

Oppervlakteversterkte Ramanspectroscopie
met behulp van een monomode-nanofotonisch-plasmonisch platform

Surface Enhanced Raman Spectroscopy
Using a Single Mode Nanophotonic-Plasmonic Platform

Frédéric Peyskens

Promotoren: prof. dr. ir. R. Baets, prof. dr. ir. N. Le Thomas
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Fotonica

Vakgroep Informatietechnologie
Voorzitter: prof. dr. ir. D. De Zutter
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2015 - 2016



ISBN 978-90-8578-895-9
NUR 965
Wettelijk depot: D/2016/10.500/27



Universiteit Gent
Faculteit Ingenieurswetenschappen en
Architectuur
Vakgroep Informatietechnologie
Photonics Research Group

Promotoren:

Prof. dr. ir. Roel Baets
Prof. dr. ir. Nicolas Le Thomas

Jury:

Prof. dr. ir. Rik Van de Walle (voorzitter)	Ghent University
Prof. dr. ir. Peter Bienstman (secretaris)	Ghent University
Prof. dr. ir. Roel Baets (promoter)	Ghent University
Prof. dr. ir. Nicolas Le Thomas (promoter)	Ghent University
Prof. dr. ir. Pol Van Dorpe	imec, KU Leuven
dr. Stephane Clemmen	Ghent University
Prof. dr. Stefan Maier	Imperial College London
Prof. dr. Andre Skirtach	Ghent University

Universiteit Gent
Faculteit Ingenieurswetenschappen en Architectuur
Vakgroep Informatietechnologie
Technologiepark-Zwijnaarde 15 iGent, B-9000 Gent, België
Tel.: +32 (0) 9264 3316
Fax.: +32 (0) 9264 3593

Dit werk kwam tot stand in het kader van een persoonlijk BOF (Bijzonder Onderzoeksfonds) doctoraatsmandaat van de Universiteit Gent en van de ERC grant InSpectra.



Proefschrift tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Fotonica
Academiejaar 2015-2016

Dankwoord

Het dankwoord, een kort stukje tekst dat in de meeste boeken slechts sporadisch bekeken of resoluut overgeslagen wordt, krijgt in deze lijvige turf waarschijnlijk de eer om het meest gelezen te worden. Hiermee beweer ik absoluut niet dat de rest van dit werk slechts een lange en saaie epiloog is. Het valt echter niet te ontkennen dat het resultaat van vier jaren hard labeur ook te danken is aan het advies en de steun van heel wat mensen. Zij verdienen dan ook een prominente plaats aan het begin van deze thesis en ik zal proberen om hen in de volgende paragrafen de eer aan te doen die hen toekomt.

Dit doctoraat zou in eerste instantie niet mogelijk geweest zijn zonder de steun en het vertrouwen van mijn promotoren Prof. Roel Baets en Prof. Nicolas Le Thomas. Ik herinner me mijn introgesprek met Roel in maart 2012 nog goed. Dankzij zijn enthousiasme was ik al snel overtuigd van het voorgestelde project en op 1 april 2012 (geen grap) ging ik officieel van start in de Photonics Research Group. Een paar maanden later vervoegde ook Nicolas het Raman-on-chip (ROC) team. Ik wil hen in het bijzonder bedanken voor de ondersteuning, voor de uitgebreide feedback en voor de vrijheid die zij mij gegeven hebben. De zelfstandigheid en kritische geest die ik hierdoor gaandeweg heb aangeleerd zijn ongetwijfeld van groot belang voor het uitbouwen van een verdere (onderzoeks)carrière.

Het overgrote deel van de chipfabricage werd uitgevoerd in samenwerking met imec (Leuven). Ik wil dan ook een aantal mensen bedanken zonder wie het nooit mogelijk was geweest om mijn chips te maken. Eerst en vooral Prof. Pol Van Dorpe voor de vele discussies over plasmonics, de fabricage-tips en uiteraard ook om op te treden als mijn supervisor binnen imec. Josine, bedankt voor al je hulp, tips en geduld bij het uitvoeren van mijn ongebruikelijke e-beam requests. Ook dank aan Pieter voor zijn begeleiding bij de opstartfase en Kim voor alle hulp in de cleanroom. Daarnaast wil ik ook Steven en Liesbet van de UGent-cleanroom bedanken:

Steven voor je allround cleanroom expertise en Liesbet voor de vele prachtige SEM beelden en geduld. Na verloop van tijd werd ik soms de *bowtie-man* genoemd en ik vrees dat jij het grootste slachtoffer was van mijn zoektocht naar het ideale bowtie beeld.

Het voordeel van een grote onderzoeksgroep is dat je gedurende vier jaar heel wat nieuwe mensen leert kennen. Het nadeel is dat het onmogelijk wordt om voor iedereen een afzonderlijk bedankje te schrijven. Toch zou ik graag een aantal personen willen bedanken voor de blijvende indruk die ze gemaakt hebben. Thanks to Ananth for your guidance during the startup of my PhD and all the measurement tips. Ashim, my ROC *partner in crime* from the beginning, you are truly a remarkable guy. Together with Pieter W we made Chemnitz a much more fun place during the SES 2014 conference. Speaking of remarkable guys, a special thanks to “*good morning Sir*” Alfonso and “*dang tiki dang dang*” Sarvagya for the support, numerous chats and dinners at our beloved de Brug. I was lucky (or unlucky ☺) to share an office with both of you and observe your special *bromance*. Merci ook aan Bart K, Pieter G en Raphaël voor de bijzonder interessante verkenning van San Jose. Wij hebben laten zien dat daar nog veel meer te beleven valt dan CLEO alleen. Thanks to the *new guys*, “*tovarish*” Anton, “*dale dale*” Alejandro and “*plansjeeee*” Sören for the nice chats, the dinners and the many nightly explorations of Ghent. Also thanks to all the other *party people* (just too many to list). Ook dank aan Kristien, Mike en Ilse² voor alle ICT en administratieve hulp.

Uiteraard zijn er ook heel wat personen buiten de onderzoeksgroep die een bijzonder dankwoord verdienen. Ook hier is het helaas onmogelijk om iedereen apart in de bloemetjes te zetten. Daarom dat ik hen allemaal wil bedanken om al zo veel jaren klaar te staan met een luisterend oor en goede raad. Ik kan niet genoeg benadrukken hoe dankbaar ik jullie daar voor ben. En dat weten jullie zelf beter dan wie ook, want ik heb het jullie ongetwijfeld al meer gezegd dan ik hier ooit zou kunnen neerschrijven. Toch zou ik graag nog een aantal mensen persoonlijk willen bedanken omdat ze tijdens de afgelopen vier jaar (en vaak al veel langer) prominent aanwezig waren.

Laurenz, Jonas C en Jonas V oftewel “*Sterreman*” Laurie, “*Petit Bateau*” Claeys en “*koning van Zimbabwe*” Plankie, na het middelbaar zijn we elk onze eigen weg gegaan, maar gelukkig is het contact gebleven dankzij de ondertussen legendarische *gezelschapsspellekensavonden*. Ooit begonnen met Risk en inmiddels uitgebreid met strategische hersenbrekers als Havana, Puerto Rico, Imperial 2030, Archipelago, . . . en Camel Up, of

is het nu toch Cup, niet te vergeten. Gedurende de afgelopen 4 jaar vormden al die avondjes een fantastische afwisseling op het doctoraatswerk waar ik altijd naar uitkeek (en dat nog steeds doe voor alle duidelijkheid ☺).

Van studiegenoot naar goede vriend en later zelfs nog even collega. Tom, buiten jou ken ik werkelijk niemand die na wat cocktails of een *Despe* nog steeds zo gefascineerd over fysica kan babbelen. Bedankt voor de vele (al dan niet fysica gerelateerde) babbels tijdens de bijna wekelijkse pizza-avonden in La Rustica. Na zoveel discussies zijn we het er hopelijk over eens dat het juiste antwoord 9π is. Dankzij jou heb ik ook de Polé Polé, ondertussen één van mijn favoriete bars in Gent, leren kennen. Ik ben ervan overtuigd dat de eigenaars jou daar ook bijzonder dankbaar voor zijn. *EDM* Pieter, nog een pizza-regular, bedankt voor de toffe concertavonden en de vele babbels. Ook dank aan alle andere *vriendjes en vriendinnetjes* van de Toegepaste Natuurkunde, Geoff, Jeroen VS, Jeroen V, Céline, Yves, . . .

Amélie, het is nu bijna 10 jaar geleden dat we elkaar leerden kennen op rij 4 van auditorium A in de Plateau. Wie had toen kunnen denken dat we zoveel jaar later beide zouden doctoreren in achtereenvolgens bijna hetzelfde gebouw, hetzelfde gebouw en dan weer bijna hetzelfde gebouw. Bedankt voor al de (ontspannende, grappige, serieuze) babbels tijdens de vele lunches en koffiepauzes (en ook daarbuiten).

Mijn laatste burgiejaar is er eentje dat ik niet snel zal vergeten (de meeste momenten toch ☺) en dat is mede dankzij Bart, met wie ik de beste VTK show ooit heb mogen regisseren. Bart, je kan als geen ander relativeren (wat is 100 euro nu?) en ik ben blij dat we na onze wilde studentenjaren contact gehouden hebben.

Men zou het een paar jaar geleden nooit gedacht hebben, maar samen met Eva heb ik me er toch aan gewaagd om de eerste stappen (1,2,3-5,6,7-...) in het salsawereldje te zetten. Achteraf bekeken een zeer welkome ontspanning tijdens het schrijven van een doctoraat. Ik kan het iedereen alleen maar aanraden, de salsa dan toch ☺. . .

Oma, jij hebt mij *ook een beetje grootgebracht*. Ik kan niet bedenken hoeveel woensdagnamiddagen ik bij jou ben geweest en ben je dan ook enorm dankbaar voor al de goede zorgen en je interesse. Merci. Valérie, onze gekruiste Belgisch-Spaanse mus, bedankt voor de leuke herinneringen aan de vele zomers die we samen hebben doorgebracht op Perreveld en voor de vele babbels. Ook dank aan mijn andere nichten en neef voor de bijzondere reünie vorige zomer en aan alle familieleden die me steeds gesteund hebben.

Tot slot wil ik ook mijn ouders een zeer oprechte dankjewel zeggen.

Bedankt voor jullie onvoorwaardelijke steun, geduld en interesse. En dat niet alleen gedurende de afgelopen vier jaar. Jullie hebben mij altijd volop gesteund in mijn keuzes, ook al waren die niet steeds even voor de hand liggend. Ik wil jullie ook bedanken voor de prachtige reizen, die me ongetwijfeld een veel ruimere kijk hebben gegeven en waaraan ik een ontelbaar aantal mooie herinneringen heb overgehouden. Ik zou hier nog heel lang kunnen doorgaan, want het is nagenoeg onmogelijk te beschrijven wat jullie allemaal voor mij gedaan hebben en hoeveel jullie voor mij betekenen. Dit werk zou er nooit geweest zijn zonder jullie. Kortom, ongelooflijk bedankt voor alles . . .

Oudenaarde, 5 Mei 2016
Frédéric Peyskens

Table of Contents

Dankwoord	i
Nederlandse samenvatting	xi
English summary	xvii
1 Introduction	1-1
1.1 Introduction	1-1
1.2 Outline	1-3
1.3 Publications	1-5
1.3.1 Publications in international journals	1-5
1.3.2 Publications in international conferences	1-6
References	1-8
2 Theory of Plasmon Optics	2-1
2.1 Introduction	2-1
2.2 Maxwell revisited	2-1
2.3 Surface Plasmon Polaritons	2-3
2.4 Localized Surface Plasmons	2-8
2.4.1 Optical excitation of plasmon modes	2-12
2.4.2 Field enhancement	2-14
2.4.3 Coupled nanoplasmonic particles	2-16
2.5 Conclusion	2-19
References	2-19
3 Raman Spectroscopy	3-1
3.1 Introduction	3-1
3.2 Infrared absorption spectroscopy	3-2
3.3 Fluorescence spectroscopy	3-5
3.4 Spontaneous Raman spectroscopy	3-8
3.5 Surface Enhanced Raman Spectroscopy	3-12

3.5.1	Qualitative treatment of SERS	3-13
3.5.2	Size and shape effects on field enhancement	3-14
3.5.3	Fluorescence quenching versus Raman enhancement	3-18
3.6	Conventional Raman spectroscopy setup	3-27
3.7	Conclusion	3-31
	References	3-31
4	On-Chip Raman Spectroscopy	4-1
4.1	Introduction	4-1
4.2	Optical Waveguides	4-2
4.3	Silicon nitride photonics	4-6
4.4	Raman radiation near a waveguide	4-9
4.4.1	Dipole emission near a waveguide	4-10
4.4.2	Spontaneous on-chip Raman scattering	4-14
4.5	Evanescence excitation and collection of Raman spectra	4-18
4.5.1	Measurement setup	4-18
4.5.2	On-chip bulk Raman sensing	4-21
4.6	Conclusion	4-23
	References	4-25
5	Fabrication and Characterization of Nanophotonic-Plasmonic Chips	5-1
5.1	Introduction	5-1
5.2	E-beam lithography	5-2
5.2.1	Forward and backward scattering of an electron beam	5-2
5.2.2	Electron Dose	5-4
5.2.3	Shot pitch	5-5
5.2.4	Field stitching	5-5
5.3	Scheme for nanoplasmonic-photonic patterning	5-6
5.3.1	Patterning plasmonic antennas	5-8
5.3.1.1	FOx-12	5-8
5.3.1.2	PMMA	5-12
5.3.1.3	ma-N 2400.6	5-14
5.3.2	Patterning SiN waveguides	5-16
5.4	Optical characterization	5-17
5.4.1	Waveguide loss	5-17
5.4.2	Plasmon resonance measurement	5-23
5.5	Conclusion	5-25
	References	5-27

6	Plasmon Resonances of Evanescently Coupled Nanoplasmonic Antennas	6-1
6.1	Introduction	6-1
6.2	Hybrid nanophotonic-plasmonic chips	6-2
6.3	Simulation model	6-4
6.4	Bright mode resonance tuning	6-6
6.4.1	Comparison between theory and experiment	6-10
6.4.2	Q -factor of integrated nanoantennas	6-10
6.4.3	Interaction efficiency	6-12
6.4.4	Raman enhancement	6-17
6.5	Excitation of dark plasmon modes	6-18
6.6	Conclusion	6-23
	References	6-23
7	Integrated Surface Enhanced Raman Spectroscopy	7-1
7.1	Introduction	7-1
7.2	On-chip SERS platform	7-2
7.3	Influence of the plasmon resonance	7-6
7.4	Influence of the number of antennas	7-7
7.5	Analytical on-chip SERS model	7-9
7.5.1	SERS signal	7-10
7.5.1.1	Forward scattering	7-12
7.5.1.2	Backward scattering	7-16
7.5.2	Relation to the measurable signal	7-18
7.5.3	Background signal	7-19
7.5.4	Signal-to-Noise Ratio	7-21
7.6	Generalized on-chip SERS model	7-23
7.7	Comparison between theory and experiment	7-26
7.7.1	Simulation setup	7-26
7.7.2	Fit of the experimental data	7-30
7.7.3	Comparison of FOM_F	7-32
7.8	ROC versus SEROC	7-33
7.9	Comparison with other SERS substrates	7-37
7.10	Conclusion	7-40
	References	7-41
8	Background Mitigation	8-1
8.1	Introduction	8-1
8.2	Y-splitter design	8-1
8.2.1	Signal	8-4

8.2.2	Background	8-6
8.2.3	Signal-to-Noise Ratio	8-13
8.3	Preliminary experimental results	8-13
8.4	Conclusion	8-15
	References	8-16
9	Conclusion and Outlook	9-1
9.1	Conclusion	9-1
9.2	Outlook	9-4
9.2.1	Background reduction	9-4
9.2.2	Non-Au SERS	9-4
9.2.3	Nanoparticle trapping	9-4
9.2.4	Quantum Photonics	9-5
	References	9-5

Nederlandse samenvatting –Summary in Dutch–

Gedurende de afgelopen jaren is er een toenemende vraag naar draagbare en compacte spectroscopische toestellen voor het uitvoeren van accurate, gevoelige en in situ detectie van een grote verscheidenheid aan biologische en chemische substanties. Ramanspectroscopie neemt een bijzondere plaats in omdat men via deze techniek de fundamentele vibraties van een molecule kan detecteren. Aangezien deze vibraties karakteristiek zijn voor elke molecule (een soort moleculaire vingerafdruk), vormt het dus een uitgelezen methode om specifieke substanties te identificeren aan de hand van hun karakteristieke spectrum. Alle traditionele spectroscopische technieken (dus niet alleen Ramanspectroscopie) vereisen echter vaak het gebruik van dure en omvangrijke instrumentatie. Als gevolg daarvan is het moeilijk om dergelijke methodes op grote schaal toe te passen buiten een typische laboratorium omgeving. Hoge index contrast fotonisch geïntegreerde circuits (ICs) kunnen een oplossing bieden voor dit probleem. Eerst en vooral hebben hoge index contrast fotonische ICs zeer gelokaliseerde modes, waardoor de sterkte van het evanescente veld zeer groot is voor een gegeven optisch vermogen. Dit resulteert in een intense interactie tussen het fotonische circuit en zijn omgeving en is van groot belang voor on-chip sensing toepassingen. Bovendien kunnen verschillende optische functionaliteiten geïntegreerd worden op één enkele fotonische chip. Daarnaast leidt integratie tot goedkopere en betrouwbare toestellen die massa-reproduceerbaar zijn. Tot slot biedt integratie de mogelijkheid tot multiplexing.

