, two were withheld, their dependence on noise was simulated using Monte-Carlo simulations. The Lorentzian fitting method was found to be the most robust method in the presence of noise. Since actual experiments proved to be very noisy, this method was withheld to be used in the sensing experiments. The detection limit is ultimately limited by the magnitude of sensor noise.

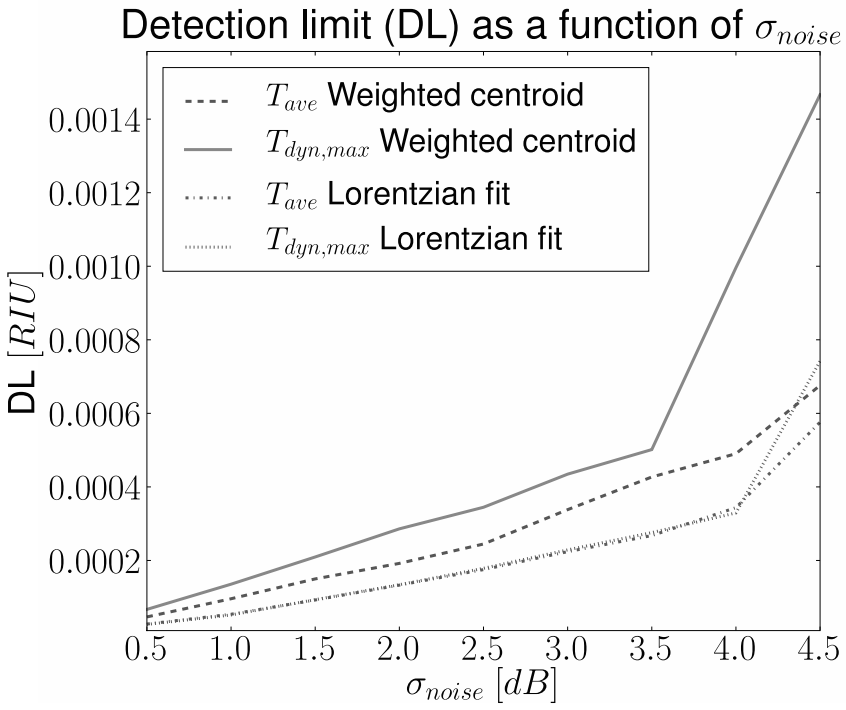


Figure 6.11: Evolution of detection limit as a function of noise, determined using the weighted centroid algorithm and the Lorentzian fitting algorithm. The number of Monte Carlo simulations was 1000.

References

[1] Knut Johansen, Ralph Stalberg, Ingemar Lundtröm, and B. Liedberg. *Surface plasmon resonance: instrumental resolution using photo diode arrays*. Measurement Science and Technology, 11:1630–1638, 2000.

[2] K.S. Johnston, S.S. Yee, and K. S. Booksh. *Calibration of surface plasmon resonance refractometers using locally weighted parametric regression*. Anal. Chem., 69(10):1844–1851, 1997.

[3] T. M. Chinowsky, L. S. Jung, and S. S. Yee. *Optimal linear data analysis for surface plasmon resonance biosensors*. Sensors and Actuators B -Chemical, 54:89–97, 1999.

[4] K. Kurihara, K. Nakamura, and K. Suzuki. *Asymmetric SPR sensor response curve-fitting equation for the accurate determination of SPR resonance angle*. Sensors and Actuators B -Chemical, 86(1):49–57, 2002.

[5] C Thirstup and Zong W. *Data analysis for surface plasmon resonance sensors using dynamic baseline algorithm*. Sensors and Actuators B -Chemical, 106:796–802, 2005.

[6] J Homola. *Surface plasmon resonance based sensors*, volume 4 of *Springer Series on Chemical Sensors and Biosensors*. Springer, Berlin, 2006.

7 Sensing Experiments

THIS chapter gives an overview of the measurements performed on integrated surface plasmon interferometer devices. The measurement procedure is first explained in the first section. the following two section deal with the bulk sensing experiments and with a surface sensing experiment respectively. The latter results are preliminary in the sense that only one single measurement run was performed, due to both time and sample constraints. In the course of this work, measurements were performed both on Si-integrated surface plasmon interferometers that were back-lapped and devices that were polished.

7.1 Measurement setup

The horizontal measurement setup is shown schematically in figure 7.1, and a picture of the actual measurement setup is provided in figure 7.2.

Laser light is coupled to a polarization-maintaining fiber, which is connected to to a lensed fiber for incoupling. The lensed fiber is not a polarization maintaining fiber, hence an additional polarization controller is needed. This controller induces polarization changes in an optical fiber by applying torque on this fiber. The sample containing the surface plasmon interferometer sensors is mounted on a 2-D translation stage (lateral and vertical translation) with sub micron controls. The sample is held in place by a vacuum chuck. The lensed fiber is also mounted on a 3-D translation stage,

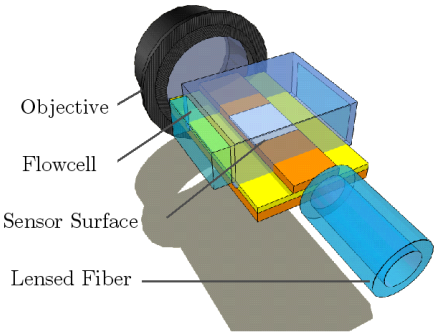


Figure 7.1: Schematical representation of the measurement setup, depicted is the SPI sensor (embedded Au layer in the sensor region), the incoupling lensed fiber and outcoupling objective. Analyte fluid is delivered through the flowcell on top of the sensing area.

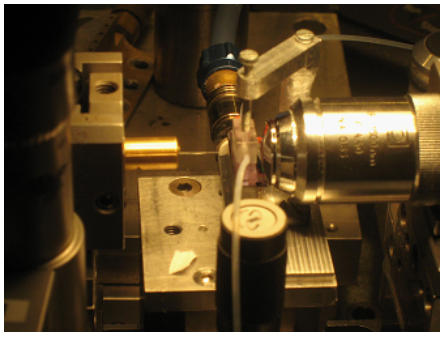


Figure 7.2: Photograph of the actual measurement setup. The flowcell is kept into position using a screw, tubing connects the cell with the syringe pump and a waste container.

controlled by piezoelectric actuators with feedback controls for accurate and stable alignment. Coarse alignment is possible through a binocular microscope mounted above the sample. The spot of the lensed fiber is focused on a cleaved facet of a 10 μm wide ridge waveguide. This waveguide then tapers down to a 3 μm broad ridge waveguide containing the sensor region. After tapering out to a broad output waveguide, the transmitted light is collected by an objective lens and focused onto a power detector. This objective lens is also mounted on a 3-D translation stage.

The power detector is connected to a measurement PC through GPIB, and is controlled through Python using pyVISA, so automatic wavelength sweeps are possible. To avoid stray light from falling onto the power detector, an adjustable pinhole aperture is placed in front of the detector. During alignment the detector can be removed from the optical path, so the image of the outcoupling facet can be projected to an infrared CCD camera, facilitating alignment of the setup.

During flow cell experiments, a flow cell is placed on top of the sample. This cell is held in place mechanically with a screw attached to the vacuum chuck holder. Tubing connected to this flow cell provide the in-and outputs of the analyte to be measured. Flow rate is controlled through a syringe pump.

As a function of wavelength the transmission will be strongly oscillating.

This is due to the fact that a tunable laser is used for scanning the wavelength range, combined with the fact that cleaved facets and the lensed fiber/ objective create a cavity. Since the laser has a fixed linewidth, the only way of determining the correct transmission curve is by scanning the wavelength range with a sufficiently fine resolution. Scanning the whole range takes time, making this a slow measurement technique [1].

Alternatively a broadband source could be used in combination with an optical spectrum analyzer. The advantage of this approach is that the entire spectrum can be monitored in real time. The sensor detection limit then depends on the resolution of the optical spectrum analyzer. The disadvantage is that the output power of broadband sources is far less than the output power of tunable lasers. Due to the high losses, both during in- and outcoupling and during propagation of the optical power through the sensing region, this approach was deemed unpractical.

7.2 Bulk Refractive Index Sensing: Flowcell Experiment

7.2.1 Proof-of-Principle Measurement

Sample dimensions for proof-of-principle measurements were, $d_{Si} = 150$ nm, $d_{Au} = 37$ nm and $L = 8$ μm. The thickness of the Au layer is the optimal possible value within the constraints posed due to fabrication issues (thickness of the Si layer in the sensing section is fixed at 150 nm and cannot be optimized). Deionized water, an aqueous solution of $2m\%$ of $NaCL$ and a solution of $10m\%$ of $NaCL$ were flown over the sensor surface, using a flow cell to avoid evaporation. At 20°C the refractive index of an aqueous solution of $NaCl$ varies with 0.0017151 RIU (Refractive Index Units) per mass $\%$ [2, 3]. The spectra for these measurements are presented in Fig. 7.3, transmission was measured fiber to detector. Standard deviation error bars represent measurements of several spectra with an interval of 10 min to ensure that the steady-state response was measured. A two-point moving average filter and spectral normalisation has been utilised to analyse he spectral data. Sensitivity S_{bulk} $[nm/RIU]$ for this sensor was determined by calculating the wavelength of destructive interference using the dip finding algorithms presented earlier, the standard deviation error σ is partly due to the dip finding algorithm (dependence of the dip position as a function of treshold value is taken into account, so the values shown here are calculated according to the $T_{threshold} = T_{ave}$) and partly due to measurement noise.