Het meest prominente fotonische platform is het SOI platform, gebaseerd op silicium golfgeleiders op een silica substraat. Dergelijke golfgeleiders zijn transparant tussen 1.1 en 4 μm . Het therapeutische golflengtevenster voor de meeste biologische media is echter tussen 750 en 930 nm omdat er binnen dit bereik minimale licht-geïnduceerde schade en waterabsorptie is. Voor biologische toepassingen is het dus noodzakelijk

om het SOI platform uit te breiden naar kortere golflengtes en tegelijkertijd de mogelijkheid behouden om gebruik te maken van standaard CMOS technologie. Silicium nitride (SiN) is waarschijnlijk het meest voor de hand liggende materiaal omdat het transparant is tussen 0.4 en 5 μm , wat het bijzonder geschikt maakt voor het zichtbare en nabije infrarood (IR) spectrum. Binnen dit golflengtebereik is er ook een grote verscheidenheid aan performante en goedkope lichtbronnen en detectors.

In de context van de ERC Advanced Grant InSpectra is onze groep gestart met het onderzoek naar geïntegreerde Raman spectroscopie systemen, gebruik makend van het SiN platform. We hebben voor de eerste maal experimenteel aangetoond dat Raman spectra geëxciteerd en geëxciteerd kunnen worden door het evanescente veld van een monomode SiN golfgeleider. Bovendien hebben we ook vastgesteld dat zo'n geïntegreerd Raman platform een betere pomp-naar-Raman conversie-efficiëntie heeft in vergelijking met een conventionele Raman microscoop. Nochtans kan een dergelijk platform niet alle nadelen van conventionele systemen omzeilen. Zo blijft het moeilijk om zeer kleine deeltjesaantallen of nanopartikels te detecteren. De plasmonresonanties van specifiek ontworpen metallische nanostructuren (nanoplasmonische antennes) kunnen een oplossing bieden voor dit probleem. Het elektrisch veld rond zo'n resonante structuur kan enorm versterkt worden omdat het samengeperst wordt nabij het metaaloppervlak. Als gevolg daarvan wordt de pomp (excitatie) intensiteit versterkt. Bovendien wordt ook het uitgezonden Raman licht versterkt door hetzelfde mechanisme. Het uiteindelijke resultaat is een aanzienlijke versterking van het oorspronkelijke Raman signaal waardoor extreem zwakke Raman emitters, geadsorbeerd op het metaaloppervlak, gedetecteerd kunnen worden. Dit proces wordt Surface Enhanced Raman Spectroscopy (SERS) genoemd.

De meeste SERS toepassingen maken echter nog steeds gebruik van conventionele systemen, gebaseerd op excitatie en collectie met behulp van een microscoop. Ondanks het feit dat men recent ook gestart is met het integreren van nanoplasmonische antennes op fotonische ICs, zijn dergelijke ICs tot op heden alleen gebruikt om het SERS signaal van externe, niet-geïntegreerde, metallische nanopartikels te bestuderen. Een dergelijke aanpak is echter niet kwantitatief als gevolg van de grote onzekerheid op de precieze Raman versterking en de koppeling tussen het pomp licht en de metallische nanodeeltjes. Om kwantitatieve SERS spectra te bekomen, is een complete controle over de Raman versterking en de koppeling met de onderliggende golfgeleider echter noodzakelijk. Dit uitdagend probleem

vormt het centrale doel van dit doctoraatsonderzoek. De verschillende aspecten die onderzocht werden, zijn hieronder samengevat:

- Verschillende fabricagemethodes voor de hybride integratie van nanoplasmonische antennes en fotonische ICs.
- Invloed van de geometrische antenneparameters op de plasmonresonantie.
- Ontwikkeling van een analytisch on-chip SERS model om de relevante design parameters te achterhalen.
- Experimentele karakterisatie van on-chip SERS signalen, geëxciteerd en gecollecteerd door de fundamentele golfgeleider mode.
- Strategieën om de signaal-ruis verhouding te verbeteren.

Zo'n geïntegreerde monomode SERS probe kan gecombineerd worden met andere fotonische, microfluidische of biologische functionaliteiten om een volledig geïntegreerd platform te ontwikkelen voor de detectie van extreem zwakke Raman signalen.

De typische resolutie die nodig is om nanoplasmonische antennes te definiëren, kan normaal niet bereikt worden met standaard deep-UV lithografie. Electron beam lithografie (EBL) is wel in staat om de vereiste resolutie te bereiken, maar is typisch veel trager. Omdat de resist waarin de antennes gedefinieerd worden veel dunner is dan de dikte van een typische golfgeleider, is het aangeraden om eerst de antennes te fabriceren op een vlak substraat en pas nadien de golfgeleiders rond de antennes te etsen. Dit verhoogt de uniformiteit en leidt normaal ook tot een betere definitie van de antennes. Een dergelijk proces vereist wel 2 EBL stappen. In de eerste stap worden de antennes gedefinieerd en verschillende EBL resists werden onderzocht om de optimale proces-strategie te bepalen. Er werd vastgesteld dat een combinatie van ma-N 2400.6 en ma-N 2403 resist de beste resultaten geeft in termen van reproduceerbaarheid, patroondefinitie en verliezen in de chip. De beste golfgeleiderverliezen die met dit geoptimaliseerd proces bekomen werden, zijn 3 dB/cm rond $\lambda = 785$ nm.

Vervolgens werd de invloed van de geometrische antenneparameters op de plasmonresonantie onderzocht in het golflengtebereik 700–1000 nm. In onze studie hebben we 3 verschillende antenne types (single rod, double rod en bowtie antenne) theoretisch en experimenteel vergeleken. De theoretisch voorspelde trends werden allemaal bevestigd door onze experimenten en interactie-efficiënties tot 19% rond 800 nm werden waargenomen voor

een bowtie antenne. We toonden theoretisch ook aan dat deze interactie-efficiëntie geoptimaliseerd kan worden door de golfgeleidergeometrie aan te passen. Uit numerieke simulaties werd vastgesteld dat de bowtie antenne de hoogste Raman versterking heeft van de drie onderzochte antennes. Bovendien bezitten bowtie antennes met een grote apex hoek naast een brede dipool resonantie ook een smalle (hogere Q -factor) quadrupool resonantie.

Na deze plasmonresonantie-studie hebben we onze aandacht gericht op het hoofddoel van dit doctoraatswerk, namelijk de excitatie en collectie van SERS signalen via een hybride nanofotonisch-plasmonisch platform. Om de pure SERS eigenschappen van een dergelijk hybride platform te bestuderen, is het belangrijk om het SERS signaal te isoleren en elke andere contributie van spontaan Raman licht te vermijden. Daartoe werd 4-nitrothiophenol (NTP) gebruikt. NTP vormt een monolaag op het goudoppervlak van de antennes door middel van een Au-S binding, zonder daarbij te binden op het SiN. Metingen bevestigen inderdaad dat er geen NTP signaal gegenereerd wordt door een referentie golfgeleider zonder antennes. Daarnaast toonden we aan dat het SERS signaal toegeschreven kan worden aan een plasmonresonantie effect zodat er een stabiele en reproduceerbare Raman versterkingsfactor mee geassocieerd kan worden. Bovendien werd vastgesteld dat het SERS signaal gemaximaliseerd kan worden door de golfgeleider met een optimaal aantal antennes te functionaliseren. Een essentieel aspect binnen onze studie bestond uit het ontwikkelen van een on-chip SERS model om de relevante design parameters te achterhalen. Ons eerste (ideaal) model hield echter geen rekening met variaties in de antennegeometrie als gevolg van fabricagefouten. Om dergelijke variaties in rekening te brengen werd een veralgemeend model ontwikkeld. Op basis van een random fit aan dit veralgemeend model is het mogelijk om een schatting te maken van de variatie op de antenneparameters (die op hun beurt het gevolg zijn van variaties in de antennegeometrie). Vervolgens werden deze experimenteel gefitte data vergeleken met de theoretische voorspellingen. We vonden een excellente correspondentie tussen het theoretisch voorspelde en experimenteel bepaalde Raman vermogen dat terug in de golfgeleider koppelt. Het ontwikkelde hybride nanofotonisch-plasmonisch platform wordt dus gekenmerkt door een complete kwantitatieve controle op de Raman versterking en koppeling van het versterkte Raman licht met de onderliggende golfgeleider.

Daarenboven hebben we ook een trade-off vastgesteld tussen signaal optimalisatie en background reductie. Het SiN genereert zelf een aanzien-

lijke Raman background bovenop het NTP Raman signaal, wat aanleiding geeft tot zogenaamde shot noise. Shot noise kan de kleinste SERS signalen maskeren en moet dus vermeden worden om de absolute detectielimiet te verbeteren. Strategieën die tegelijkertijd het signaal optimaliseren en de shot noise minimaliseren zijn in dat opzicht dus heel interessant. Om dit te realiseren hebben we een specifieke Y-splitter configuratie onderzocht waarbij de korte rechte arm van de splitter gefunctionaliseerd is met antennes, terwijl de twee gebogen armen gebruikt worden voor respectievelijk excitatie en collectie. Dit leidt tot een efficiënte scheiding van het SERS signaal en de SiN background. De eerste experimentele resultaten tonen aan dat de background voor zo'n Y-splitter designs al met een factor 100 kan verkleind worden.

English summary

Over the past years, there has been an increasing demand for hand-held devices capable of accurate, sensitive, and in situ spectroscopic detection of a variety of substances. These devices span a large range of potential applications in physics, chemistry, biology and environmental science. Raman spectroscopy is a particularly interesting method as it probes the fundamental vibrations of a molecule (which act as a molecular fingerprint). In that respect it is widely employed to identify specific substances from the characteristic spectrum. However, a common disadvantage for all traditional spectroscopic systems (so not only Raman systems) is the use of expensive and bulky instrumentation, which prohibits their dissemination in non-laboratory environments. The use of high index contrast photonic integrated circuits (PICs) can provide a potential solution to this bottleneck. First of all, high index contrast PICs have tightly confined guided modes, allowing a very strong electric field strength of the evanescent tail for a given optical power. This leads to an intense interaction between the PIC and its surroundings, which is of paramount importance for on-chip evanescent sensing. Moreover, PICs enable miniaturization and integration of several optical components on a single chip. Integration also leads to devices which are reliable, mass producible and cheap. Finally, integration also allows for the possibility of multiplexing.

The most prominent PICs are still based on the Silicon-on-Insulator (SOI) platform. Such SOI waveguides are transparent in the 1.1 to 4 μm wavelength range. However, the therapeutic wavelength window for biological media, which induces minimal photodamage to cells and has negligible water absorption, is 750 – 930 nm. For biological (sensing) applications it is hence required to extend silicon photonics to shorter wavelengths, while maintaining the ability to make use of CMOS tools at these shorter wavelengths. To this end, silicon nitride (SiN) is probably the most convenient material as it is transparent from 0.4 to 5 μm , which makes it particularly important for the visible and very near-IR wavelengths (< 1 μm). In this wavelength region there is also a plethora of high-performance

and low-cost light sources and detectors.

In the context of the ERC Advanced Grant InSpectra, our group embarked upon the challenging task of realizing on-chip Raman spectroscopy using a SiN platform. We reported on the first experimental demonstration of on-chip Raman sensing using the evanescent field of a single mode SiN photonic waveguide and showed that such an on-chip Raman platform offers a clear advantage over conventional confocal Raman spectroscopy in terms of conversion efficiency from pump to Stokes power. Nevertheless, this on-chip Raman platform still suffers from the same drawback of conventional Raman spectroscopy in the sense that spontaneous Raman scattering is usually extremely weak. As a result, it becomes increasingly difficult to detect very small amounts of analyte or single nanoparticles (or molecules) through spontaneous Raman spectroscopy. Specifically designed metallic nanostructures can exhibit a so-called plasmon resonance, associated to a subwavelength confinement of the electric field, which results in a considerable enhancement of the electric field near the metal surface. This field enhancement is beneficial for the excitation and collection of Raman signals and nanoplasmonic antennas can locally boost the Raman signal in order to allow detection of small amounts of analyte adsorbed to the metal surface. This process is termed Surface Enhanced Raman Spectroscopy (SERS).

Most SERS applications still rely on off-chip systems and while researchers also started integrating nanoplasmonic antennas on top of waveguides, PICs have only been used to probe SERS signals from external, non-integrated, metallic nanoparticles. Such an approach is however poor in terms of quantitative results owing to the large uncertainty on the absolute Raman enhancement and coupling between the excitation beam and the metallic nanoparticles. A complete control of the plasmonic enhancement and coupling with the underlying waveguide is however necessary in order to obtain quantitative SERS spectra. This challenging problem forms the central goal of this PhD work and to resolve it, we have investigated several aspects of hybrid nanophotonic-plasmonic chips:

- Different fabrication schemes for the hybrid integration of nanoplasmonic antennas on SiN waveguides.
- Plasmon resonance tuning as a function of the geometrical antenna parameters.
- Development of an analytical on-chip SERS model in order to outline the relevant design parameters and figure of merit.

- Experimental characterization and demonstration of on-chip SERS signals, excited and collected by the same fundamental waveguide mode.
- Strategies for background mitigation in order to improve the Signal-to-Noise Ratio.

These on-chip single mode SERS probes can be combined with other photonic, fluidic or biological functionalities in order to develop a highly dense integrated platform for the multiplexed detection of extremely weak Raman signals.