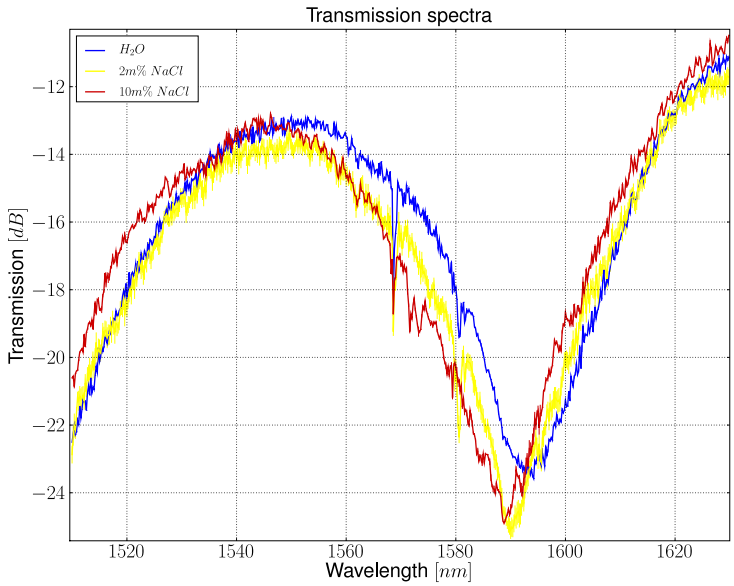


Figure 7.3: Proof-of-principle measurement, overview of the measured transmission spectra for different $NaCl$ concentrations, spectra averaged over three consecutive measurements and subtracted from a reference signal.

Using this data we can estimate the general sensitivity S_{bulk} of this device to be 245.837 nm/RIU. This value was obtained by calculating the slope of the regression line of the measured spectral minima as a function of the refractive index of the analyte. The standard deviation of this calculated slope σ_S is also calculated. A confidence interval for the slope of the regression line can be computed from the estimate S_{bulk} using this computed standard deviations and the appropriate critical value t^* from the $t(n-2)$ distribution in a t-table[1], where n is the number of degrees of freedom[2] [4]. The confidence interval for the sensitivity value takes the form $S_{bulk} \pm t^* \sigma_S$. The 95% confidence interval for this proof-of-principle measurement was $234.677 < S_{bulk} < 256.996$. The wavelength of minimal transmission versus the refractive index of the analyte flown over the sensor surface is depicted in figure 7.4, exact values are given in table 7.1. This confidence level is basically a measure for the linearity of the sensor response. Detection

[1] t-table taken from: www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf

[2] Assuming that the measurement samples come from a Gaussian (normal) distribution, with an unknown mean and deviation.

limits in this table have been calculated using the calculated mean value of the regression line slope S_{bulk}.

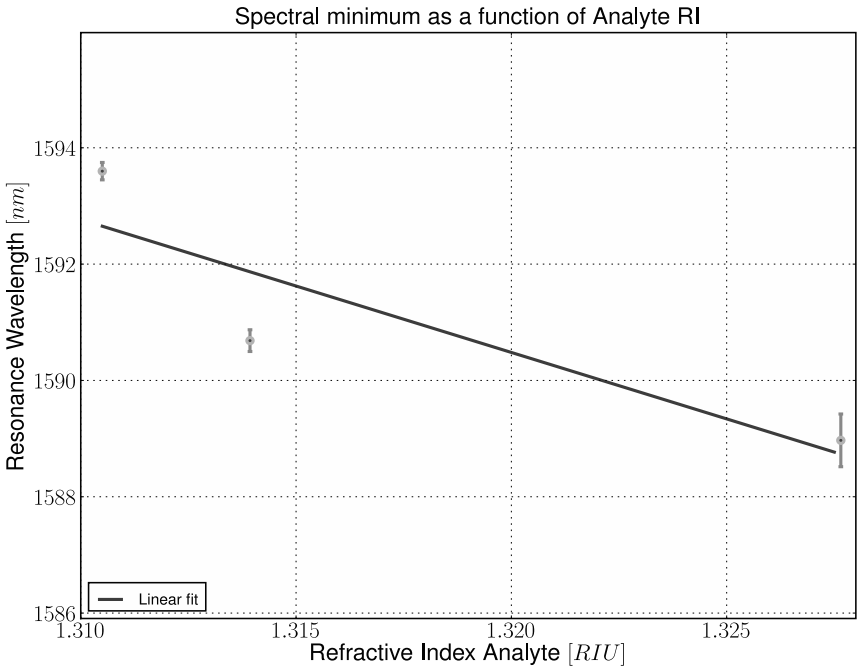


Figure 7.4: Proof-of-principle measurement, λ_{min} as a function of $NaCl$ concentration, the solid line is a linear fit to calculated λ_{min} values.

The observed positions of the maximum attenuation and sensitivity are in agreement with the full wave simulations [5] as can be seen in Fig. 6.3. However, experimental sensitivity is substantially lower than theoretically predicted values, although too few data points are available to make reliable predictions. Therefore we decided to perform additional measurements, which will be described below.

7.2.2 Determination of bulk sensitivity

To obtain a quantitative value for the bulk sensitivity, additional bulk sensing experiments, with a larger array of $NaCl$ concentrations, were performed. The three measured samples were nominally identical, and the experimental conditions were similar as with previously obtained results. Transmission spectra are shown in figures 7.5, 7.6 and

Sensitivity S_{bulk}in nm/RIU is shown in figures 7.8, 7.9 and 7.10, the exact values for sensitivity, spectral minima wavelengths and detection limits

Analyte	RI	$\lambda_{min}[nm]$	$\delta\lambda_{min}[nm]$	DL$[RIU]$
H_2O	1.3105	1593.663	0.079	3.233×10^{-4}
$2m\%NaCl$	1.3139	1590.709	0.159	6.486×10^{-4}
$10m\%NaCl$	1.3276	1588.743	0.476	1.939×10^{-3}
Sensitivity: 245.837 **nm/RIU**				

Table 7.1: Proof-of-principle measurement, λ_{min} as a function of $NaCl$ concentration, values calculated with the Lorentzian curve fitting algorithm

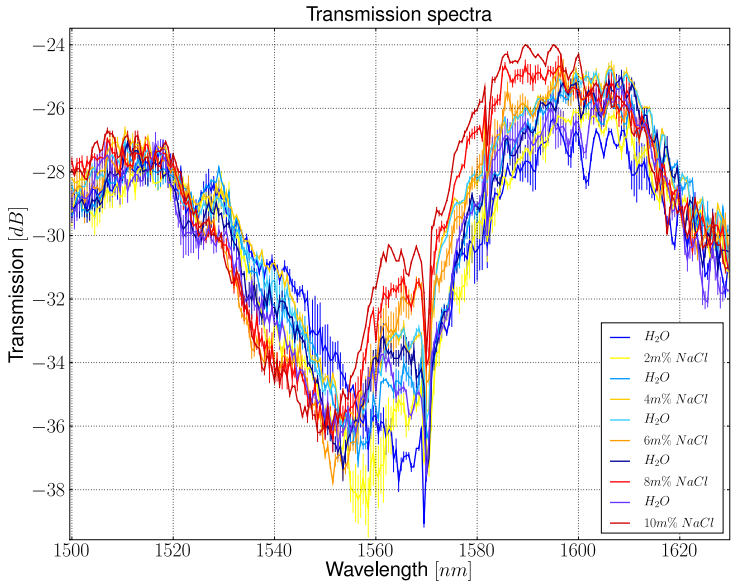


Figure 7.5: Measurement run 1, overview of the measured transmission spectra for different $NaCl$ concentrations, spectra averaged over three consecutive measurements and subtracted from a reference signal.

can be found in tables C.1, C.2 and C.3 in appendix C. Confidence intervals and detection limits were calculated in a similar fashion as earlier.