The required resolution to pattern nanoplasmonic antennas typically exceeds the one that can be achieved through standard deep-UV lithography. Despite the relatively low throughput of e-beam lithography (EBL), it is nevertheless necessary because EBL can achieve the required resolution. Patterning of nanoplasmonic antennas generally requires an e-beam resist which is much thinner than the thickness of the waveguides on which they have to be integrated. For an improved definition of these structures it is therefore desirable to fabricate the antennas first, and only afterwards etch a waveguide around them. This is realized by a two-step e-beam lithography process starting from a blank SiN slab wafer. The antennas are defined in the first step. Different processes were considered based on the type of e-beam resist. We investigated the pros and cons of each strategy and eventually concluded that a process using the negative ma-N 2400.6 resist is the optimal choice because it allows the definition of smooth and reproducible structures, while at the same time not having a detrimental impact on the eventual waveguide performance. The waveguides are defined in the second step using ma-N 2403 and the best waveguide losses obtained through this combined process were about 3 dB/cm around $\lambda = 785$ nm.

Subsequently, we investigated the plasmon resonance tuning in the 700 – 1000 nm wavelength region, as a function of the geometrical antenna parameters. In our study we theoretically and experimentally compared single rod, double rod and bowtie antennas. The theoretically predicted resonance trends were confirmed by our experiments and experimental interaction efficiencies as high as 19% around 800 nm were observed for bowtie antennas. We also verified that this interaction efficiency can be optimized by tuning the waveguide geometry. Moreover, numerical simulations of the field enhancement were performed and it was shown that of the three investigated antenna types, the bowtie antenna has the highest

maximum Raman enhancement. Apart from the broad dipolar resonance we established that bowtie antennas with a large apex angle also exhibit a narrower (higher Q) dark resonance.

After our study of the plasmon resonance tuning of integrated nanoplasmonic antennas in the 700 – 1000 nm region, we turned our attention to the investigation of on-chip SERS, which was the key challenge of this PhD. In order to study the pure SERS properties of the platform, it is important to isolate the SERS signal and avoid any other contribution from waveguide generated spontaneous Raman scattering. To this end, 4-nitrothiophenol (NTP) was used as it forms a self-assembled monolayer on top of the gold through a Au-S bond without binding to the SiN waveguide. Measurements on a non-functionalized reference waveguide (without any gold) indeed confirmed that no NTP remained on the SiN after cleaning. We also showed that the SERS spectra could be attributed to a plasmon resonance effect, meaning that a stable and reproducible enhancement factor can be associated with it. Moreover we found that an optimum antenna number exists such that the SERS signal is maximized. An essential aspect within our study was the development of an analytical on-chip SERS model in order to establish the relevant design parameters and figure of merit. Fabrication induced variations among each of the antennas will however result in differences between the antenna parameters. In order to incorporate these potential differences, we have extended our ideal on-chip SERS model to a generalized model. Based on a randomized fit to this generalized model, we assessed the experimental spread on the design parameters and subsequently compared the experimental data with the theoretical predictions. An excellent correspondence was found between the theoretically predicted and experimentally observed absolute Raman power that is coupled back into the waveguide mode. As a result, the suggested platform enables a complete quantitative control on the Raman enhancement and subsequent coupling of the enhanced Stokes light with the underlying waveguide.

Furthermore we identified a trade-off between signal optimization and background reduction. The SiN itself generates a considerable Raman background signal on top of the Raman signal of the analyte, resulting in unwanted shot noise. This background associated shot noise will mask the smallest signal features and has to be mitigated in order to improve on the absolute detection limit. Strategies which simultaneously optimize the signal and mitigate the noise are therefore very interesting. We investigated a dedicated Y-splitter where an array of antennas is patterned on the short

straight arm, while the two bend arms are used for excitation and collection respectively. This results in a more efficient separation of the SERS signal and the SiN background. Preliminary results already show that the background can be reduced by a factor of 100 for such Y-splitter designs.

1

Introduction

1.1 Introduction

The main optical spectroscopy techniques which are used to detect molecular vibrations are based on infrared absorption and Raman spectroscopy. They are widely employed to provide chemical and physical information, to identify substances from the characteristic spectral patterns (so-called fingerprinting) and to quantitatively determine the amount of a particular substance in a sample. [1] The attraction of Raman spectroscopy for chemical analysis comes from a combination of advantages. First, Raman spectra can be acquired noninvasively. Moreover, Raman scattering probes fundamental vibrations with high spectral resolution. In addition, Raman has some added features such as the compatibility with aqueous samples. [2]

A common disadvantage for all traditional spectroscopic systems (so not only Raman systems) is the use of expensive and bulky instrumentation, which prohibits their dissemination in non-laboratory environments. In recent years, there has however been a large demand for hand-held devices that are capable of accurate, sensitive, and in situ spectroscopic detection of a variety of substances. The use of high index contrast (HIC) photonic integrated circuits (PICs) can provide a potential solution to this bottleneck. First of all, high index contrast PICs have tightly confined guided modes,

allowing a very strong electric field strength of the evanescent tail for a given optical power. This leads to an intense interaction between the PIC and its surroundings, which is of paramount importance for on-chip evanescent sensing. Moreover, PICs enable miniaturization and integration of several optical components on a single chip. Integration also leads to devices which are reliable, mass producible and cheap. Finally, integration also allows for the possibility of multiplexing. [3]

The most prominent PICs are realized on an SOI (Silicon-On-Insulator) waveguide platform. However, SOI waveguides are only relevant in the 1.1 to 4 μm wavelength range due to the absorption by Si. In order to extend silicon photonics to shorter wavelengths (below 1.1 μm), while maintaining the ability to make use of CMOS technology, it is probably most convenient to use silicon nitride (SiN). SiN is a HIC material and is transparent from 0.4 to 5 μm , which makes it particularly important for the visible and very near-IR wavelengths ($< 1 \mu m$). This wavelength region covers the therapeutic window (750 – 930 nm) for biological media because these wavelengths induce minimal photodamage to cells and have negligible water absorption. Moreover, there is a large availability of low-cost and high-performance light sources and silicon-based detectors in this wavelength range. [3]

Recently, our group established that PICs offer a promising alternative to standard confocal microscopy to probe spontaneous Raman spectra. [4] Nevertheless, such an on-chip Raman platform suffers from the same drawback of conventional Raman spectroscopy in the sense that spontaneous Raman scattering is usually extremely weak. As a result, it becomes increasingly difficult to detect very small amounts of analyte or single nanoparticles (or molecules) through spontaneous Raman spectroscopy. Nanoplasmonic antennas are specifically designed metallic nanostructures which allow a considerable enhancement of the electric field near the metal surface. Hence, such antennas can locally boost the signal in order to allow detection of small amounts of analyte adsorbed to the metal surface.

While researchers also started integrating nanoplasmonic antennas on top of such waveguides [5–12], PICs have only been used to probe SERS signals from external, non-integrated, metallic nanoparticles [13–15]. Such an approach is however poor in terms of quantitative results owing to the large uncertainty on the Raman enhancement and coupling between the excitation beam and the metallic nanoparticles. In order to obtain quantitative SERS spectra, a complete control of the plasmonic

enhancement and coupling with the underlying waveguide is necessary. This challenging problem forms the central goal of this PhD work. In order to resolve it, several aspects have been investigated:

- Different fabrication schemes for the hybrid integration of nanoplasmonic antennas on SiN waveguides.
- Plasmon resonance tuning as a function of the geometrical antenna parameters.
- Development of an analytical on-chip SERS model in order to outline the relevant design parameters and figure of merit.
- Experimental characterization and demonstration of on-chip SERS signals excited and collected by the same fundamental waveguide mode.
- Strategies for background mitigation in order to improve the Signal-to-Noise Ratio of the Raman sensor.

A brief overview of the different chapters is given in the next section.

1.2 Outline

In Chapter 2 we give a theoretical overview of plasmon optics, with a strong focus on localized surface plasmons. A coupled mode theory model is used to explain the field enhancement mechanism of isolated and coupled metallic nanoparticles. Moreover, the optical excitation efficiency of plasmon modes is explained through the associated dipole moment of the mode.

Chapter 3 provides an introduction to IR absorption, fluorescence and Raman spectroscopy and their respective pros and cons. Since the focus of our work is Raman spectroscopy, we provided a more thorough overview of this particular technique. At first, spontaneous Raman spectroscopy is discussed. Secondly, a qualitative treatment of Surface Enhanced Raman Spectroscopy (SERS) is given in order to introduce the rationale behind it. The coupled mode theory as outlined in Chapter 2 is then used to quantitatively describe the simultaneous effect of fluorescence quenching and Raman enhancement. Finally, we discuss a conventional Raman spectroscopy setup.

While Chapter 3 focused on free space spectroscopy, Chapter 4 discusses an on-chip spectroscopy platform. First, some general theory on optical waveguides is discussed, after which the silicon nitride photonics platform is introduced. Subsequently, a theoretical treatment of dipole emission near a waveguide is outlined. This theory is then applied to derive the amount of Raman scattering, generated by an analyte on top of the waveguide, that couples into the fundamental waveguide mode. Moreover, a detailed description of the measurement setup for on-chip Raman spectroscopy and the corresponding alignment procedures is given.

Having discussed the spontaneous on-chip Raman case, we turn our focus to the main topic of this PhD in Chapters 5 to 8. In Chapter 5, different fabrication strategies for the hybrid integration of nanoplasmonic antennas and SiN waveguides are discussed. Secondly, the optical characterization of the eventual devices is described as well. The impact of the metal ion milling on the waveguide loss is discussed. Moreover, the measurement setup for characterizing on-chip plasmon resonances is described as well.

Chapter 6 contains an overview of the simulated antenna geometries and their resonance tuning properties. These simulated trends are subsequently compared with the experimental data. Furthermore, the interaction efficiency of these antennas with the evanescent waveguide mode is studied theoretically and experimentally. Finally, the excitation of dark plasmon modes is discussed.

In Chapter 7 we turn our attention to the main result of this PhD work, i.e. the first experimental demonstration of SERS signals generated from integrated bowtie nanoantennas, excited and collected by a single mode silicon nitride waveguide. The influence of the plasmon resonance and the number of antennas on the SERS signal and the inherent Raman background is initially discussed. Subsequently, we developed an analytical on-chip SERS model in order to identify the relevant design parameters and figure of merit. This model also illustrates the interplay between signal optimization and noise reduction. A formula for the SERS Signal-to-Noise Ratio SNR is derived as well. It is found that a minimum number of antennas is required to obtain an $SNR > 1$ in forward collection. Since the actual fabricated antennas will show differences among each other, we have extended our ideal on-chip SERS model to a more generalized model which incorporates potential changes among all antennas on the same chip. Based on a randomized fit procedure to this generalized model, we were able to estimate the experimental spread on the design parameters

and subsequently compare them with the theoretical predictions. An excellent correspondence was found between the theoretically estimated and experimentally obtained absolute Raman powers. As a result, the developed single mode SERS probe allows a complete quantitative control of the Raman enhancement and subsequent coupling of the SERS signal to the underlying waveguide.

While the development of such a quantitative single mode SERS probe was the key challenge of this PhD work, we have identified potential improvements to the current platform as well. A major issue in all on-chip Raman experiments is the presence of the inherent SiN background. This background sets a constraint to the absolute detection limit and has to be mitigated as much as possible. Chapter 8 deals with our ongoing work on background mitigation in order to improve the *SNR* and hence pave the way towards highly dense and sensitive on-chip SERS probes. Concluding remarks and a future outlook are briefly discussed in Chapter 9.

1.3 Publications

1.3.1 Publications in international journals

1. Peyskens, F.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform. *ACS Photonics* **2016**, *3*(1), 102–108.
2. Dhakal, A.; Raza, A.; Peyskens, F.; Subramanian, A.Z.; Clemmen, S.; Le Thomas, N.; Baets, R. Efficiency of evanescent excitation and collection of spontaneous Raman scattering near high index contrast channel waveguides. *Opt. Express* **2015**, *23*(21), 27391–27404.
3. Subramanian, A.Z.; Ryckeboer, E.M.P.; Dhakal, A.; Peyskens, F.; Malik, A.; Kuyken, B.; Zhao, H.; Pathak, S.; Ruocco, A.; De Groote, A.; Wuytens, P.C.; Martens, D.; Leo, F.; Xie, W.; Dave, U.D.; Muneeb, M.; Van Dorpe, P.; Van Campenhout, J.; Bogaerts, W.; Bienstman, P.; Le Thomas, N.; Van Thourhout, D.; Hens, Z.; Roelkens, G.; Baets, R. Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip (Invited). *Photon. Res.* **2015**, *3*(5), B47–B59.

4. Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Bright and dark plasmon resonances of nanoplasmonic antennas evanescently coupled with a silicon nitride waveguide. *Opt. Express* **2015**, *23*, 3088–3101.
5. Dhakal, A.; Subramanian, A.Z.; Wuytens, P.; Peyskens, F.; Le Thomas, N.; Baets, R. Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides. *Opt. Lett.* **2014**, *39*, 4025–4028.
6. Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Jansen, R.; Claes, T.; Rottenberg, X.; Peyskens, F.; Selvaraja, S.; Helin, P.; Du Bois, B.; Leyssens, K.; Severi, S.; Deshpande, P.; Baets, R.; Van Dorpe, P. Low-Loss Singlemode PECVD Silicon Nitride Photonic Wire Waveguides for 532900 nm Wavelength Window Fabricated Within a CMOS Pilot Line. *IEEE Photon. J.* **2013**, *5*(6), 2202809.

1.3.2 Publications in international conferences

1. Peyskens, F.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R., In *Hybrid Single Mode Nanophotonic-Plasmonic Waveguides for On-Chip Surface Enhanced Raman Spectroscopy*, accepted for publication in META 2016, Spain.
2. Peyskens, F.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R., In *Surface Enhanced Raman Spectroscopy on Single Mode Nanophotonic-Plasmonic Waveguides*, accepted for publication in CLEO 2016, United States.
3. Dhakal, A.; Raza, A.; Wuytens, P.C.; Peyskens, F.; Skirtach, A.; Baets, R. In *Lab-on-a-chip Raman sensors outperforming Raman microscopes*, accepted for publication in CLEO 2016, United States.
4. Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.Z.; Skirtach, A.; Le Thomas, N.; Baets, R. In *Nanophotonic Lab-On-A-Chip Raman Sensors: a Sensitivity Comparison with Confocal Raman Microscope*, Proceedings of BioPhotonics 2015 Conference, Florence, Italy, May 20–22, 2015.
5. Clemmen, S.; Zhao, H.; Peyskens, F.; Dhakal, A.; Wuytens, P.C.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Coherent*

- anti-Stokes Raman spectroscopy on chip*, 28th IEEE Photonics Conference (IPC 2015), United States, 2015; pp. 623-624.
6. Dhakal, A.; Peyskens, F.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Enhanced Spontaneous Raman Signal Collected Evanescently by Silicon Nitride Slot Waveguides*, CLEO: 2015 - Laser Science to Photonic Applications, United States, 2015; pp. paper STh4H.3.
 7. Baets, R.; Subramanian, A.Z.; Dhakal, A.; Peyskens, F.; Wuytens, P.C.; Ryckeboer, E.M.P.; Roelkens, G.; Le Thomas, N. In *Spectroscopic sensing enabled by silicon photonics*, Asia Communications and Photonics Conference (ACP), China, 2014.
 8. Baets, R.; Dhakal, A.; Peyskens, F.; Wuytens, P.C.; Skirtach, A.; Le Thomas, N.; Subramanian, A.Z. In *Resonant enhancement mechanisms in lab-on-chip raman spectroscopy on a silicon nitride waveguide platform*, IEEE Photonics Conference 2014 (IPC), United States, 2014; pp. 500-501.
 9. Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Evanescent Raman spectroscopy using photonic waveguides*, International Conference on Raman Spectroscopy-ICORS 2014, Germany, 2014; pp. WeA-O-006.
 10. Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Wuytens, P.C.; Van Dorpe, P.; Le Thomas, N.; Baets, R. In *Enhancement of Raman scattering efficiency by on-chip nanoplasmonic antennas*, Surface-Enhanced Spectroscopies 2014 conference, Germany, 2014.
 11. Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Silicon-nitride waveguides for on-chip Raman spectroscopy*, SPIE Photonics Europe 14, 2014.
 12. Neutens, P.; Subramanian, A.Z.; Ul Hassan, M.; Chen, C.; Jansen, R.; Claes, T.; Rottenberg, X.; Du Bois, B.; Leyssens, K.; Helin, P.; Severi, S.; Dhakal, A.; Peyskens, F.; Lagae, L.; Deshpande, P.; Baets, R.; Van Dorpe, P. In *Characterization of PECVD silicon nitride photonic components at 532 and 900 nm wavelength*, 4th Conference on Silicon Photonics and Photonic Integrated Circuits, Belgium, 2014; pp. article 91331F (6 pages).

13. Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Raman spectroscopy using photonic waveguides*, IPS Benelux 2013, Netherlands, 2013.
14. Le Thomas, N.; Dhakal, A.; Peyskens, F.; Subramanian, A.Z.; Claes, T.; De Vos, K.; Ryckeboer, E.M.P.; Bockstaele, R.; Bienstman, P.; Baets, R. In *Biological sensing with integrated silicon and silicon nitride photonics*, The 2nd BioPhotonics Conference, Taiwan, 2013; pp. 28-29.
15. Dhakal, A.; Peyskens, F.; Subramanian, A.Z.; Le Thomas, N.; Baets, R. In *Enhancement of light absorption, scattering and emission in high index contrast waveguides*, OSA- Advanced Photonics Congress, Optical Sensors, United States, 2013; pp. ST2B.5.
16. Peyskens, F.; Subramanian, A.Z.; Dhakal, A.; Le Thomas, N.; Baets, R. In *Enhancement of Raman Scattering Efficiency by a Metallic Nano-antenna on Top of a High Index Contrast Waveguide*, CLEO 2013, United States, 2013; pp. paper CM2F.5.pdf.
17. Baets, R.; Subramanian, A.Z.; Dhakal, A.; Selvaraja, S.; Komorowska, K.; Peyskens, F.; Ryckeboer, E.M.P.; Yebo, N.A.; Roelkens, G.; Le Thomas, N. In *Spectroscopy-on-chip applications of silicon photonics*, Photonics West, United States, 2013; pp. 86270I-1 - 86270I-10.

References

- [1] Smith, E.; Dent, G. *Modern Raman Spectroscopy - A Practical Approach*; John Wiley & Sons Ltd, Chichester, UK, 2005.
- [2] McCreery, R.L. *Raman Spectroscopy for Chemical Analysis*; John Wiley & Sons Inc., Hoboken, NJ, USA, 2000.
- [3] Subramanian, A.Z.; Ryckeboer, E.M.P.; Dhakal, A.; Peyskens, F.; Malik, A.; Kuyken, B.; Zhao, H.; Pathak, S.; Ruocco, A.; De Groote, A.; Wuytens, P.C.; Martens, D.; Leo, F.; Xie, W.; Dave, U.D.; Muneeb, M.; Van Dorpe, P.; Van Campenhout, J.; Bogaerts, W.; Bienstman, P.; Le Thomas, N.; Van Thourhout, D.; Hens, Z.; Roelkens, G.; Baets, R. Silicon and silicon nitride photonic circuits

- for spectroscopic sensing on-a-chip (Invited). *Photon. Res.* **2015**, *3*(5), B47–B59.
- [4] Dhakal, A.; Subramanian, A.Z.; Wuytens, P.; Peyskens, F.; Le Thomas, N.; Baets, R. Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides. *Opt. Lett.* **2014**, *39*, 4025–4028.
- [5] Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Bright and dark plasmon resonances of nanoplasmonic antennas evanescently coupled with a silicon nitride waveguide. *Opt. Express* **2015**, *23*, 3088–3101.
- [6] Peyskens, F.; Subramanian, A.Z.; Dhakal, A.; Le Thomas, N.; Baets, R. In *Enhancement of Raman Scattering Efficiency by a Metallic Nano-antenna on Top of a High Index Contrast Waveguide*, CLEO 2013, San Jose, California, United States, June 09-14, 2013.
- [7] Arnaud, L.; Bruyant, A.; Renault, M.; Hadjar, Y.; Salas-Montiel, R.; Apuzzo, A.; L  rondel, G.; Morand, A.; Benech, P.; Le Coarer, E.; Blaize, S. Waveguide-coupled nanowire as an optical antenna. *J. Opt. Soc. Am. A* **2013**, *30*, 2347–2355.
- [8] Arango, F. B.; Kwadrin, A.; Koenderink, A. F. Plasmonic Antennas Hybridized with Dielectric Waveguides. *ACS Nano* **2012**, *6*, 10156–10167.
- [9] F  vrier, M.; Gogol, P.; Aassime, A.; M  gy, R.; Delacour, C.; Chelnokov, A.; Apuzzo, A.; Blaize, S.; Lourtioz, J-M.; Dagens, B. Giant Coupling Effect between Metal Nanoparticle Chain and Optical Waveguide. *Nano Lett.* **2012**, *12*, 1032–1037.
- [10] F  vrier, M.; Gogol, P.; Lourtioz, J-M.; Dagens, B. Metallic nanoparticle chains on dielectric waveguides: coupled and uncoupled situations compared. *Opt. Express* **2013**, *21*, 24504–24513.
- [11] F  vrier, M.; Gogol, P.; Barbillon, G.; Aassime, A.; M  gy, R.; Bartenlian, B.; Lourtioz, J-M.; Dagens, B. Integration of short gold nanoparticles chain on SOI waveguide toward compact integrated bio-sensors. *Opt. Express* **2012**, *20*, 17402–17410.

- [12] Chamanzar, M.; Xia, Z.; Yegnanarayanan, S.; Adibi, A. Hybrid integrated plasmonic-photonic waveguides for on-chip localized surface plasmon resonance (LSPR) sensing and spectroscopy. *Opt. Express* **2013**, *21*, 32086–32098.
- [13] Lin, S.; Zhu, W.; Jin, Y.; Crozier, K.B. Surface-Enhanced Raman Scattering with Ag Nanoparticles Optically Trapped by a Photonic Crystal Cavity. *Nano Lett.* **2013**, *13*, 559-563.
- [14] Kong, L.; Lee, C.; Earhart, C.M.; Cordovez, B.; Chan, J.W. A nanotweezer system for evanescent wave excited surface enhanced Raman spectroscopy (SERS) of single nanoparticles. *Opt. Express* **2015**, *23*, 6793-6802.
- [15] Measor, P.; Seballos, L.; Yin, D.; Zhang, J.Z.; Lunt, E.J.; Hawkins, A.R.; Schmidt, H. On-chip surface-enhanced Raman scattering detection using integrated liquid-core waveguides. *Appl. Phys. Lett.* **2007**, *90*, 211107-1-211107-3.

2

Theory of Plasmon Optics

2.1 Introduction

In this chapter we will give a theoretical treatment of plasmon optics. First the Maxwell equations and corresponding boundary conditions are introduced. Then a short discussion on Surface Plasmon Polaritons (SPPs) is given. Subsequently we present a thorough analysis on Localized Surface Plasmons (LSPs) and their associated field enhancement.

2.2 Maxwell revisited

All classical electromagnetic interactions are governed by Maxwell's equations, which unify position \mathbf{r} and time t dependent electric $\mathbf{E}(\mathbf{r}, t)$ and magnetizing $\mathbf{H}(\mathbf{r}, t)$ fields through a set of four partial differential equations:

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.1)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} + \mathbf{J}(\mathbf{r}, t) \quad (2.2)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \quad (2.3)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (2.4)$$

where $\mathbf{D}(r, t)$ is the electric displacement field, $\mathbf{B}(r, t)$ the magnetic induction field, $\mathbf{J}(\mathbf{r}, t)$ the external electric current density and $\rho(\mathbf{r}, t)$ the external electric charge density. [1] The electromagnetic properties of a medium are most commonly described in terms of the macroscopic polarization $\mathbf{P}(r, t)$ and magnetization $\mathbf{M}(r, t)$:

$$\mathbf{D}(r, t) = \epsilon_0 \mathbf{E}(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) \quad (2.5)$$

$$\mathbf{H}(r, t) = \frac{\mathbf{B}(\mathbf{r}, t)}{\mu_0} - \mathbf{M}(\mathbf{r}, t) \quad (2.6)$$

where ϵ_0 and μ_0 are the vacuum permittivity and permeability, respectively. For a linear, non-dispersive and isotropic medium these relations can be written as

$$\mathbf{D} = \epsilon_0 \epsilon \mathbf{E} \quad (2.7)$$

$$\mathbf{B} = \mu_0 \mu \mathbf{H} \quad (2.8)$$

where ϵ and μ are the relative permittivity and permeability of the medium and $\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}$ and $\mathbf{M} = \chi_m \mathbf{H}$, with χ_e and χ_m the electric and magnetic susceptibility. For non-linear media, \mathbf{D} and \mathbf{B} can depend on higher powers of \mathbf{E} and \mathbf{H} respectively. Anisotropic media are characterized by introducing a tensorial $\vec{\epsilon}$ and $\vec{\mu}$. In case of an inhomogeneous medium the material parameters ϵ and μ are functions of the spatial coordinate \mathbf{r} . If on top of that ϵ and μ are functions of frequency, the medium is called temporally dispersive. When the constitutive relations (2.7) and (2.8) are convolutions over space, then the medium is spatially dispersive. [2]

At the interface between two different media (characterized by different permittivity and permeability), the fields need to satisfy a set of boundary conditions in order to be faithful solutions of the above equations:

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \quad (2.9)$$

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = \rho_s \quad (2.10)$$

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{J}_s \quad (2.11)$$

$$\mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 \quad (2.12)$$

in which \mathbf{n} is the normal vector (from medium 1 to medium 2) to the interface separating medium 1 and 2, $\rho_s(\mathbf{r}, t)$ is a surface charge density and $\mathbf{J}_s(\mathbf{r}, t)$ a surface current density. In most cases there are no sources in the individual domains, hence ρ_s and \mathbf{J}_s vanish. [2]

2.3 Surface Plasmon Polaritons

In this section Maxwell's equations are applied to the specific case depicted in Figure 2.1. We consider a surface wave propagating along the interface between a dielectric and metallic medium, characterized by the relative permittivities ϵ_D and ϵ_M respectively. We first discuss the situation in which a wave with magnetic field \mathbf{H} parallel to the interface (p-polarized wave) propagates along the surface $z = 0$. Furthermore we require an exponential decay in the positive and negative z -direction. [3, 4]

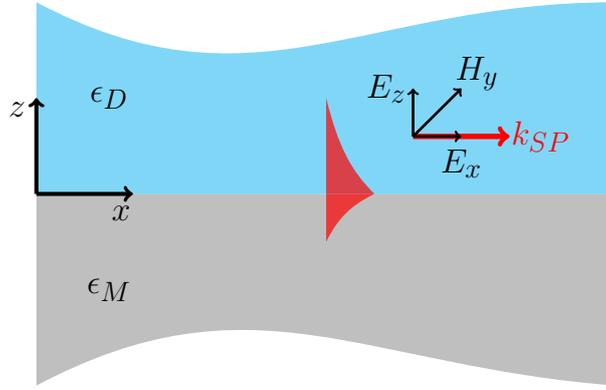


Figure 2.1: Surface plasmon polariton (red area) at the interface of a semi-infinite dielectric (ϵ_D) and metallic (ϵ_M) medium.

The fields in each medium i ($i = \{D, M\}$) are hence given by

$$\mathbf{E}^i = (E_x^i, 0, E_z^i) e^{-\kappa_i |z|} e^{j(q_i x - \omega t)} \quad (2.13)$$

$$\mathbf{H}^i = (0, H_y^i, 0) e^{-\kappa_i |z|} e^{j(q_i x - \omega t)} \quad (2.14)$$

in which κ_i is the decay constant along the z -direction, q_i the magnitude of a wavevector along the propagation direction and ω the angular frequency. Substituting equations (2.13) and (2.14) into the second Maxwell equation (2.2) results in

$$-\kappa_M H_y^M = -j\omega\epsilon_0\epsilon_M E_x^M \quad (2.15)$$

$$\kappa_D H_y^D = -j\omega\epsilon_0\epsilon_D E_x^D \quad (2.16)$$

$$q_M H_y^M = -\omega\epsilon_0\epsilon_M E_z^M \quad (2.17)$$

$$q_D H_y^D = -\omega\epsilon_0\epsilon_D E_z^D \quad (2.18)$$

while substitution in equation (2.1) gives

$$\kappa_M E_x^M - jq_M E_z^M = j\omega\mu_0 H_y^M \quad (2.19)$$

$$-\kappa_D E_x^D - jq_D E_z^D = j\omega\mu_0 H_y^D. \quad (2.20)$$

Combining equations (2.15), (2.17) and (2.19) or (2.16), (2.18) and (2.20) yields

$$\kappa_i^2 = q_i^2 - \epsilon_i \frac{\omega^2}{c^2}. \quad (2.21)$$

The boundary conditions (2.9) and (2.11) imply

$$\frac{\kappa_D}{\epsilon_D} H_y^D + \frac{\kappa_M}{\epsilon_M} H_y^M = 0 \quad (2.22)$$

$$H_y^D - H_y^M = 0 \quad (2.23)$$

This system only has a solution provided the determinant is zero:

$$\boxed{\frac{\kappa_D}{\epsilon_D} + \frac{\kappa_M}{\epsilon_M} = 0} \quad (2.24)$$

which is the so-called surface plasmon condition. [3] A plasmon is a collective oscillation of the free electrons in a metal. It can be described as a quantum of plasma oscillation. Such plasmon oscillations can be considered as oscillations of the electron gas of a metal, arising as a result of an external electric field that causes displacements of the electron gas with respect to the fixed ionic cores. At the surface of a metal, plasmons appear in the form of surface plasmons (SPs) while the combined excitation consisting of an SP and a photon is called a surface plasmon polariton (SPP). [4, 5]

The boundary conditions furthermore imply $q_D = q_M = k_{SP}$ (invariance per translation), such that the surface plasmon condition can also be expressed as

$$\boxed{k_{SP}(\omega) = \frac{\omega}{c} \sqrt{\frac{\epsilon_D \epsilon_M}{\epsilon_D + \epsilon_M}}}. \quad (2.25)$$

If the metal can be characterized by a Drude model, the permittivity is given by

$$\epsilon_M(\omega) = 1 - \frac{\omega_P^2}{\omega^2 + i\omega\Gamma} \quad (2.26)$$

in which ω_P is the bulk plasma frequency of the metal and Γ a measure for the friction experienced by the electrons during their motion in the metal.

[1] The plasma frequency can be expressed in terms of the electron mass m_e , charge e and density n_e :

$$\omega_P = \frac{n_e e^2}{\epsilon_0 m_e}. \quad (2.27)$$

A derivation of the Drude model can be found in [1]. For a Drude metal, the dispersion relation (2.25) has two solutions

$$\begin{aligned} \omega_{\pm}^2(k_{SP}) = & k_{SP}^2 c^2 \left(\frac{1 + \epsilon_D}{\epsilon_D} \right) + \frac{\omega_P^2}{2} \\ & \pm \sqrt{k_{SP}^4 c^4 \left(\frac{1 + \epsilon_D}{\epsilon_D} \right)^2 + \frac{\omega_P^4}{4} + \frac{k_{SP}^2 c^2 \omega_P^2}{2} \left(1 - \frac{1}{\epsilon_D} \right)}. \end{aligned} \quad (2.28)$$

which are plotted in Figure 2.2 for a metal-air interface ($\epsilon_D = 1$). The ω_- curve represents the surface plasmon polariton (SPP) solution. For large k_{SP} (so when the denominator in equation (2.25) becomes very small), the dispersion curve converges to the classical surface plasmon (SP) frequency ω_{SP} which is determined by

$$\text{Re}(\epsilon_M(\omega_{SP})) = -\epsilon_D. \quad (2.29)$$

or $\omega_{SP} = \frac{\omega_P}{\sqrt{1 + \epsilon_D}}$ which boils down to $\frac{\omega_P}{\sqrt{2}}$ for a metal-air interface. This non-retarded SP regime holds as long as the phase velocity ω/k_{SP} is much smaller than the speed of light. In the retarded regime ($k_{SP} \ll \omega_{SP}/c$) the SPP curve approaches the free space (FS) light line $\omega = ck_{SP}$. [3, 4]

The SPP field decay is determined by

$$\kappa_i = \frac{\omega}{c} \sqrt{\frac{-\epsilon_i^2}{\epsilon_M + \epsilon_D}} \quad (2.30)$$

which can be obtained by substituting (2.25) into (2.21). Because the dielectric constant of a metal is usually much larger than that of a dielectric, the SPP field has a shorter penetration depth ($L_i \propto 1/\kappa_i$) in the metal. Due to the SPP field concentration close to the metal, the field is significantly enhanced at the surface. [3, 4] This surface enhancement effect will be considered in more detail in the next section.

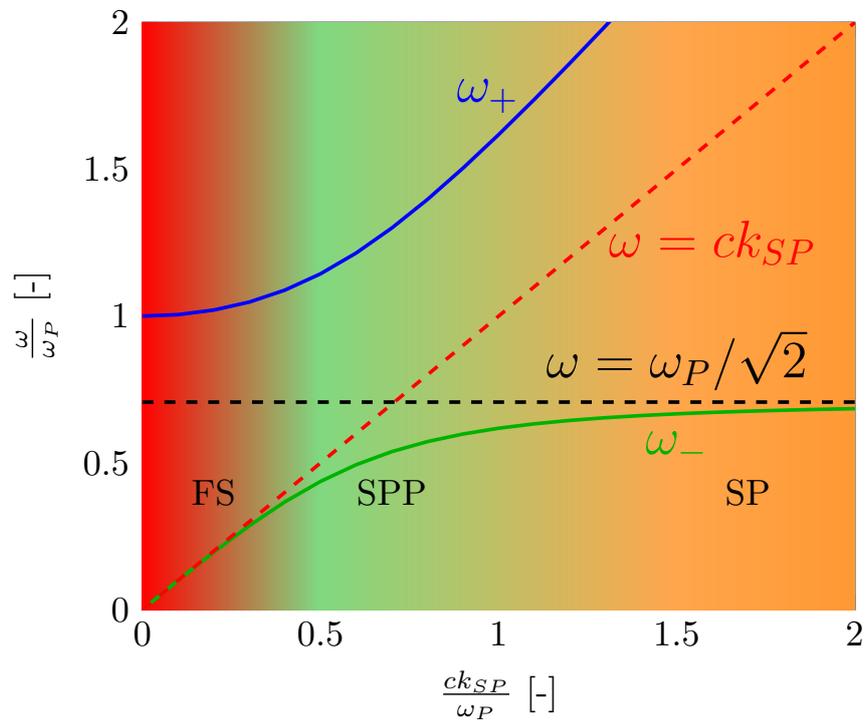


Figure 2.2: Dispersion relation for surface plasmon polaritons (SPP). For small k_{SP} the SPP regime (green area) approaches the free space (FS) light line (red area), while for large k_{SP} the SPP regime converges to the non-dispersive surface plasmon (SP) regime.

In second instance we can also consider an s-polarized wave, where the electric field vector is now perpendicular to the plane of incidence. The solutions of Maxwell's equations that are wavelike in the x -direction and localized to the interface can then be written as

$$\mathbf{E}^i = (0, E_y^{s,i}, 0) e^{-\kappa_i^s |z|} e^{j(q^s x - \omega t)} \quad (2.31)$$

$$\mathbf{H}^i = \frac{cE_y^{s,i}}{j\omega} (\text{sgn}(z)\kappa_i^s, 0, jq) e^{-\kappa_i^s |z|} e^{j(q^s x - \omega t)} \quad (2.32)$$

The continuity of the tangential components of the electric and magnetic field then yields $E_y^{s,D} = E_y^{s,M} = 0$. So an s-polarized SPP cannot exist in the structure depicted in Figure 2.1. [6]

The wave vector k_{SP} entering the SPP dispersion relation is a 2D wave vector in the plane of the surface. Hence, if light hits the surface in an arbitrary direction, then the external radiation dispersion line will always lie above the light line ck_{SP} , inhibiting an intersection with the SPP line. Due to this wave vector mismatch, SPPs cannot be excited by light incident on an ideal surface. Special configurations have been designed to provide conservation of the wavevector. In the so-called Kretschmann configuration (see Figure 2.3), the metal is illuminated through a dielectric prism (permittivity ϵ_P) at an incidence angle which is greater than the total internal reflection angle. When the in-plane component of the wavevector in the prism matches with the SPP wavevector of an air-metal interface, light can be coupled to SPPs. This occurs at an incidence angle θ_{SP} which satisfies

$$\frac{\omega}{c} \sqrt{\epsilon_P} \sin(\theta_{SP}) = k_{SP}. \quad (2.33)$$

Under these resonant conditions (i.e. light incident at an angle θ_{SP}), a sharp minimum is observed in the reflectivity from the prism interface because light can be coupled to SPPs with almost 100% efficiency. [3, 4] This phenomenon can be used for high sensitivity surface plasmon resonance (SPR) sensing. Due to the SPR, the reflectance has a minimum at a specific wavelength and angle. If the refractive index of the dielectric on top of the metal changes, the resonance angle θ_{SP} will shift. Changes in the refractive index are then usually detected through variations in the reflected laser light at the back of the film. SPR sensing is a powerful analytical technique since it can detect sub-monolayer quantities of analyte at a metal surface and provides kinetics data through continuous optical measurements. One drawback is that SPR sensors lack a localized sensing volume, which in

many cases requires a thick polymer layer or matrix that covers the metal surface and then must be suffused with a dedicated capture antibody to generate a detectable binding signal. [5] For a more detailed overview of other SP excitation strategies and SPR we refer the reader to the following reference [4]. Highly localized sensing volumes can be obtained through Localized Surface Plasmons, which will be discussed in the next section.

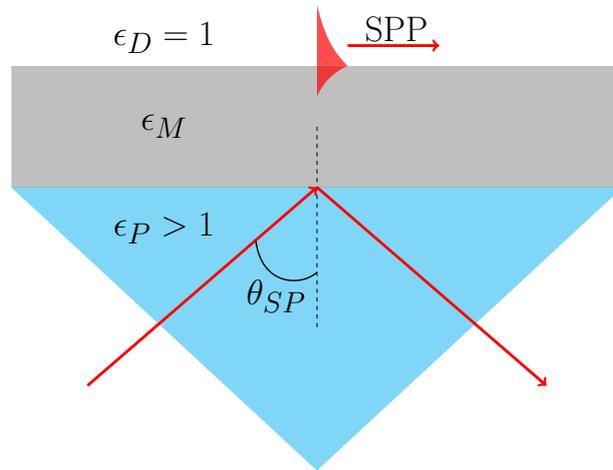


Figure 2.3: Kretschmann configuration to excite SPPs by wavevector matching using photon tunnelling in a total internal reflection geometry.

2.4 Localized Surface Plasmons

In addition to surface plasmons on a plane surface, localized surface plasma excitations can be considered in geometries involving bound electron plasmas such as metallic nanoparticles or voids. Such surface plasma excitations in bounded geometries are called localized surface plasmons (LSPs). The LSP resonance wavelength can be determined in the electrostatic approximation by solving the classical Dirichlet problem. Such an approximation is valid if the characteristic size, a , of the system is small compared to the wavelength of light λ , i.e. $a < \lambda$. [4] It should be emphasized that the SPP considered in the previous section is a different type of excitation than the LSP. An SPP is determined by the dispersion relation (2.25) and is a propagating surface mode along a continuous

metal-dielectric interface. In contrast, LSPs are confined to isolated metal particles and are characterized by discrete, complex frequencies which depend on the size and shape of the object to which the surface plasmon is confined, as well as its dielectric constant and the environmental refractive index. LSPs can be resonantly excited with light of appropriate wavelength (and polarization) irrespective of the excitation light wavevector. Therefore, LSPs also decay radiatively. In contrast, the SPP mode can only be excited if both the wavelength and wavevector of the excitation light match the SPP wavelength and wavevector. Because of their strong dependence on the refractive index of the surrounding, LSPs can also be used for refractive index shift sensing (similar to the previously discussed SPR sensing). For more information on this topic we refer the reader to [5, 7–10]. Since LSPs are confined to the particle, this results in a significant electromagnetic field enhancement near the surface of small metallic particles due to the small volume of the LSP mode. [4] Because the field enhancement associated to LSPs is essential for the description of Surface Enhanced Raman Spectroscopy, which is the main topic of this PhD, we will give a thorough overview of the physics involved in the generation of enhanced electric fields. To this end, the LSP modes of a spherical particle will be studied in detail as they can be calculated analytically and provide a good physical understanding of LSP optics.

We consider a spherical metallic particle with radius R much smaller than the wavelength λ as shown in Figure 2.4(a). As a result of the external electric field the free electrons in the metal move with respect to the fixed positive ionic cores. The incident wave is polarized along the z -axis and propagates along the y -direction, such that the driving electric field can be represented by

$$\mathbf{E}_0 = E_z \cos(ky - \omega t) \mathbf{e}_z \quad (2.34)$$

where $k = 2\pi/\lambda$ and $\omega = kc$. In the long wavelength approximation, $ky \ll \omega t$, such that $\mathbf{E}_0 \approx E_z \cos(\omega t) \mathbf{e}_z$. Since $\lambda \gg R$ we can furthermore assume that the incident electric field does not change much over the volume of the sphere. In this case, Maxwell's equations reduce to the laws of electrostatics. Since there is no net charge density inside the metal sphere ($\rho = 0$) and the relative permittivity of both the metal and the surrounding dielectric is constant in space, we get

$$\nabla \times \mathbf{E} = 0 \Rightarrow \mathbf{E} = -\nabla\varphi \quad (2.35)$$

$$\nabla \cdot \mathbf{D} = \rho = 0 \Rightarrow \nabla \cdot \mathbf{E} = 0 \quad (2.36)$$

where φ is the electrostatic potential which satisfies Laplace's equation because

$$\nabla \cdot \mathbf{E} = 0 \Rightarrow \nabla \cdot (-\nabla\varphi) = 0 \Rightarrow \Delta\varphi = 0. \quad (2.37)$$

Converting to spherical coordinates (r, θ, ϕ) , which are related to the cartesian coordinates (x, y, z) according to

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta, \quad (2.38)$$

yields the following solutions for the Laplace equation [7]:

$$\varphi_{(l,m)}(r, \theta, \phi) = \left(A_l r^l + B_l r^{-(l+1)} \right) P_l^{|m|}(\cos \theta) e^{jm\phi}. \quad (2.39)$$

The solutions are quantified by the integer numbers l and m which can take on the following values

$$l = 0, 1, 2, \dots \quad (2.40)$$

$$m = -l, -l+1, \dots, -1, 0, 1, \dots, l-1, l. \quad (2.41)$$

The radial part is dependent on the constants A_l and B_l , which are determined by the specific boundary conditions of the problem. The angular part is determined by the associated Legendre polynomials $P_l^{|m|}(\cos \theta)$. The first solutions for $P_l^{|m|}(\cos \theta)$ can be found in [7].

For a single metallic sphere, the problem is completely spherically symmetric, so the solution should be independent of ϕ . This corresponds to the $m = 0$ solutions, so the ϕ and m dependence will be omitted furtheron. In addition, the electric field has to converge to 0 when $r \rightarrow \infty$, while it has to be finite when $r \rightarrow 0$. Since

$$\mathbf{E} = -\nabla\varphi = -\frac{\partial\varphi}{\partial r} \mathbf{e}_r - \frac{1}{r} \frac{\partial\varphi}{\partial\theta} \mathbf{e}_\theta \quad (2.42)$$

these boundary conditions can only be satisfied if

$$\varphi_l = \begin{cases} A_l r^l P_l^0(\cos \theta), & r < R \\ B_l r^{-(l+1)} P_l^0(\cos \theta), & r \geq R \end{cases} \quad (2.43)$$

corresponding to

$$\mathbf{E}_l = \begin{cases} \frac{A_l(l+1)}{r^{l-1}} \left(-\frac{l}{l+1} P_l^0(\cos \theta) \mathbf{e}_r + \frac{\cos \theta P_l^0(\cos \theta) - P_{l+1}^0(\cos \theta)}{\sin \theta} \mathbf{e}_\theta \right), & r < R \\ \frac{B_l(l+1)}{r^{(l+2)}} \left(P_l^0(\cos \theta) \mathbf{e}_r + \frac{\cos \theta P_l^0(\cos \theta) - P_{l+1}^0(\cos \theta)}{\sin \theta} \mathbf{e}_\theta \right), & r \geq R. \end{cases} \quad (2.44)$$

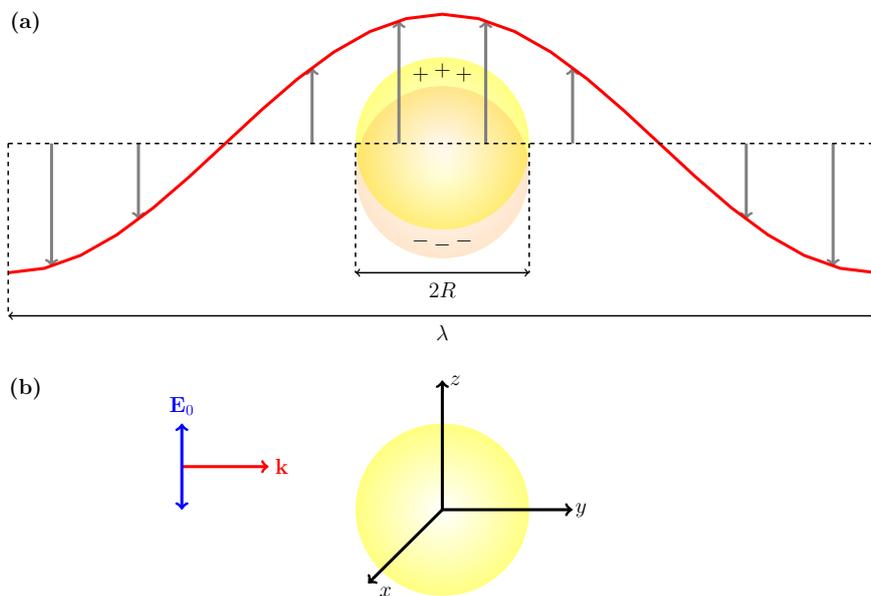


Figure 2.4: (a) Spherical metallic particle (radius R) illuminated by a wave with wavelength $\lambda \gg R$. The electrons (-) move with respect to the fixed ionic cores (+) as a result of the driving electric field which is almost constant over the volume of the sphere (electrostatic approximation). (b) Axis system: the electric field of the incident light is polarized along the z -axis and propagates along the y -direction.