7.2.3 Summary bulk sensing results

To begin this summary we will briefly recap the theoretically obtained bulk sensitivity and detection limit values for this specific sample. In table 7.2 we list the theoretically calculated value based on Mach-Zehnder interferometer

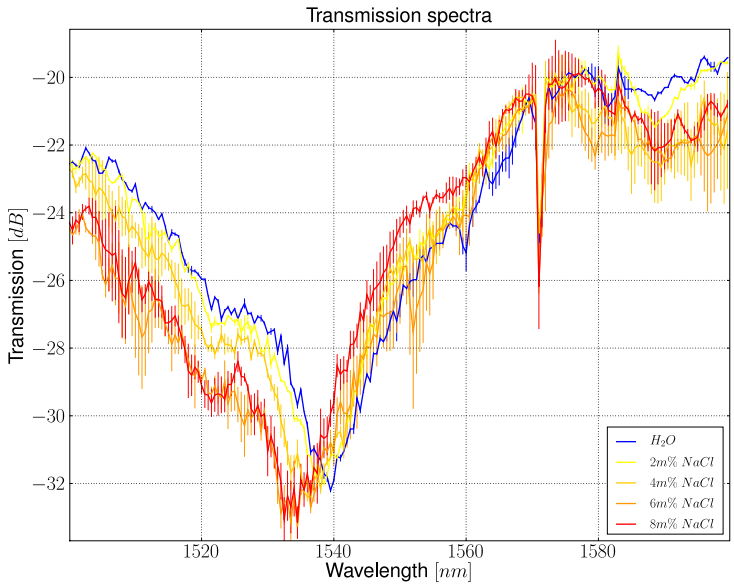


Figure 7.6: Measurement run 2, overview of the measured transmission spectra for different $NaCl$ concentrations, spectra averaged over three consecutive measurements and subtracted from a reference signal.

theory, the calculated value using the full-wave simulation and Lorentzian datafit algorithm presented in chapter 6 and the experimentally obtained values. Sensitivity values shown are the mean values of the linear regression. Sensitivity S_{bulk} varies significantly among these three measured samples, from as high as 580.962 nm/RIU (sample 3), to as low as 433.605 nm/RIU (sample 2) and 435.454 nm/RIU (sample 1). All of these values are remarkably higher than the previous estimate of 245.837 nm/RIU, but even more peculiar, are also significantly larger than the theoretically obtained sensitivity. While the earlier calculated 95% confidence intervals give some estimate of the accuracy of the calculated sensitivity values, the linear regression algorithm does not take the uncertainty on the wavelengths of minimal interference into account. To calculate the influence of the uncertainty values of the wavelength minima, we have added Gaussian white noise with a standard deviation equal to the calculated standard deviation of the spectral minima to the sensitivity curves and calculated the value of the sensitivity S_{bulk} in a Monte-Carlo simulation. The resulting sensitivity values are normally distributed, the mean value and standard deviation of these distributions are

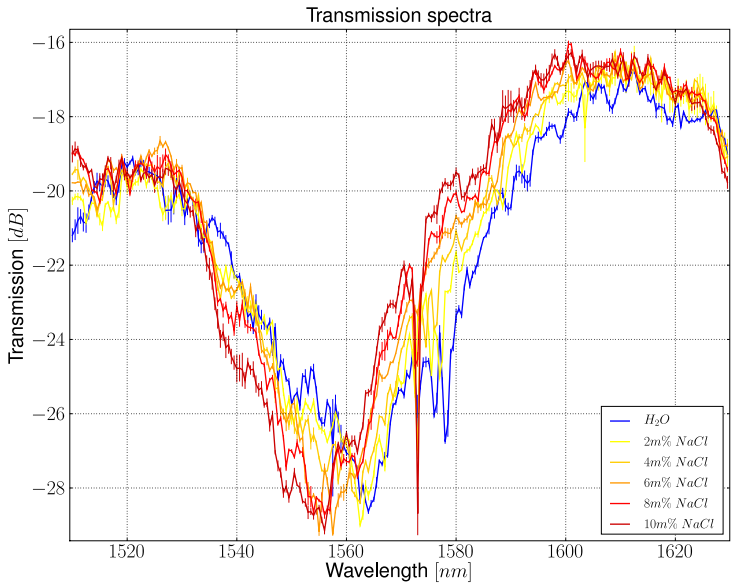


Figure 7.7: Measurement run 3, overview of the measured transmission spectra for different $NaCl$ concentrations, spectra averaged over three consecutive measurements and subtracted from a reference signal.

shown in table 7.3. The probability density functions of these Monte-Carlo simulations are shown in figure 7.11

From figure 7.11 it is clear that the first and second measurement run have a considerable overlap with each other, but also with the theoretically determined sensitivity value (both the approximative value and the full-wave calculated value). Calculated sensitivity of the proof-of-principle measurement is too small, while the third measurement run has a very high sensitivity. Both probability density functions do not overlap with theoretical values. While the proof-of-principle measurement contains too few data point to be of any statistical relevance, the calculated sensitivity value of the third measurement run cannot be discarded. Why this measurement has a sensitivity that is higher than the two previous measurement runs remains an open questions, as the three samples are nominally identical.

The exact position of the spectral minimum as a function of analyte concentration is difficult to predict theoretically, and also within the same sample batch there are significant differences. As a illustrative example, the wavelength of destructive interference for H_2O is different for all all four

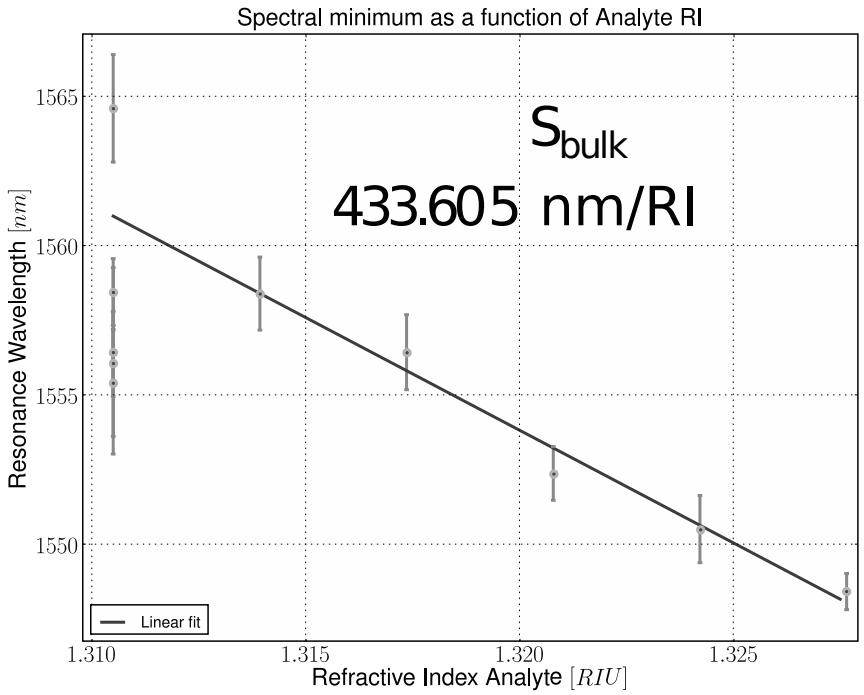


Figure 7.8: Measurement run 1, λ_{min} as a function of $NaCl$ concentration, the solid line is a linear fit to calculated λ_{min} values.

measured samples.

Slight variations in the physical dimensions may be responsible for this difference, be it in the thickness of the Au layer, or the length of the developed section when processing the BCB layer. SEM pictures of these samples are not conclusive due to the low contrast between the Au and cladding layers, this makes is difficult to accurately determine both Au thickness and sensing region length L. These two parameters are the only two microscopic differences between the otherwise identical samples. A slightly different phase shift upon coupling from the dielectric input mode to both surface plasmon modes can shift the position of the minima too either larger or smaller values, while the sensitivity would remain the same.

Gold roughness might also be a contributing factor to the deviation of the experimental result as compared to the theoretical values. Surface roughness contributes both to the position of the wavelength of destructive interference for H_2O and the obtained sensitivity values. Due to a lack of time the gold roughness of these samples was not characterized.

A third and possible cause for the discrepancy obtained here could be

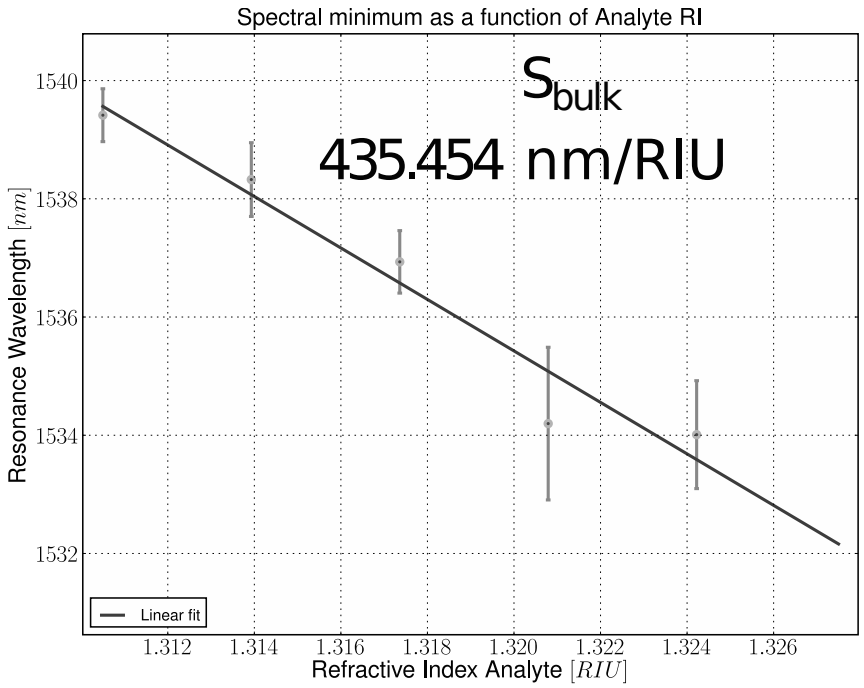


Figure 7.9: Measurement run 2, λ_{min} as a function of $NaCl$ concentration, the solid line is a linear fit to calculated λ_{min} values.

the thermo-optic effect that causes the refractive index of the silicon and silicon-oxide to vary as a result of temperature fluctuations. A difference in refractive index would lead to a difference in propagation constant, again dramatically influencing sensor performance. Temperature controlled experiments were unfortunately not performed in this thesis.

Measured extinction ratios between constructive and destructive interference all fall within the range $13 - 16\ dB$, this value is mostly limited by polarization coupling in the waveguide itself and by performance of polarizing optics [6].

7.3 Avidin-Biotin sensing

Now that sensitivity to bulk refractive index changes, though not yet fully understood, has been obtained. The next step is to test the sensitivity to the adsorption of thin biomolecular layers. Samples were chemically functionalized with Biotin (using the process described in 5.6), and Phosphate buffered saline (PBS) buffer solution, PBS buffer solution with a concentration of

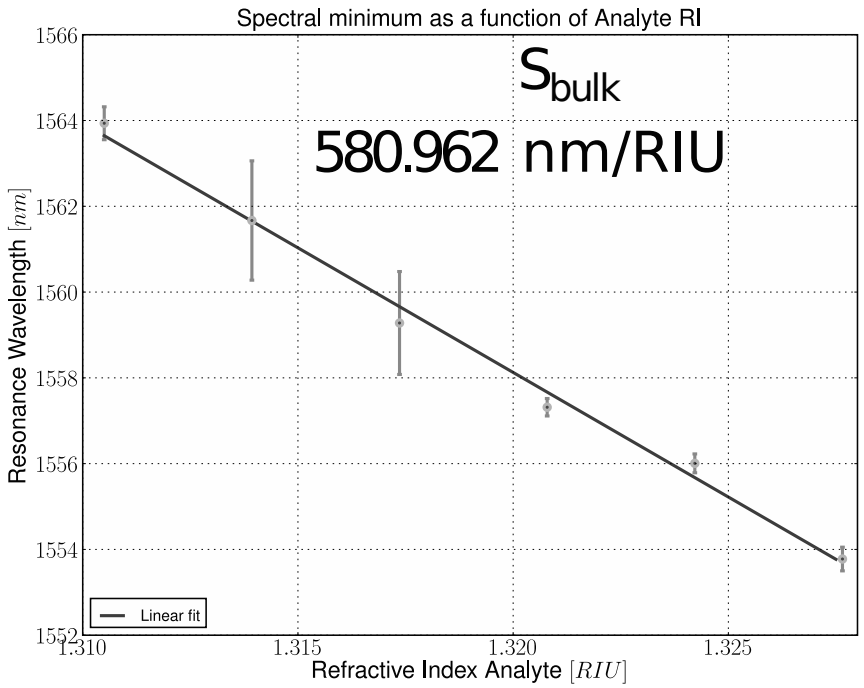


Figure 7.10: Measurement run 3, λ_{min} as a function of $NaCl$ concentration, the solid line is a linear fit to calculated λ_{min} values.

32.5 $\mu g/ml$ of Avidin, and again a PBS buffer solution were flown over the sample consecutively. The position of the spectral minimum was tracked as a function of time and is shown in figure 7.12

This device under study has a theoretical sensitivity of 120 pm/nm of additional adlayer thickness. This result is based on the approximative formula 4.35 ($n_{Avidin-Biotin} \approx 1.455$, adlayer thickness $1 - 20$ nm), and is well below the noise level in this specific experiment. As such, no measurable shift in resonance wavelength is to be expected. Since the size of an Avidin molecule is about 10 nm, the total shift assuming full coverage of the shift would be approximately 1.2 nm. Noise levels comparable to the ones observed in the proof-of-principle experiment would be required in order to draw conclusions. Nevertheless, the spectral minimum does shift, as can be seen in figure 7.12, surprisingly, instead of a blueshift, we obtain a redshift when the PBS-Avidin solution is flown over the sensor surface. After the consecutive second PBS step, to distinguish bulk refractive index shift from adlayer shift. the minimum is blue shifted again to lower wavelengths, the net result being a very small redshift versus the original baseline signal,

	Sensitivity $[nm/RIU]$	Detection Limit $[RIU]$
Theory	391	
Full-wave simulation	382.597	$1.670 \times 10^{-5} - 1.152 \times 10^{-5}$
Proof-of-principle	245.837	$1.939 \times 10^{-3} - 3.233 \times 10^{-4}$
1	433.605	$1.451 \times 10^{-2} - 9.961 \times 10^{-4}$
2	435.454	$2.965 \times 10^{-3} - 1.027 \times 10^{-3}$
3	580.962	$2.394 \times 10^{-3} - 3.128 \times 10^{-4}$

Table 7.2: Comparison of the different calculated and experimentally obtained sensitivity values and detection limits.

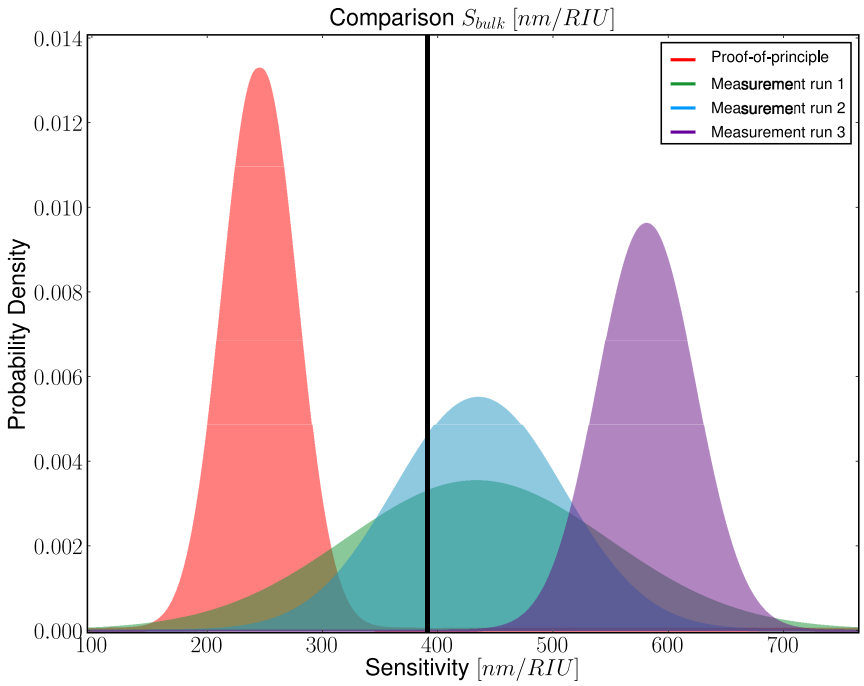


Figure 7.11: Probability density functions of the Monte-Carlo simulation of the sensitivity value S_{bulk} for the four experiments. The black vertical line indicates the theoretically determined sensitivity value.

although this shift is significantly smaller than the experimental errors also shown in figure 7.12. While theoretical considerations do allow the possibility of a negative surface sensitivity $S_{surface}$ for sufficiently thin Au layers, the thickness of the Au is sufficiently large to exclude this possibility. Also note the slope of the curve in figure 7.12. The spectral minimum appears to manifest itself for ever slightly smaller wavelengths as a function of time.

Measurement	$S_{bulk}[nm/RIU]$	$S_{bulk,MC}[nm/RIU]$	$\delta S_{bulk,MC}$
Proof-of-principle	245.837	245.834	30.160
1	433.605	433.911	113.847
2	435.454	435.541	72.534
3	580.962	580.971	41.552

Table 7.3: Comparison of the mean and standard deviation of the different sensitivities. Values calculated using a Monte-Carlo simulation (1 million iterations).