These equations are obtained by using the following properties of the associated Legendre functions [11]:

$$\frac{dP_l^0(\cos\theta)}{d\theta} = \frac{l \cos\theta P_l^0(\cos\theta) - l P_{l-1}^0(\cos\theta)}{\sin\theta} \quad (2.45)$$

$$l P_{l-1}^0(\cos\theta) = (2l+1) \cos\theta P_l^0(\cos\theta) - (l+1) P_{l+1}^0(\cos\theta). \quad (2.46)$$

At the boundary between the metal and the dielectric the normal vector \mathbf{n} equals \mathbf{e}_r such that the boundary condition (2.9) at $r = R$ yields

$$A_l(l+1)R^{l-1} = B_l(l+1)R^{-(l+2)} \Rightarrow B_l = A_l R^{2l+1}. \quad (2.47)$$

By introducing $C_l = A_l R^l$, one obtains the following expressions for the electric field

$$\mathbf{E}_l = \begin{cases} \frac{C_l(l+1)}{R} \left(\frac{r}{R}\right)^{l-1} \left(-\frac{l}{l+1} P_l^0(\cos\theta) \mathbf{e}_r + \frac{\cos\theta P_l^0(\cos\theta) - P_{l+1}^0(\cos\theta)}{\sin\theta} \mathbf{e}_\theta \right), & r < R \\ \frac{C_l(l+1)}{R} \left(\frac{R}{r}\right)^{l+2} \left(P_l^0(\cos\theta) \mathbf{e}_r + \frac{\cos\theta P_l^0(\cos\theta) - P_{l+1}^0(\cos\theta)}{\sin\theta} \mathbf{e}_\theta \right), & r \geq R. \end{cases} \quad (2.48)$$

By imposing the boundary condition (2.10), one finds the condition to which all resonance frequencies ω_l should satisfy (see [12])

$$\epsilon_D(\omega_l) - \epsilon_M(\omega_l) \left(-\frac{l}{l+1} \right) = 0 \Rightarrow (l+1)\epsilon_D(\omega_l) + l\epsilon_M(\omega_l) = 0. \quad (2.49)$$

For a metal characterized by a Drude model and a non-dispersive dielectric this is equivalent to

$$\omega_l = \omega_P \sqrt{\frac{l}{l+(l+1)\epsilon_D}}. \quad (2.50)$$

If the order l increases ($l \rightarrow \infty$), the resonance frequency ω_l converges to the surface plasmon polariton frequency of a flat metal-dielectric surface

$$\omega_\infty = \omega_{SPP} = \frac{\omega_P}{\sqrt{1+\epsilon_D}}. \quad (2.51)$$

The discussion in the next sections mainly follows the theoretical analysis presented in [12, 13].

2.4.1 Optical excitation of plasmon modes

The surface charge density $\sigma_l(\theta)$ associated to the l th mode can be evaluated using the normal component of the electric field at the resonance

frequency ω_l and for $r \rightarrow R$:

$$\begin{aligned}
\sigma_l(\theta) &= (\epsilon_0 (\epsilon_M(\omega_l) - 1) \mathbf{E}_{in}(R, \theta) - \epsilon_0 (\epsilon_D - 1) \mathbf{E}_{out}(R, \theta)) \cdot \mathbf{e}_r. \\
&= \epsilon_0 (\epsilon_M(\omega_l) - 1) E_{max,l} \left(-\frac{l}{l+1} P_l^0(\cos \theta) \right) \\
&\quad - \epsilon_0 (\epsilon_D - 1) E_{max,l} P_l^0(\cos \theta). \\
&= \epsilon_0 E_{max,l} P_l^0(\cos \theta) \left((1 - \epsilon_M(\omega_l)) \left(\frac{l}{l+1} \right) - (\epsilon_D - 1) \right) \\
&= \epsilon_0 E_{max,l} P_l^0(\cos \theta) \left(\left(\frac{\omega_P^2}{\omega_l^2} \right) \left(\frac{l}{l+1} \right) - (\epsilon_D - 1) \right) \\
&= \epsilon_0 E_{max,l} P_l^0(\cos \theta) \left(\left(\frac{l + (l+1)\epsilon_D}{l} \right) \left(\frac{l}{l+1} \right) - (\epsilon_D - 1) \right) \\
&= \frac{2l+1}{l+1} \epsilon_0 E_{max,l} P_l^0(\cos \theta) \tag{2.52}
\end{aligned}$$

in which $E_{max,l} = \frac{C_l}{R}(l+1)$ is the maximum field just outside the metal sphere at $r = R$ and $\theta = 0$. [12] When the mode order l increases, the field gets more compressed closer to the metal surface because the field decays much faster ($\propto r^{-(l+2)}$) for larger l . Such a concentration of energy near the metal surface can be characterized through the effective mode volume $V_{eff,l}$ which is defined through the modal energy U_l :

$$U_l = \frac{1}{2} \epsilon_0 \epsilon_D E_{max,l}^2 V_{eff,l}. \tag{2.53}$$

The modal energy U_l itself is given by

$$U_l = \frac{1}{2} \int \int \varphi_l \sigma_l d^2 r \tag{2.54}$$

which eventually yields the following expression for the effective modal volume

$$V_{eff,l} = \frac{4\pi R^3}{(l+1)^2 \epsilon_D} = \frac{4\pi R^3}{3} \frac{3}{(l+1)^2 \epsilon_D} < \frac{4\pi R^3}{3}. \tag{2.55}$$

This volume is always smaller than the volume of the nanosphere (which in itself is already much smaller than λ^3). Electric fields near a metallic nanoparticle can hence be concentrated in volumes much smaller than the diffraction limit. As the mode index l increases, the effective mode volume decreases roughly with l^{-2} . The higher order modes are therefore very desirable to achieve enormous peak energy densities and concentrate electric fields to nanometer-size volumes. [13]

In order to make use of the extreme energy densities of higher order modes, one has to be able to couple an external excitation source to them. The ability to do so depends on the effective dipole moment \mathbf{p}_l of the mode. For single symmetrical nanoparticles which are much smaller than the wavelength, \mathbf{p}_l is given by

$$\begin{aligned}
\mathbf{p}_l &= \iint_{r=R} \sigma_l (R\mathbf{e}_r) dA. \\
&= \frac{2l+1}{l+1} R\epsilon_0 E_{max,l} \iint_{r=R} P_l^0(\cos\theta) \mathbf{e}_r dA \\
&= 2\pi \frac{2l+1}{l+1} R^3 \epsilon_0 E_{max,l} \int_0^\pi d\theta P_l^0(\cos\theta) \sin\theta \underbrace{\cos\theta}_{P_1(\cos\theta)} \\
&= 2\pi R^3 \epsilon_0 E_{max,l} \frac{2l+1}{l+1} \int_0^\pi d\cos\theta P_l^0(\cos\theta) P_1^0(\cos\theta) \\
&= 2\pi R^3 \epsilon_0 E_{max,l} \frac{3}{2} \delta_{l,1} = 2\pi R^3 \epsilon_0 E_{max,l} \delta_{l,1}. \tag{2.56}
\end{aligned}$$

The dipole mode ($l = 1$) is hence the only one which can couple to external radiation because it is the only mode with a non-vanishing dipole moment \mathbf{p}_l . For all higher order modes with $l \geq 2$, the dipole moment vanishes. This holds as long as the spherically symmetric particle is much smaller than the wavelength. By placing two metallic nanoparticles in each others vicinity, it is possible to achieve a coupling to the higher order modes as will be explained furtheron. Plasmon resonances with small dipole moments are typically referred to as dark modes, while modes with a larger dipole moment are referred to as bright modes. The probability to excite a plasmon mode using an incident plane wave is roughly proportional to the square of its dipole moment. Hence, bright modes will be visible in an optical spectrum and will be radiatively broadened. Dark modes cannot be excited with a pure plane wave and are hence only broadened by intrinsic damping (so they will typically have a higher Q -factor). [8] More information on bright and dark plasmon modes can be found in [8, 14–19].

2.4.2 Field enhancement

We will first consider the field enhancement that can be obtained for the case of a single metallic nanoparticle in which only the dipole mode is active. In order to calculate the field enhancement of the dipole mode, $E_{max,1}$ is compared to the field E_{foc} of a tightly focused light beam in

the absence of any metallic nanoparticles. [13] To this end, we consider a Gaussian beam with frequency ω (wavelength λ) and half angle θ_a which is focused to a diffraction limited spot with radius $w = \frac{\lambda}{\pi\theta_a}$. The field in the focal spot E_{foc} is related to the incident power $|s_+|^2 = \frac{n}{Z_0} \pi \left(\frac{w}{2}\right)^2 E_{foc}^2$ where n is the refractive index of the surrounding dielectric and Z_0 is the free space impedance. In the presence of a metal sphere, the incident light can be coupled into the dipole mode ($l = 1$) but not into the higher order modes because of their vanishing dipole moment. This process is reciprocal to the radiative decay of the dipole mode into the free space radiation modes. The incoupling coefficient κ_{in} can then be related to the radiative decay rate γ_{rad} of the dipole mode into a cone with solid angle Ω :

$$\kappa_{in} = \sqrt{\gamma_{rad} \Omega} \approx \frac{\theta_a}{2} \sqrt{\frac{3\gamma_{rad}}{2}} \quad (2.57)$$

in which $\gamma_{rad} = \frac{2\omega_1}{3\epsilon_D} \left(\frac{2\pi R}{\lambda_1}\right)^3 = \frac{2\omega_1}{3\epsilon_D} \chi^3$. [13] Simultaneously, all modes also experience nonradiative decay due to the imaginary part of the metal dielectric function at roughly the same rate that is equal to the metal loss in the Drude model $\gamma_{nr,ad,l} \approx \gamma$. The decay rate for all modes can thus be written as

$$\gamma_l = \begin{cases} \gamma_{rad} + \gamma & l = 1 \\ \gamma & l \geq 2. \end{cases} \quad (2.58)$$

The rate equation for the amplitude $A_1 = \sqrt{U_1}$ of the dipole mode is then given by

$$\frac{dA_1}{dt} = j(\omega - \omega_1)A_1 - \frac{\gamma_{rad} + \gamma}{2} A_1 + \kappa_{in} s_+. \quad (2.59)$$

In a steady state regime we obtain the field enhancement factor at a distance d ($\chi_d = 2\pi d/\lambda_D$) from the metal surface

$$\begin{aligned} F_S &= \left| \frac{E_{max,1}}{E_{foc}} \right| \left(\frac{R}{R+d} \right)^3 \\ &= \frac{\sqrt{2}}{\sqrt{\left(Q^{-1} + \frac{2\chi^3}{3\epsilon_D}\right)^2 + \delta^2}} \left(\frac{\chi}{\chi + \chi_d} \right)^3 \\ &= \frac{\sqrt{2}}{\sqrt{\left(Q^{-1} + \frac{\gamma_{rad}}{2\omega}\right)^2 + \delta^2}} \left(\frac{\chi}{\chi + \chi_d} \right)^3 \end{aligned} \quad (2.60)$$

where the Q -factor of the metal sphere is given by $Q = \frac{\omega}{\gamma}$ and the normalized excitation detuning δ is determined by $\delta = 2\left(1 - \frac{\omega_1}{\omega}\right)$. [13] The

Q -factor of gold in air is typically around 10 (at resonance). [20] So for small spheres ($R < \lambda/20$), $Q^{-1} > \chi^3$, such that $F_S \approx \sqrt{2}Q = 14$ near the resonance ($\omega = \omega_1$) and when evaluated near the surface of the sphere. The radiative decay rate γ_{rad} clearly reduces the maximum achievable field enhancement. The field enhancement for such a single metal sphere is relatively low mainly because the higher order modes, which are not subject to radiative decay, can simply not couple with the external field. The dipole mode which couples effectively with external radiation hence acts as an efficient antenna, but also as a bad resonator. On the other hand, higher order modes are efficient resonators which can confine the light into much smaller volumes, but are bad antennas. In order to obtain very high field enhancements it is hence required to have metallic structures that act as good antennas as well as good resonators. As outlined above it is impossible to achieve this with a single mode in a symmetric and small metal nanoparticle. Placing nanoparticles in close proximity to each other does allow the combination of antenna and resonator properties in order to boost the field enhancement. [13] This will be discussed in the next subsection.

2.4.3 Coupled nanoplasmonic particles

We will consider the case of two metallic nanospheres as shown in Figure 2.5. The amplitude A_l^i of the l -th mode on each of the two spheres ($i = \{1, 2\}$) is defined as in the previous section

$$A_l^i = \sqrt{U_l^i} = \sqrt{\frac{1}{2}\epsilon_0\epsilon_D V_{eff,l}^i E_{max,l}^i}. \quad (2.61)$$

The coupling energy between the different modes is obtained as a surface integral of the electric potential $\varphi_{l_1}^1$ of the l_1 -th mode of sphere 1 multiplied by the surface charge density $\sigma_{l_2}^2$ of the l_2 -th mode of sphere 2, and evaluated over the surface of sphere 2 (see [13]):

$$U_{l_1 l_2}^{12} = \iint \varphi_{l_1}^1 \sigma_{l_2}^2 dA_2 = -4\kappa_{l_1 l_2}^{12} A_{l_1}^1 A_{l_2}^2. \quad (2.62)$$

We are mainly interested in the coupling coefficients κ_{1l}^{12} between the dipole mode $l_1 = 1$ in one sphere and all the modes $l_2 = l$ in the other sphere. They are given by

$$\kappa_{1l}^{12} = \frac{l+1}{2} \left(\frac{R_1}{R_1 + R_2 + \Delta} \right)^{3/2} \left(\frac{R_2}{R_1 + R_2 + \Delta} \right)^{l+1/2}. \quad (2.63)$$

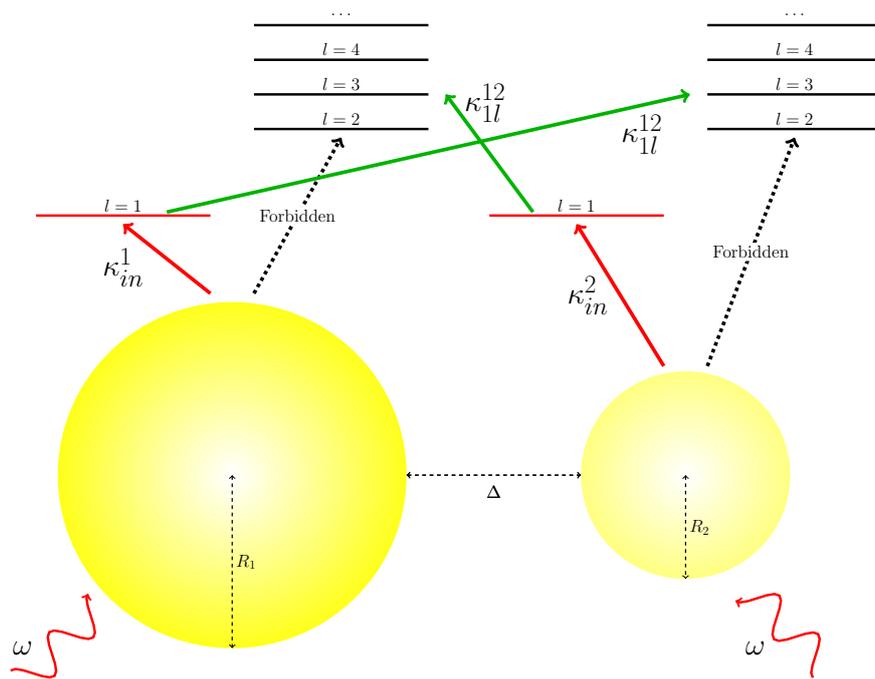


Figure 2.5: Coupled spherical particles with radii R_1 and R_2 and gap Δ inbetween. Each particle has a set of modes $\{l\}$. External radiation with frequency ω can only excite the dipolar ($l = 1$) mode on each sphere (red arrow), while exciting higher order modes is impossible (black dashed arrow). The dipolar mode on each sphere can however couple with the higher order modes on the other sphere (green arrows). The coupling coefficients κ between several modes are also depicted.