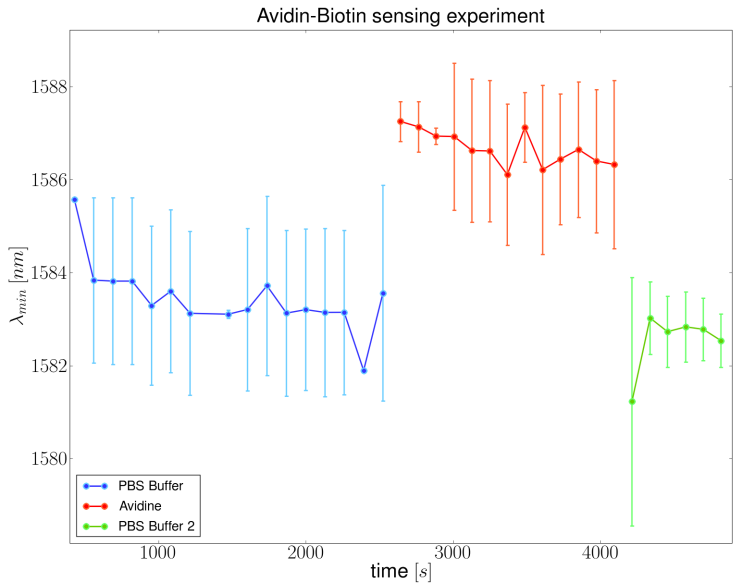


Figure 7.12: sensorgram, λ_{min} as a function of time and solution flown over the sensor surface, blue: PBS buffer solution, red: PBS buffer solution with a concentration of 32.5 $\mu g/ml$ of Avidin, green: PBS buffer

This spectral shift phenomenon could be due to sample degradation since the Au layer is exposed to various substances flowing over the surface. Possible degradation of surface chemistry might be caused by some of the reagents in the analyte containing the Avidin molecules. Temperature fluctuations could also cause this strange behavior, as this experiment could not be temperature controlled. Temperature fluctuations are believed to play a larger role in the Biotin-Avidin experiments since experiments typically take longer, meaning that the sample has more time to heat up as a result of,

and during, the measurement.

Numerical errors in the determination of the spectral minimum could also contribute to this phenomenon since the signal-to-noise ratio of the measured spectra was quite small. Additional experiments would need to be performed in order to quantify adlayer sensitivity of the device.

7.4 Conclusion

Proof-of-principle operation of a surface plasmon interferometer has been acquired. Values obtained for bulk refractive index sensitivity differ substantially in the experiments performed. The samples were nominally identical, and experimental errors could not be held accountable for the large differences in sensitivity. The proof-of-principle experiment contains too few data points to be considered relevant, two out of three consecutive experiments are in agreement both with one another and theoretical values, the third experiment however differs substantially. The cause of this discrepancy has not been determined, though several possibilities identified. More experimental verification is needed to rule out which of these (Au thickness, Au roughness and temperature fluctuations) is responsible.

The biosensing experiment did not provide a reliable proof-of-principle, nor did it provide a value for the surface sensitivity $S_{surface}$ of the device. Odd behavior of the position of the spectral minimum was encountered in this experiment, since redshifts were obtained where blueshifts were expected. Only two different samples could be measured, one of them provided no useable signal, the second one lead to noisy, difficult to interpret spectral data. Ultimately errors on the obtained spectral minima were too large to draw conclusions. Again possible causes for this discrepancy were identified, yet additional measurements need to be carried out in order to fully quantify the biosensing capabilities.

References

[1] W. Bogaerts. *Nanophotonic Waveguides and Photonic Crystals in Silicon-on-Insulator*. PhD thesis, Ghent University, 2004.

[2] D. R. Lide. *CRC Handbook of Chemistry and Physics*. CRC Publishing Company, London, 2004.

[3] Hiu Su and Xu Guang Huang. *Fresnel-reflection-based fiber sensor for the on-line measurement of solute concentration in solutions*. Sensors and Actuators B -Chemical, 126:579–582, 2007.

[4] Y. Nievergelt. *Total least squares: state-of-the-art regression in numerical analysis*. SIAM Review, 36(2):258–264, 1994.

[5] P. Debackere, S. Scheerlinck, P. Bienstman, and R. Baets. *Surface plasmon interferometer in silicon-on-insulator: novel concept for an integrated biosensor*. Optics Express, 14(16):7063–7072, 2006.

[6] J. Homola, J. Čtyrocký, M. Skalský, J. Hradilová, and P. Kolářová. *A Surface Plasmon Resonance Based Integrated Optical Sensor*. Sensors and Actuators B -Chemical, 38-39:286–290, 1997.

8

Conclusion and Outlook

Conclusion

THE objective of this work was to miniaturize, integrate and multiplex current state-of-the art surface plasmon resonance biosensors into Silicon-on-Insulator and as such build a sensitive, reliable and multiparameter biosensor. The potential for mass-fabrication of the developed technology could lead to inexpensive point-of-care devices, accommodating demand in both developed and developing countries.

In a first step, the convergence problems associated with very thin metal layers in 1D simulation were tackled. Adaptive spatial resolution methods, pioneered by Granet, and later improved by Vallius *et al*, were applied to waveguide computational problems. The method was further refined by using perfectly matched layer boundary conditions, a necessary prerequisite to handle numerical problems involving a lot of radiating modes. The new methods enables simulating surface plasmon polariton modes with eigenmode solvers instead of (slow and memory-hungry) FDTD methods, thus drastically reducing the simulation time.

Improving convergence for $2D$ problems proved to be somewhat more difficult. The biggest problem is that the method uses six different Fourier approximations of the dielectric function (or inverse of this function), and smoothening one of these generally has a negative effect on the others. As a result, convergence of the calculated eigenvalues did not improve noticeably. Although perhaps somewhat discouraging, not all the different possibilities have been exhausted here. The mathematical framework presented in this work, and implemented in CAMFR can be easily modified to use other adaptive spatial resolution functions than the choices made in this work.

Using the full-wave electromagnetic solver developed in this work, a novel concept for an integrated surface plasmon biosensor was conceived. Proposed device would be two orders of magnitude smaller than conventional surface plasmon waveguide sensors, due to the integration into a high-index contrast material system. The basic principle allows high tunability, making an excellent candidate for a vast number of applications. The underlying physical principle was investigated in detail and the potential for sensing was numerically calculated. Several parameters, such as waveguide thickness, gold thickness and sensor length, dictate sensor response. Based on these considerations, a set of theoretical design parameters for surface plasmon interferometers were laid out.

The various fabrication steps were carried out, optimized and/or characterized. Ultimately, proof-of-principle samples were fabricated. The design parameters of these samples were far from ideal from a sensitivity point of view, due to several constraints imposed by the fabrication process. More

specifically, the thickness of the Si layer in the sensing region was chosen to be $150\ nm$, which is the standard depth of the shallow etch process ($70\ nm$)in the fabrication process for SOI chips. Nevertheless, interference visibility and sensitivity were sufficient to prove the working principle.

Measurement software was developed, alongside with advanced data analysis methods to estimate the spectral position of the minimum. Robustness of several data-analysis methods was numerically determined, Lorentzian fitting algorithms proved to be the most suitable and noise-insensitive.

Qualitative agreement between measurement results and theoretical data has been obtained in a set of repeated experiments. Furthermore, we have shown this device capable of detecting bulk refractive index changes, a first and indispensable step toward label-free biosensing. The measured blue-shift shows the potential of this device to be used a a sensitive and label-free biosensor.

The authors believe this concept to be a possible and viable alternative toward a fully integrated surface plasmon lab-on-chip solution.

Outlook

The nanophotonic sensor presented in this work is currently at a point in its evolution were the full potential is far from completely realized. While a proof-of-principle for bulk refractometric sensing was presented here, experimental reproducibility remains an open issue. Several different noise sources and unwanted effects such as thermo-optic ones were identified, a thorough analysis of the importance of these different noise sources has not been performed yet. As such the device is far from practical applications but remains a promising concept.

Attempts at proving the applicability for biosensing are inconclusive, although too few experiments could be performed within the timeframe of this work to exclude possible biological applications . A potentially interesting experiment would be to perform these experiments without the chemically modified Au surfaces in order to distinguish between bulk refractive index sensing effects and wavelength shifts originating from binding events at the surface. The largest challenges here remain photonic ones, rather than chemical or microfluidic ones

Guiding and providing the necessary optical power to the sensing structure itself proved to be a larger challenge than originally thought. The current incoupling method (horizontal incoupling) is far from ideal, and originated as a last resort rather than a conscious choice. The development of efficient and broadband grating couplers, together with taper structures for

TM polarization are essential in improving the reproducibility and reliability of surface plasmon interferometer structures.

Since the general principle is quite generic in the sense that it does not rely on the material system it is implemented in, other possible configurations might also work and even boast higher sensitivity values than the SOI device. From a practical point of view this work focused mainly on Silicon-on-Insulator.

Due to limited success in applying the adaptive spatial resolution algorithm to 2D waveguide computational problems, simulations were, and still are, limited to one dimensional situations. While this is a good approximation for 10 and 3 μm broad waveguides, what happens when moving too narrower waveguides remains an interesting question. It is known that a narrower wire width would lead to several distinguishable surface plasmon modes on both the bottom and the top surface (although the top surface could potentially still be monomodal), their interaction and mutual dependence could potentially be interesting for sensing purposes. Exciting and being able to distinguish between multiple propagating surface plasmon modes would provide more information than one single mode since each of these modes will have a different field-profile and thus a different response to either bulk refractive index changes or surface adlayer changes.

A

Au Refractive Index

A theoretical model for the dielectric permittivity of metals is derived here. Experimental results for thin Au films were obtained using a variable angle spectroscopic ellipsometer and agree quite well with presented simple theoretical model in the near-infrared spectral region.

A.1 The Drude Model

In order to understand metal properties in detail, complicated quantum-mechanical treatment is required. In the present context, a classical model for metals, developed by Drude in 1900 [1], is already sufficient for describing many metallic properties. This appendix deals with the the origin of the relative permittivity in metals and, how it changes sign depending on the frequency. The Drude model (or *free-electron model*) is based on kinetic gas theory, It assumes a gas of non-interacting free electrons that have an isotropic effective mass and relaxation time. This means that electrons move freely between independent collisions and the average rate of these collisions γ is connected to the average relaxation time τ, with $\gamma = 1/\tau$. In an external electric field the electrons will obey Newton's law of motions and are accelerated in between collisions, leading to a net drift motion. Assume now that these electrons are subject to a time harmonic electric field (optical wave) of the form $E = E0e^{-j\omega t}$ where ω is the angular frequency

of oscillation. The equation of motion of the electron can be written as

$$\frac{\partial^2 x}{\partial t^2} - \frac{1}{\tau}\frac{\partial x}{\partial t} = -\frac{e}{m_e}\mathbf{E} \tag{A.1}$$

where τ is the average time between collisions for the electron, and its reciprocal $\omega_c = \tau^{-1}$ is referred to as the collision frequency.m_e is the mass of the electron and e the elementary charge. Equation A.1 has the rather simple but nontrivial solution $x(t) = x_0 e^{j\omega t}$, and upon substitution the instantaneous position of the electron is found to be

$$x(t) = \frac{e}{m_e}\frac{1}{\omega^2 + j\frac{\omega}{\tau}}\mathbf{E} \tag{A.2}$$

Polarization $\mathbf{P}$ is defined as the induced dipole moment $p = -ex$ per unit volume so if N is the number of free electrons per unit volume the polarization is

$$\mathbf{P} = -\frac{e^2}{m_e}\frac{1}{\omega^2 + j\frac{\omega}{\tau}}\mathbf{E} \tag{A.3}$$

The polarization for linear dielectrics can be defined in terms of the free-space permittivity ϵ_0, and the electrical susceptibility χ as

$$\mathbf{P} = \epsilon_0\chi\mathbf{E} = \epsilon_0(\epsilon_m - 1)\mathbf{E} \tag{A.4}$$

ϵ_m being the relative permittivity, or dielectric constant of the metal. By combining equations A.3 and A.4 we can find an expression for the frequency dependent dielectric function of a metal :

$$\epsilon_m(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + j\frac{\omega}{\tau}} \tag{A.5}$$

where ω_p is the plasma resonance frequency of the metal, determined by the electron density N, the electron charge e, and mass[1] m_e.

$$\omega_p^2 = \frac{Ne^2}{\epsilon_0 m_e} \tag{A.6}$$

The dielectric function of a metal is complex, and as such can be written as a combination of a real part and an imaginary part

$$\epsilon_r(\omega) = 1 - \frac{\omega_p^2\tau^2}{1 + \omega^2\tau^2} \tag{A.7}$$

$$\epsilon_i(\omega) = \frac{\omega_p^2\tau}{\omega(1 + \omega^2\tau^2)} \tag{A.8}$$

[1] In a modified version of the Drude model an effective electron mass can be defined which is different than the electron mass m_e, corrections are due to the quadratic dispersion relation of a free electron originating from the Pauli exclusion principle.

The electron relaxation time can be related to the conductivity of the metal by the expression [2]

$$\sigma = \tau \frac{Ne^2}{m_e} \tag{A.9}$$

In the Drude model the dielectric function of the metal is defined only by the plasma frequency and the electron collision rate. Typically, the plasma frequency ω_p lies in the UV region, for silver $\omega_p = 260\ nm$ and for gold $\omega_p = 230nm$ (for an electron density of $N \approx 10^{22} cm^{-3}$). The collision rate can be quite rapid, the time interval between individual collision is of the order of femtoseconds. If for some frequency range $(\omega\tau)^2 >> 1$ (which is the case for optical frequencies, $\omega \approx 3 \times 10^{15} s^{-1}$ for $\lambda = 633nm$), the following approximation is valid

$$\epsilon(\omega) \approx 1 - \frac{\omega_p^2}{\omega^2} \tag{A.10}$$

Since the reflectivity R of a planar surface at perpendicular incidence is related to the refractive index n ($n^2 = \epsilon$)through

$$R = \frac{|n-1|^2}{|n+1|^2}, \tag{A.11}$$

above the plasma frequency, the metal becomes effectively transparent ($\epsilon \to 1$, $R \to 0$), while below the plasma frequency all incident radiation is reflected ($\epsilon < 0$), explaining why most metals are so reflective at optical frequencies.

The above sentence also reveals some of the limitations of the classical Drude model. Assuming that all valence electrons can move freely and independently, one comes to the conclusion that metals reflect all incoming radiation for $\omega < \omega_p$, which can not, for example, explain the different colors of metals - the yellowish tint of gold, or the reddish hue of copper. The reason for this is that in reality the electronic band structure mustn't be neglected. At high frequencies, additional effects arise such as interband transitions. The range of validity of the Drude model can be extended by adding Lorentzian terms to Eq. A.5 to obtain the Lorentz-Drude model

$$\epsilon_{m,LD}(\omega) = \epsilon_{m,Drude}(\omega) + \sum_{j=1}^{k} \frac{f_j \omega_j^2}{(\omega_j^2 - \omega^2) - j\omega\gamma_j} \tag{A.12}$$

where ω_j and γ_j stand for the oscillator resonant frequencies and bandwidths respectively, and f_j are weighting factors. Physically (as mentioned earlier),

the Drude and Lorentzian terms are related to intraband (free-electron) and interband (bound-electron) transitions respectively. Even though the Lorentz-Drude model extends the range of validity of analytical approximations to metallic dielectric constants, it is not suitable for the description of sharp absorption edges observed in some metals, unless a very large number of terms is used.

Nevertheless, for many metals at optical frequencies the Drude model often provides a good agreement with experimental observations, especially in the telecom wavelength region where dispersion is relatively weak and linearly dependent on wavelength. The complex refractive index of a number of fairly common metals and a number of dielectrics for comparison is shown in figure A.1 for a wavelength of $1550\ nm$.

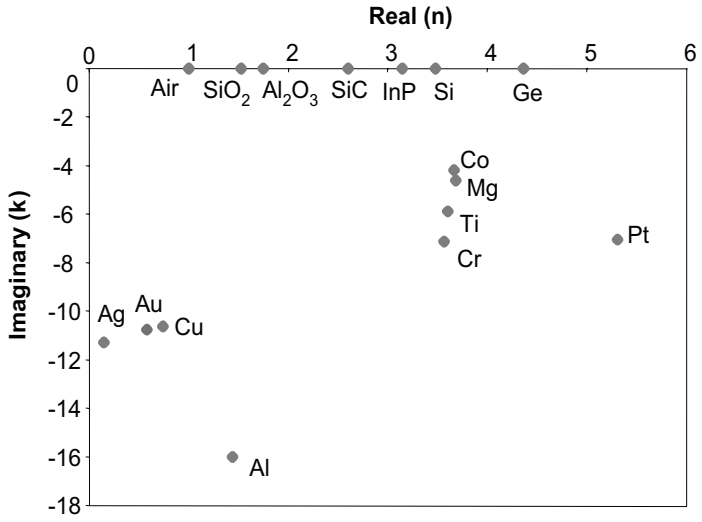


Figure A.1: Real and imaginary part of the refractive index of various metals at a wavelength of $1550\ nm$ *[3].*

A.2 Experimental Data

Figure A.2 shows the principle of operation of ellipsometry. Input light is linearly polarized at $45°$ with respect to the s and p directions. Reflection upon the sample results in a change in the polarization ellipse. If the sample is made of isotropic materials the reflection is determined by two coefficients r_s and r_p (in the case of reflection on a single substrate these are Fresnel coefficients). In ellipsometry, the polarization ellipse is typically

expressed in terms of the phase shift Δ between s and p components, and the amplitudes ratio angle Φ defined as $tan(\Phi) = E_s/E_p$. These angles provide the ratio between the amplitudes reflected in p and s components by means of a complex value ρ defined as:

$$\rho = tan(\Phi)e^{j\Delta} = \frac{E_{rp}}{E_{rs}} = \frac{r_p}{r_s} \tag{A.13}$$

where the final equality arises because incident amplitudes in s and p components are equal. The values of (Φ, Δ) are very sensitive to the angle of incidence (θ) and the wavelength (λ). Ellipsometry is an indirect problem in

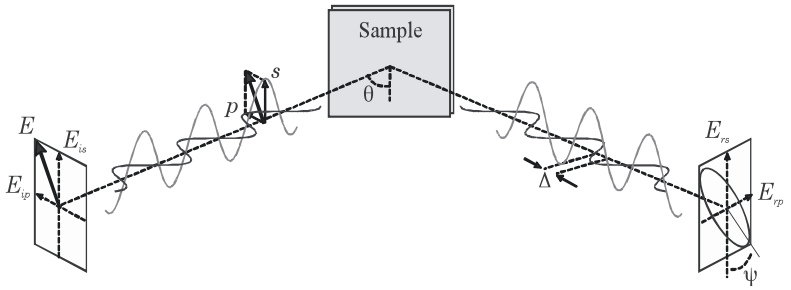


Figure A.2: Principle of ellipsometry. The phase shift Δ and the amplitude's ratio angle Φ depend on the characteristics of the sample and on the angle of incidence θ and the wavelength λ.

the sense that the measured quantities have no direct meaning. To extract the optical properties (n and k) one must evaluate Φ and Δ over the wavelength range $1.5 - 1.7\ \mu m$ for variable angles, and solving the inverse ellipsometric problem.