The coupling between the higher order modes only shifts the resonance frequencies of these modes by a small amount, which is typically smaller than the non-radiative broadening γ and will hence be omitted. [13] The energy balance equations for sphere i can then be written as

$$\frac{dA_1^i}{dt} = j(\omega - \omega_1)A_1^i - \frac{\gamma_1^i}{2}A_1^i + \kappa_{in}^i s_+ - j \sum_{l=1}^{\infty} \sqrt{\omega_1 \omega_l} \kappa_{1l}^{ij} A_l^j. \quad (2.64)$$

$$\frac{dA_l^i}{dt} = j(\omega - \omega_l)A_l^i - \frac{\gamma}{2}A_l^i - j\sqrt{\omega_1 \omega_l} \kappa_{1l}^{ji} A_1^j, \quad l \geq 2 \quad (2.65)$$

The dipolar mode $l = 1$ is the only mode which can couple to the external radiation field (κ_{in}^i) and hence also exhibits radiative damping apart from non-radiative damping ($\gamma_1^i = \gamma_{rad}^i + \gamma$). The higher order modes on the other hand can only be excited by the dipolar mode of the other sphere and do not decay radiatively ($\gamma_{l \geq 2} = \gamma$). At steady state one finds

$$\frac{E_{max,l}^2}{E_{max,1}^1} = \frac{\omega_{1l}}{(\omega_l - \omega) + j\frac{\gamma}{2}} \left(\frac{l+1}{2}\right)^2 \left(\frac{R_1}{S}\right)^{3/2} \left(\frac{R_2}{S}\right)^{l+1/2} \left(\frac{R_1}{R_2}\right)^{3/2} \quad (2.66)$$

where $\omega_{1l} = \sqrt{\omega_1 \omega_l}$ and $S = R_1 + R_2 + \Delta$. A similar expression between $E_{max,l}^2$ and $E_{max,l}^1$ can be found analogously. The electric field at a certain location in the gap is the sum of all modal contributions from both spheres. After tedious calculations outlined in [13] one can show that the maximum field enhancement for a zero gap configuration ($\Delta = 0$) of equal spheres ($R_1 = R_2 = R$), excited at the dipolar mode frequency $\omega = \omega_1$ is given by

$$F_D \approx \frac{9\sqrt{2}}{4} Q^2 \quad (2.67)$$

where Q is again the metal quality factor $Q = \omega/\gamma$. The field enhancement factor F_D of this so-called dimer configuration incorporates contributions from all higher order modes, of which the electric fields are determined by an expression of the form (2.66). In the case of a nanolens, $R_1 \gg R_2$. Now the large sphere acts as an antenna that allows the coupling of energy into the system, while the smaller sphere behaves like a cavity in which the energy can be concentrated. The maximum field enhancement for a nanolens (again for $\Delta = 0$) is

$$F_{NL} \approx 2\sqrt{2}Q^2. \quad (2.68)$$

Both equations (2.67) and (2.68) overestimate the enhancement because detuning and finite gap $\Delta > 0$ are not taken into account. They do

however provide the qualitative understanding that higher order modes in coupled nanoparticles allow higher field enhancements as compared to isolated single nanoparticles. For a single nanoparticle the maximum field enhancement scales with Q while coupled particles give rise to a Q^2 scaling. In order to boost the field enhancement one hence needs a metallic nanostructure that acts as an efficient antenna to interact with the external fields and moreover realizes a small effective mode volume. [13] High field enhancements are very useful to boost extremely weak Raman signals through a process called Surface Enhanced Raman Spectroscopy (SERS) as will be explained in the next chapter.

2.5 Conclusion

In this chapter we have outlined the physical mechanism underlying the field enhancement of Localized Surface Plasmon Resonances. For single metallic nanoparticles the field enhancement approximately scales with the Q -factor of the metal, while coupled nanoplasmonic particles exhibit a field enhancement which scales with the square of the Q -factor. Such coupled structures are hence of great interest for enhancing sensing processes where the observable quantity strongly depends on the electric field. In the next chapter we will discuss how extremely weak Raman signals can be boosted by the aforementioned plasmonic field enhancement.

References

- [1] Maier, S.A. *Plasmonics: Fundamentals and Applications*; Springer US, 2007.
- [2] Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press, New York, 2006.
- [3] Pitarke, J.M.; Silkin, V.M.; Chulkov, E.V.; Echenique, P.M. Theory of surface plasmons and surface-plasmon polaritons. *Rep. Prog. Phys.* **2007**, *70*, 1–87.
- [4] Zayats, A.V.; Smolyaninov, I.I. Near-field photonics: surface plasmon polaritons and localized surface plasmons. *J. Opt. A: Pure Appl. Opt.* **2003**, *5*, S16-S50.

- [5] Mayer, K.M.; Hafner, J.H. Localized surface plasmon resonance sensors. *Chem. Rev.* **2011** *111*, 3828–3857.
- [6] Zayats, A.V.; Smolyaninov, I.I.; Maradudin, A.A. Nano-optics of surface plasmon polaritons. *Phys. Rep.* **2005**, *408*, 131-314.
- [7] Willets, K.A.; Van Duyne, R. P. Localized surface plasmon resonance spectroscopy and sensing. *Annu. Rev. Phys. Chem.* **2007** *58*, 267-297.
- [8] Halas, N.J.; Lal, S.; Chang, W-S.; Link, S.; Nordlander, P. Plasmons in strongly coupled metallic nanostructures. *Chem. Rev.* **2011**, *111*, 3913–3961.
- [9] Mock, J.J.; Smith, D.R.; Schultz, S. Local refractive index dependence of plasmon resonance spectra from individual nanoparticles. *Nano Lett.* **2003** *3*, 485–491.
- [10] Haes, A.J.; et al. A localized surface plasmon resonance biosensor: first steps toward an assay for Alzheimers disease. *Nano Lett.* **2004** *4*.
- [11] Wolfram-Mathworld-Associated Legendre Polynomial. <http://mathworld.wolfram.com/AssociatedLegendrePolynomial.html> (accessed Sep 28, 2015).
- [12] Sun, G.; Khurgin, J.B.; Yang, C.C. Impact of high-order surface plasmon modes of metal nanoparticles on enhancement of optical emission. *Appl. Phys. Lett.* **2009** *95*, 171103-1-171103-3.
- [13] Sun, G.; Khurgin, J.B.; Bratkovsky, A. Coupled-mode theory of field enhancement in complex metal nanostructures. *Phys. Rev. B* **2011** *84*, 045415-1–045415-10.
- [14] Sonnefraud, Y.; Verellen, N.; Sobhani, H.; Vandenbosch, G.A.E.; Moshchalkov, V.V.; Van Dorpe, P.; Nordlander, P.; Maier, S.A. Experimental realization of subradiant, superradiant, and Fano resonances in ring/disk plasmonic nanocavities. *ACS Nano* **2010** *4*(3), 1664–1670.
- [15] Giannini, V.; Fernández-Domínguez, A. I.; Heck, S. C.; Maier, S. A. Plasmonic nanoantennas: fundamentals and their use in controlling the radiative properties of nanoemitters. *Chem. Rev.* **2011** *111*, 3888-3912.

-
- [16] Verellen, N.; Van Dorpe, P.; Vercruyssen, D.; Vandenbosch, G.A.E.; Moshchalkov, V.V. Dark and bright localized surface plasmons in nanocrosses. *Opt. Express* **2011** *19*, 11034–11051.
- [17] Chu, M-W.; Myroshnychenko, V.; Hsuan Chen, C.; Deng, J-P.; Mou, C-Y.; García de Abajo, F.J. Probing Bright and Dark Surface-Plasmon Modes in Individual and Coupled Noble Metal Nanoparticles Using an Electron Beam. *Nano Lett.* **2009** *9*, 399–404.
- [18] Liu, M.; Lee, T-W.; Gray, S.K.; Guyot-Sionnest, P.; Pelton, M. Excitation of Dark Plasmons in Metal Nanoparticles by a Localized Emitter. *PRL* **2009** *102*, 107401-1–107401-4.
- [19] Hao, F.; Larsson, E.M.; Ali, T.A.; Sutherland, D.S.; Nordlander, P. Shedding light on dark plasmons in gold nanorings. *Chem. Phys. Lett.* **2008** *458*, 262-266.
- [20] Bosman, M.; Zhang, L.; Duan, H.; Fen Tan, S.; Nijhuis, C.A.; Qiu, C-W.; Yang, J.K.W. Encapsulated annealing: enhancing the plasmon quality factor in lithographically defined nanostructures. *Sci. Rep.* **2014** *4*, 5537-1–5537-6.

3

Raman Spectroscopy

3.1 Introduction

The main optical spectroscopy techniques which are employed to detect vibrations in molecules are based on the processes of infrared absorption and Raman scattering. They are widely used to provide information on chemical structures and physical forms, to identify substances from the characteristic spectral patterns (so-called fingerprinting) and to quantitatively determine the amount of a substance in a sample. Raman scattering is less widely used than infrared absorption, largely due to problems with interfering fluorescence. However, recent advances in instrument technology have simplified the equipment and reduced the problems substantially. These advances, together with the ability of Raman spectroscopy to examine aqueous solutions, samples inside glass containers and samples without any preparation, have led to a rapid growth in the application of the technique. [1] In this chapter we will briefly introduce infrared absorption and fluorescence spectroscopy and provide a rationale for performing Raman spectroscopy as a (complementary) alternative. Apart from spontaneous Raman Spectroscopy we will also discuss Surface Enhanced Raman Spectroscopy (SERS), which will be the major topic of this dissertation.

3.2 Infrared absorption spectroscopy

A quantum mechanical analysis of molecules reveals that each electronic state can be subdivided in a set of vibrational (and rotational) energy levels. These levels originate from atomic motions inside the molecule (see Figure 3.1(a)). We will limit ourselves to the vibrational levels. When such vibrations can be approximated by a harmonic oscillator model, the energy spacings are equidistant ($\Delta E_{v1} = \Delta E_{v2} = \Delta E_{v3} = \Delta E_{v4} = \dots$) and given by

$$\Delta E_v = \hbar\omega_v \quad (3.1)$$

where $\omega_v = \frac{2\pi c}{\lambda_v} = 2\pi f_v$ is a characteristic vibration frequency (depending on the specific chemical bond). For diatomic molecules (such as I_2 , N_2 , H_2) the vibrational frequencies f_v are typically in the 10 – 100 THz range, so in the mid-infrared (mid-IR). [2] If a photon with energy $\hbar\omega$ is incident on a molecule, then it can be absorbed and excite a molecular vibration if its energy matches the transition energy between two different vibrational levels, i.e. $\omega = \omega_v$. In an assembly of a large number of molecules in thermal equilibrium, the relative populations of all energy levels is determined by a Boltzmann distribution

$$\frac{N_v}{N_0} = \exp\left(-\left(\frac{E_v - E_0}{kT}\right)\right) < 1 \quad (3.2)$$

where N_v and N_0 are the populations of the v -th energy level (with energy $E_v = \hbar\omega_v(v + 1/2)$) and the ground state (with energy $E_0 = \hbar\omega_v/2$) respectively. As a result most molecules will reside in the ground state and the $v = 0 \rightarrow 1$ transition will dominate the infrared absorption spectrum (this is the fundamental transition as shown in Figure 3.1(b)). Other allowed transitions such as $v = 1 \rightarrow 2$, $v = 2 \rightarrow 3, \dots$ will give rise to absorption at the same frequency as the fundamental transition (in case of a harmonic oscillator). The intensities of these transitions will however be relatively low because of the lower population of the higher energy levels. These transitions are termed hot band transitions because the population of the higher energy levels will increase with increasing temperature. In order to explain the overtone transitions we first need to introduce an additional selection rule for IR transitions, which is associated with the change in dipole moment of the molecule.

While the absorption frequency is determined by the molecular vibrational frequency, the absorption intensity depends on how effectively the photon

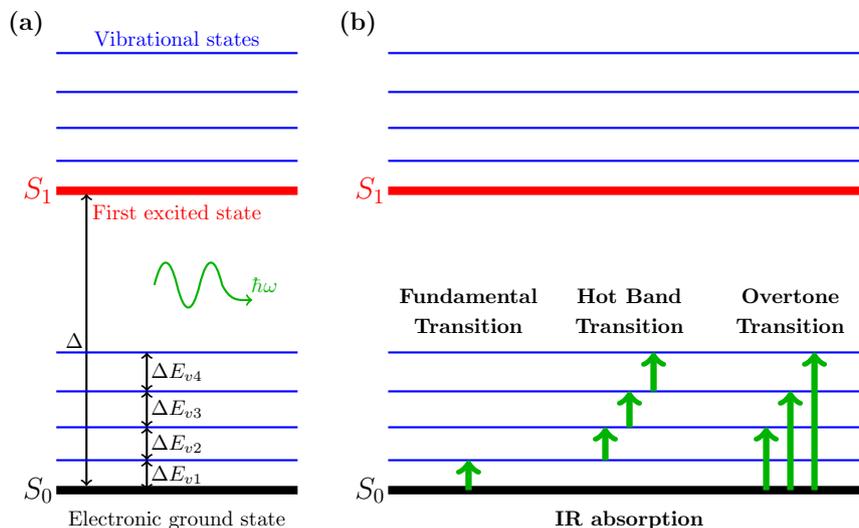


Figure 3.1: (a) Energy level diagram of a molecule. Each electronic state (S_0, S_1, \dots) is subdivided in a set of vibrational states with energy spacings $\Delta E_{v1}, \Delta E_{v2}, \Delta E_{v3}, \Delta E_{v4}, \dots$. A photon with energy $\hbar\omega$ can excite a molecular vibration when it is absorbed. (b) Different IR absorption transitions.

energy can be transferred to the molecule. This transfer rate is determined by the change of the dipole moment of the molecule and it can be shown that the absorption intensity is proportional to the square of the change in dipole moment. If a molecular vibration does not change the dipole moment, then that vibration mode will be IR inactive. In Figure 3.2 we depict a set of molecules and their associated modes. The modes in the left column are IR inactive because their dipole moment does not change upon vibration (due to the symmetry). In the right column a few IR active modes are shown. All heteronuclear diatomic molecules (AB) have a non-zero derivative of the dipole moment at the equilibrium position as a result of which the mode is IR active. For the polyatomic ABA molecule the antisymmetric stretching mode (center) or the bending mode (bottom) also have non-zero dipole moment derivatives in contrast to the symmetric stretching mode of the ABA molecule. This renders them IR active. Furtheron we will show that the IR inactive modes from Figure 3.2 can be probed by Raman scattering. [3]

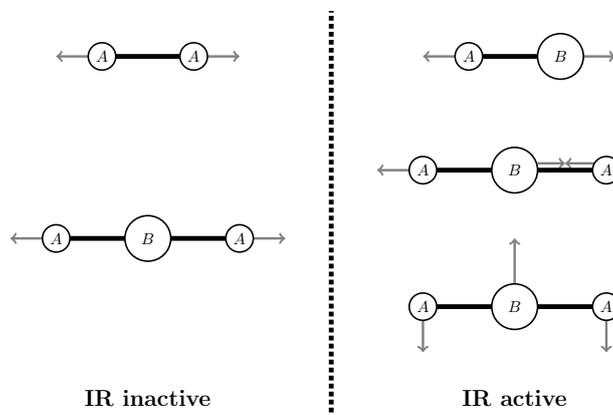


Figure 3.2: Different Molecular configurations. The blue arrows represent the vibration modes. Left column: IR inactive modes. Right column: IR active modes.