In this work a VASE type (Variable Angle Spectroscopic Ellipsometers) ellipsometer from the J.A. Woollam Co., Inc with rotating analyzer was used. The spectral range is 230 to 1700 nm. For the evaluation the WVASE software is used. Results are shown in figure A.3 and A.4. The values are compared with the simple Drude model that was described in the previous section, with experimental data taken from Palik [4], and the model that has been used in this work for all numerical simulations.

While the measured values for the real part of the refractive index in the VASE experiment become very noisy, based on the Drude model and Palik's experimental data the conjecture can be made that the data should be linear in this region. That is why a linear fit was of the VASE data was used in this work. The imaginary part of the effective index is relatively noise-free (as compared to the real part), and as such can easily be approximated by a linear fit in the wavelength range $1500\ nm$ - $1650\ nm$.

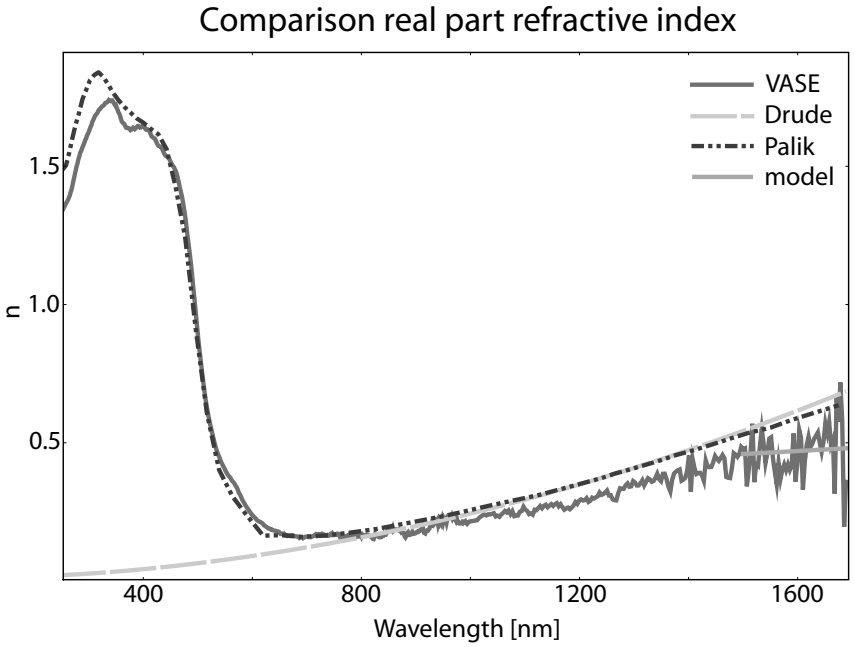


Figure A.3: Comparison real part refractive index determined by VASE measurements, Drude theoretical values, Palik's experimental data and the model used in this work.

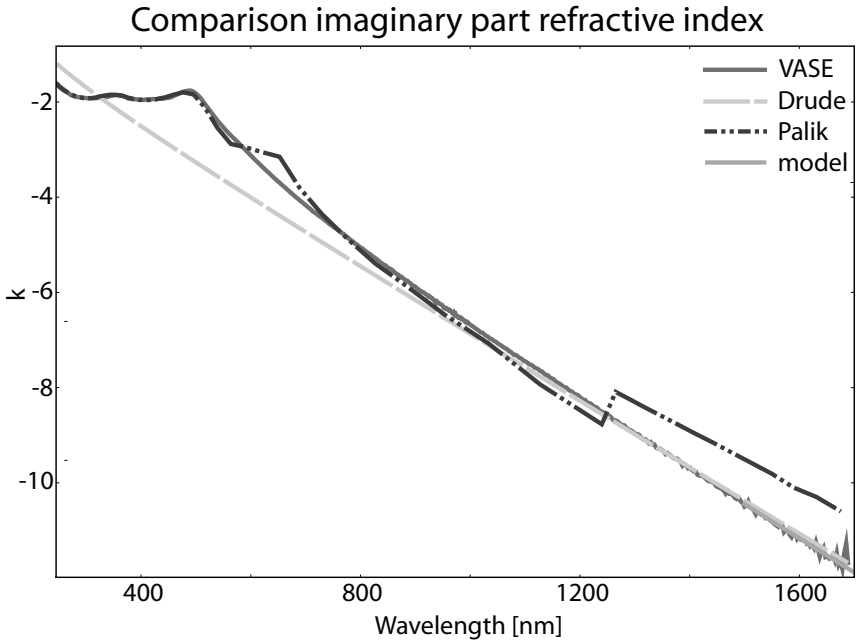


Figure A.4: Comparison imaginary part refractive index determined by VASE measurements, Drude theoretical values, Palik's experimental data and the model used in this work.

References

[1] P. Drude. *Zur Elektronentheorie I*, volume 566. Annalen der Physik 1, 1900.

[2] M. A. Ordal, R. J. Bell, R. W. Alexander, Long J.L., and M.R. Querry. *Optical properties of fourteen metals in the infrered and far-infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V and W*. Applied Optics, 24:4493–4499, 1985.

[3] S. Scheerlinck. *Metal grating based interfacing to and probing of nanophotonic circuits*. PhD thesis, Ghent University, 2009.

[4] E. L. Palik. *Handbook of Optical Constants of Solids*. Elsevier, New York, 1985.

B

Modal properties wire waveguides

UNLIKE Si rib waveguides, which can be monomodal, even for cross-sectional dimensions of several micrometers, wire waveguide dimensions must be significantly smaller than $1\ \mu m$ to suppress propagation of higher-order modes. The results in this appendix were obtained with $2D$ simulations for wire waveguides (or Si wires as they are commonly called) and suggest that the waveguides can be multimode even for relatively small dimensions. This is particularly the case for TM polarization. The modal characteristics of SOI wire waveguides have been extensively studied for waveguides with a symmetrical top cladding layer [1], for an asymmetric environment little research has been performed.

Simulation of the real part of the effective index of the waveguide modes as a function of waveguide width is shown in figure B.2. The cross-section of the simulated SOI wire waveguide is shown in figure B.1. The material stack used throughout this work has a $220\ nm$ thick amorphous Si core layer, separated from the Si substrate by a $2\ \mu m$ thick amorphous SiO_2 bottom cladding layer. The top cladding layer is either air, H_2O or BCB.

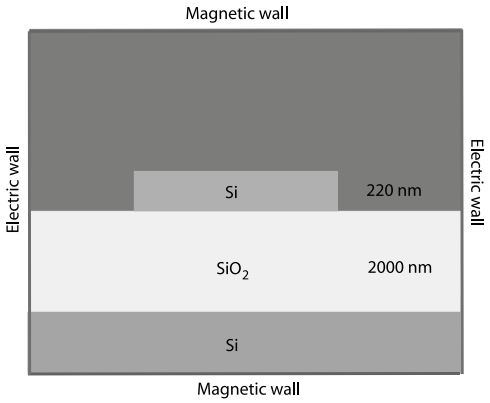


Figure B.1: Cross section of an SOI wire waveguide, indicated are the typical dimensions and the boundary conditions used here.

Hard wall boundary conditions were used, rather than perfectly matched layer (PML) boundary conditions which are often used when calculating leaky or radiative modes. Using hard (electric-magnetic) wall boundary conditions, all the modes can be calculated through a real mode solving engine, which is substantially faster than the complex mode solving engine required when using PML boundary conditions. Although only real propagation constants are obtained, from the real propagation constant it is easy to determine if a mode is guided or not [2]. The numerical results were obtained using Fimmwave [3], a fully vectorial mode expansion tool.

The wire supports transverse electric (TE) like and transverse magnetic (TM) like modes, all of which are actually hybrid modes. The TE-like modes are dominated by an electric field component in the plane, the TM like modes are dominated by their magnetic in-plane H-field. The level of hybridness depends on core cross-section. A square-cross section, symmetrically clad wire has the most hybrid modes. Nomenclature in figure B.2 is derived from the properties (more specifically the number of maxima minus one in the vertical direction and the horizontal direction) of the dominant E or H field component at a waveguide width of $3\ \mu m$.

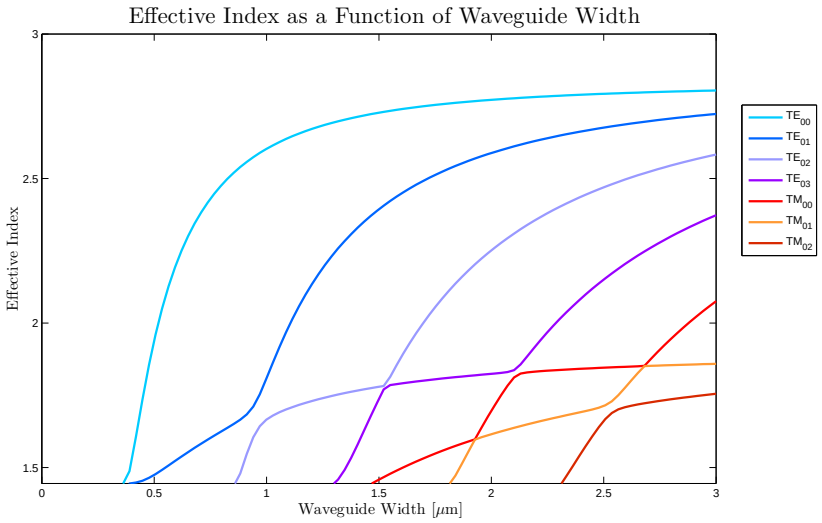


Figure B.2: Simulation of the effective index of the wire modes as a function of core width, at a wavelength of $1550\ nm$*. The indices indicate horizontal mode numbers as the waveguides are vertically monomode. The* Si *layer thickness was* $220\ nm$*, the top cladding was air(*$n = 1.0$*)*

From figure B.2 it is apparent that the SOI wire waveguide is very birefringent, with the TM modes having the lower effective indices. The TE_{00} mode is the first to be guided as the width increases, then the TM_{00} mode, the TE_{01} mode, the TE_{02} mode and the TM_{01} mode in turn. The intersection points of the effective index curves with the bottom of the plot window represent the cut-off widths of the corresponding modes. Due to the asymmetric nature of the surrounding cladding layers, whenever the effective index of a TM guided mode crosses the effective index of a TE guided modes, a strong anti-coupling effect occurs resulting in a mini-stop band between the two modes. Both the TE and TM mode are completely hybrid at this anti crossing point. This effect is typical for asymmetric cladding layers and does not manifest itself in a symmetrical environment

Upon tapering from a wire waveguide with a $10\mu m$ broad core, the effective index of the waveguide modes changes according to the lines depicted in figure B.2, if the taper satisfies the adiabatic criterium. Since the TE_{00} mode does not cross any other modes while going from $10\mu m$ to a single mode waveguide, tapering down will be very efficient for TE modes. The TM_{00} mode however encounters several higher order TE modes, at each crossing the two modes become completely hybrid, and a mini-stopband is exhibited. The narrower the wire waveguide, the larger this

stopband becomes. Therefore, tapers, even the ones that satisfy the adiabatic criterium, cannot and will not work in TM polarization. The input TM mode from a broad wire waveguide evolves into higher order TE modes, these modes have a cutoff-width well above the single mode width for TM polarization in wire waveguides, hence high tapering losses.

The influence of the BCB cladding layer on the effective indices of the waveguide modes can be seen in figure B.3. Generally, the effective index of all modes (both TE and TM like) increases, due to the higher refractive index of the top cladding layer. Since the entire setup has become somewhat more symmetrical, the magnitude of the mini-stopbands at the anti crossing points is reduced, which should result in a larger efficiency when tapering adiabatically from a $10\ \mu m$ broad waveguide. Still, tapering down to the single mode width is unfeasible, since the last encountered anti crossing of the TM_{00} mode and the TE_{01} mode still leads to a relatively large hybridization of the modes and consequently a large stopband (although still smaller than in the previous case.)

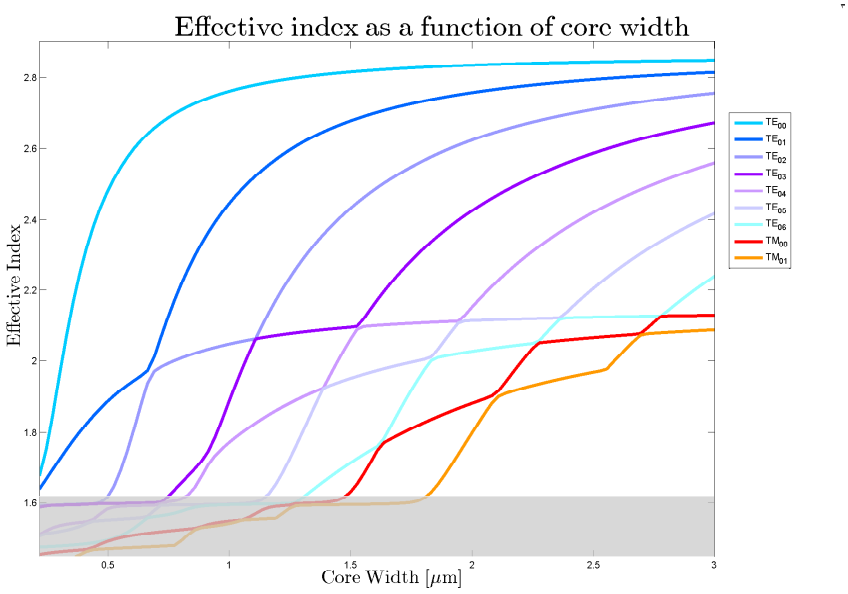


Figure B.3: Simulation of the effective index of the wire modes as a function of core width, at a wavelength of $1550\ nm$*. The indices indicate horizontal modenumbers as the waveguides are vertically monomode. The* Si *layer thickness was* $220\ nm$*, the top cladding is BCB (*$n = 1.6279$*), the shaded region is in between the BCB and the* SiO_2 *lightine, these are leaky modes.*

References

[1] Yurii Vlasov and Sharee J McNab. *Losses in single-mode silicon-on-insulator strip waveguides and bends*. Optics Express, 12(8):1622–1631, 2004.

[2] X. Xu, S. Chen, J. Yu, and X. Tu. *An investigation of the mode characteristics of SOI submicron rib waveguides using the film mode matching method*. J. Opt. A: Pure Appl. Opt., 11:015508, 2009.

[3] *Fimmwave* `http://www.photond.com/products/fimmwave.htm`.

C

Bulk sensing measurement results

Analyte	RI	$\lambda_{min}[nm]$	$\delta\lambda_{min}[nm]$	DL$[RIU]$
H_2O	1.3105	1563.744	1.071	2.471×10^{-3}
$2m\%NaCl$	1.3139	1555.959	6.397	1.475×10^{-2}
H_2O	1.3105	1557.821	1.681	3.877×10^{-3}
$4m\%NaCl$	1.3174	1554.752	0.432	9.961×10^{-4}
H_2O	1.3105	1554.659	1.820	4.198×10^{-3}
$6m\%NaCl$	1.3208	1552.360	2.187	5.044×10^{-3}
H_2O	1.3105	1553.567	2.140	4.935×10^{-3}
$8m\%NaCl$	1.3242	1550.767	1.856	4.279×10^{-3}
H_2O	1.3105	1550.591	6.292	1.451×10^{-2}
$10m\%NaCl$	1.3277	1547.823	0.455	1.049×10^{-3}

Sensitivity: 433.605 **nm/RIU**

$95\%interval$**:** $424.607 < S_{bulk} < 442.607$ **nm/RIU**

Table C.1: Measurement run 1, λ_{min} as a function of $NaCl$ concentration, values calculated with the Lorentzian curve fitting algorithm

Analyte	RI	$\lambda_{min}[nm]$	$\delta\lambda_{min}[nm]$	DL$[RIU]$
H_2O	1.3105	1539.414	0.447	1.027×10^{-3}
$2m\%NaCl$	1.3139	1538.325	0.624	1.433×10^{-3}
$4m\%NaCl$	1.3174	1536.933	0.528	1.213×10^{-3}
$6m\%NaCl$	1.3208	1534.197	1.291	2.965×10^{-3}
$8m\%NaCl$	1.3242	1534.010	0.913	2.096×10^{-3}
Sensitivity: 435.454 **nm/RIU**				
$95\%interval$**:** $433.911 < S_{bulk} < 436.997$ **nm/RIU**				

Table C.2: Measurement run 2, λ_{min} as a function of $NaCl$ concentration, values calculated with the Lorentzian curve fitting algorithm

Analyte	RI	$\lambda_{min}[nm]$	$\delta\lambda_{min}[nm]$	DL$[RIU]$
H_2O	1.3105	1563.936	0.384	6.601×10^{-4}
$2m\%NaCl$	1.3139	1561.669	1.391	2.394×10^{-3}
$4m\%NaCl$	1.3174	1559.279	1.200	2.064×10^{-3}
$6m\%NaCl$	1.3208	1557.315	0.203	3.496×10^{-4}
$8m\%NaCl$	1.3277	1553.776	0.276	4.748×10^{-4}
$10m\%NaCl$	1.3242	1553.842	0.186	3.128×10^{-4}
Sensitivity: 580.962 **nm/RIU**				
$95\%interval$**:** $580.162 < S_{bulk} < 581.741$ **nm/RIU**				

Table C.3: Measurement run 3, λ_{min} as a function of $NaCl$ concentration, values calculated with the Lorentzian curve fitting algorithm