So apart from the energy constraint, the molecular vibration needs to have a non-zero gradient in the dipole moment. From the energy diagram in Figure 3.1(b) it is clear that the photon frequency should satisfy $\omega = 2\omega_v$ in order to excite the first overtone transition in the case of a harmonic oscillator. However, if the vibration is purely harmonic there will be no net dipole moment vibrating at this frequency, so it is impossible to excite this transition (and also the higher order overtones). In contrast, for an anharmonic oscillator (in which case $\Delta E_{v1} \neq \Delta E_{v2} \neq \Delta E_{v3} \neq \Delta E_{v4} \neq \dots$) the dipole moment may still vary periodically, but nonsinusoidal, in time. This periodic function can then be decomposed in sine and cosine components where the frequencies are integral multiples of the fundamental vibrational frequencies (by Fourier analysis). As a result, the dipole moment of an anharmonic oscillator may oscillate with the fundamental frequency and integer multiples thereof such that overtone transitions can be excited. The intensity of these absorption lines will depend on the amount of anharmonicity. Usually such transitions are however relatively weak (compared to the fundamental transition). [3]

IR absorption spectroscopy can be used for the identification and study of a large class of organic compounds (such as polymers, lipids, proteins, nucleic acids). It can for example be applied to monitor polymerization processes, determine the relative amounts of lipid present in blood serum, characterize tissues and diagnose certain diseases. [4] Despite the large amount of applications, there are also some disadvantages associated to this

technique. Most relevant fundamental absorption bands are however in the mid-IR region, which requires the use of expensive sources and detectors. One way to overcome this issue is to use the overtones, which lie in the near IR region where well-developed and cheaper light sources and detectors are available. On the downside, the overtone transitions are also weaker in absolute strength. An additional problem is water absorption, which is very high in the mid-IR. As a result it is hard to study compounds in an aqueous environment. [5] From Figure 3.3 it follows that the water absorption is much lower in the visible and near-IR wavelength region (400 – 1000 nm). Optical techniques, such as fluorescence or Raman spectroscopy, which operate in this region are hence of interest for the analysis of compounds in watery solvents. Moreover, high-quality and low-cost sources and detectors are available for this wavelength region. Before we extensively describe Raman spectroscopy, we will first give a brief introduction to fluorescence spectroscopy in the next section.

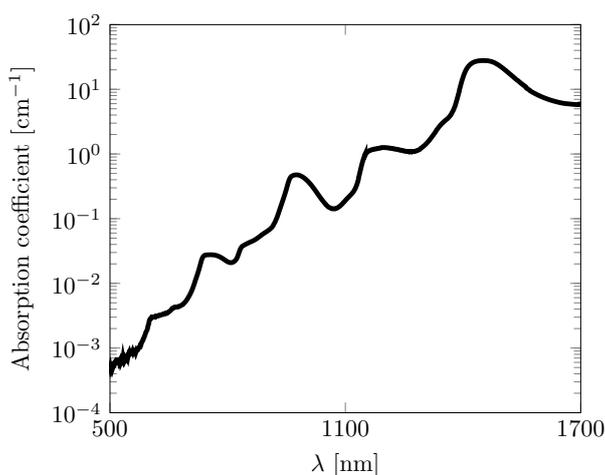


Figure 3.3: Absorption coefficient of water. [6]

3.3 Fluorescence spectroscopy

Fluorescent labelling and staining, when combined with an appropriate imaging instrument, is a sensitive and quantitative method that is widely used in molecular biology and biochemistry laboratories for a variety of

experimental, analytical, and quality control applications. Commonly used techniques, such as total nucleic acid and protein quantification and DNA sequencing, can all benefit from the application of fluorescence based methods for detection. Fluorophores, i.e. components that cause a molecule to absorb energy of a specific wavelength and then re-emit energy at a different but equally specific wavelength, permit sensitive detection of many biological molecules. Frequently they are the most sensitive option for detection of DNA, RNA, and proteins compared to traditional colourimetric methods. [7] The amount and wavelength of the emitted signal depend on both the fluorophore and its chemical environment. Fluorophores are also denoted as chromophores, which are historically speaking the part of a molecule responsible for its color. In addition, the denotation chromophore implies that the molecule absorbs light while fluorophore means that the molecule also emits light (as not all chromophores necessarily fluoresce). The umbrella term used is luminescence, whereas fluorescence denotes allowed transitions, with a lifetime in the nanosecond range, from higher to lower excited singlet states. [8, 9]

Since fluorophores play a central role in fluorescence spectroscopy and imaging, it is important to outline their interaction with incident light. In contrast to IR absorption spectroscopy, light absorption now results in transitions between different electronic levels rather than transitions between different vibrational levels of the same electronic level. In general the different electronic states will not have their minimum at the same nuclear coordinate (see Figure 3.4). Upon excitation, the molecule, which is in the vibrational S_0 ground state, will make a vertical transition rather than a transition to the S_1 minimum. This is a result of the Born-Oppenheimer approximation which states that electronic motions, when viewed from the perspective of the nuclear coordinates, occur as if the nuclei were fixed in space because the electrons move much faster than the nuclei. The Franck-Condon principle then states that an electronic transition is most likely to occur without changes to the position of the nuclei, hence electronic transitions can be considered instantaneous compared to the nuclear motion. This Franck-Condon state will quickly relax to the lowest vibrational level of the electronic excited state. For fluorophores in liquid solutions this equilibrium is reached in a time scale of the order of one picosecond. The Franck-Condon principle also applies to the emission (fluorescence). As a result of the vibrational excitation during the excitation and emission cycle, the fluorescence always

occurs at a redshifted (Stokes shifted) wavelength. Because of this redshift, the elastic (Rayleigh) scattering can be efficiently suppressed using appropriate filtering. From a quantum mechanical point of view, the overall fluorescence process is a two-step process comprising of a fast excitation and vibrational relaxation (time scale: picoseconds) and a slow fluorescence emission (time scale: nanoseconds). Hence the emission process is completely independent of the initial absorption. Both photons are not linked to each other in a coherent and instantaneous way. Consequently, some of the potentially emitted photons from the S_1 state can go missing due to e.g. nonradiative recombination. [8, 11] As we will see later, this plays a crucial role in determining the plasmonic enhancement and quenching of fluorescence processes.

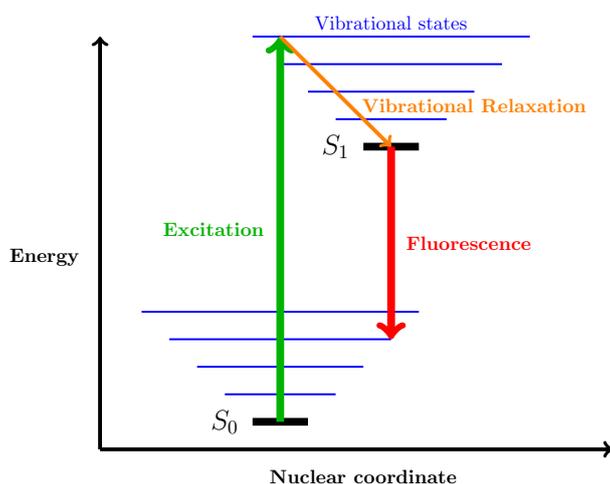


Figure 3.4: Energy level diagram for fluorescence, illustrating the Franck-Condon principle. Adapted from [8].

The spectral response of fluorophores is used in a variety of applications such as environmental monitoring, clinical chemistry, DNA sequencing, genetic analysis, cell identification, etc. If the target analyte of interest has no intrinsic fluorescence, then external fluorescent labels need to be added. [12] The addition of labels can be considered as a drawback because binding an external label can also change the intrinsic properties of the analyte under investigation. Furthermore the fluorophore can photobleach in which case it loses its applicability. Photobleaching denotes the loss of

fluorescence properties due to an irreversible reaction which changes the absorption and emission capabilities. It would hence be interesting to have a complete label-free technique which does not suffer from photobleaching, while providing the same or higher selectivity as compared to fluorescence. In these respects, Raman spectroscopy is a very promising candidate.

3.4 Spontaneous Raman spectroscopy

Similar to IR absorption spectroscopy, Raman spectroscopy also probes the vibrational levels of a molecule. [10] While for IR spectroscopy one needs a light source of which the frequency exactly matches the vibrational frequencies, a spontaneous Raman spectrum is obtained by pumping the molecule to a certain virtual state Σ^* after which the molecule relaxes almost instantaneously to one of the vibrational levels of the electronic ground state (see Figure 3.5). This virtual state lies below the first excited electronic state and the excitation energy $\hbar\omega$ satisfies

$$\hbar\omega_v \ll \hbar\omega < \Delta. \quad (3.3)$$

As will be shown furtheron, Raman spectra are extremely weak and the reason for the $\hbar\omega < \Delta$ constraint is to avoid a large fluorescence background originating from the excitation of the electronic S_1 state. A Raman process is essentially a one-step process, in which the excitation is immediately followed by a de-excitation. If one photon has been taken from the laser beam to produce a Raman process, then there will always be a scattered photon as well (in contrast to fluorescence). [11] If the system relaxes back to its original state, then no energy is exchanged. This elastic scattering process (termed Rayleigh scattering) will dominate the overall emission and proper filtering has to be applied in order to separate the Rayleigh scattering from the inelastic (Raman) scattering. Raman scattering can be subdivided into Stokes and anti-Stokes scattering. During a Stokes transition, the molecule relaxes back to a vibrationally excited S_0 state such that the excitation beam transfers part of its energy to the molecule. The emitted light will be redshifted and has a frequency of $\omega_S = \omega - \omega_v$ (for the transition depicted in Figure 3.5). For an anti-Stokes transition, the molecule is already in a vibrationally excited S_0 state and relaxes to a lower lying energy state. As a result, the emitted light will have a higher frequency $\omega_{AS} = \omega + \omega_v$ (for the transition depicted in Figure 3.5) and is hence blueshifted. The majority of the molecules will be in the electronic ground state (due to Boltzmann statistics). Therefore most

Raman events will be Stokes scattering events. The ratio of the Stokes and anti-Stokes intensities depends on the number of molecules in the ground and vibrationally excited states and can be calculated using the Boltzmann distribution

$$\frac{\exp\left(-\left(\frac{\hbar\omega_v}{2kT}\right)\right)}{\exp\left(-\left(\frac{\hbar\omega_v\left(1+\frac{1}{2}\right)}{kT}\right)\right)} = \exp\left(\frac{\hbar\omega_v}{kT}\right) \quad (3.4)$$

assuming non-degenerate levels and a Stokes transition which results from a ground state excitation while the anti-Stokes results from a $v = 1$ excitation. [1] For typical vibrational frequencies (10 – 100 THz), the ratio $\exp\left(\frac{\hbar\omega_v}{kT}\right) \gg 1$ at room temperature. So usually Raman scattering is recorded at the Stokes side. However, sometimes anti-Stokes scattering is preferred in order to mitigate fluorescence interference since fluorescence is always redshifted with respect to the excitation while anti-Stokes emission is at the blueshifted side.

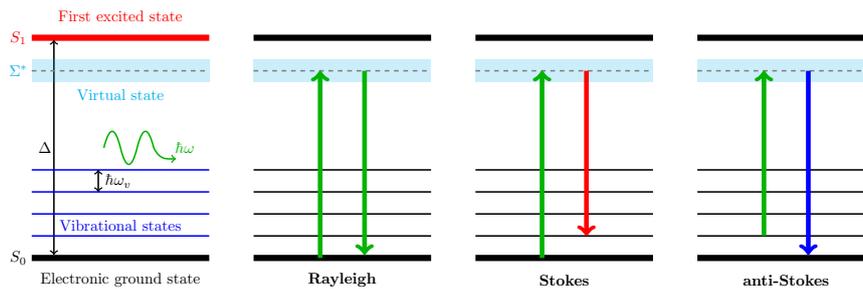


Figure 3.5: Energy level diagram for Raman spectroscopy. The molecule is pumped to a virtual state and instantaneously relaxes to one of the vibrational levels of the electronic ground state.

In contrast to IR absorption spectroscopy, where a change in dipole moment was necessary for a vibration mode to be IR active, Raman active vibrations are associated to a change in the polarizability as will be outlined below. The frequency-dependent linear induced electric dipole vector of a molecule is given by

$$\mathbf{p} = \boldsymbol{\alpha} \cdot \mathbf{E} \quad (3.5)$$

where \mathbf{E} is the electric field vector of the incident, plane wave, monochromatic radiation of frequency ω and $\boldsymbol{\alpha}$ is the polarizability tensor of the molecule. We will consider the scattering of a molecule which is free

to vibrate but does not rotate. The variation of the polarizability with vibrations of the molecule can be expressed using a Taylor series with respect to the normal coordinates of vibration. Maintaining only the first order terms (i.e. the electrical harmonic approximation) and assuming a scalar polarizability for simplicity, we get

$$\alpha = \alpha_0 + \sum_k \left(\frac{\partial \alpha}{\partial Q_k} \right)_0 Q_k \quad (3.6)$$

where α_0 is the equilibrium polarizability and Q_k are the normal coordinates of vibration associated with the molecular vibrational frequencies ω_k . If we focus on one particular normal mode of vibration Q_k then we eventually get

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial Q_k} \right)_0 Q_k = \alpha_0 + \dot{\alpha}_k Q_k. \quad (3.7)$$

Assuming a harmonic motion for the normal coordinates, the time dependence of Q_k is given by

$$Q_k = Q_{k0} \cos(\omega_k t + \delta_k). \quad (3.8)$$

For the monochromatic field \mathbf{E} we can write

$$\mathbf{E} = \mathbf{E}_0 \cos \omega t \quad (3.9)$$

such that the induced dipole moment eventually equals

$$\begin{aligned} \mathbf{p} &= \alpha_0 \mathbf{E}_0 \cos \omega t + \dot{\alpha}_k \mathbf{E}_0 Q_{k0} \cos(\omega_k t + \delta_k) \cos \omega t \\ &= \underbrace{\alpha_0 \mathbf{E}_0 \cos \omega t}_{\text{Rayleigh}} \\ &\quad + \frac{\dot{\alpha}_k \mathbf{E}_0 Q_{k0}}{2} \left(\underbrace{\cos((\omega + \omega_k)t + \delta_k)}_{\text{anti-Stokes}} + \underbrace{\cos((\omega - \omega_k)t - \delta_k)}_{\text{Stokes}} \right). \end{aligned} \quad (3.10)$$

Rayleigh scattering originates from the electric dipole oscillating at ω and induced in the molecule by the external electric field, which in itself oscillates at ω . Raman scattering arises from the electric dipoles oscillating at $\omega \pm \omega_k$ which are produced when the electric dipole oscillating at ω is modulated by the molecule oscillating at ω_k . [3]

The average radiated power by an oscillating electric dipole, induced in a molecule by an electric field of frequency ω , is given by

$$P_d = \frac{|\mathbf{p}_d|^2 \omega_d^4}{12\pi\epsilon_0 c^3} \quad (3.11)$$

where $|\mathbf{p}_d|$ is the amplitude of the induced electric dipole and ω_d the frequency of the dipole radiation. Generally ω_d will differ from ω (e.g. for Raman scattering $\omega_d = \omega \pm \omega_k$). [3, 13] For Stokes or anti-Stokes radiation, $\mathbf{p}_d = \frac{\dot{\alpha}_k \mathbf{E}_0 Q_{k0}}{2}$ such that Raman active transitions need to satisfy

$$\dot{\alpha}_k = \left(\frac{\partial \alpha}{\partial Q_k} \right)_0 \neq 0. \quad (3.12)$$

This means that a Raman active vibration mode needs to have a non-zero gradient of the polarizability. Returning to Figure 3.2, the IR inactive modes will now be Raman active since both vibrations will generate a change in the polarizability. On the other hand, the ABA antisymmetric stretching and bending mode have a zero polarizability derivative at equilibrium, which makes these modes Raman inactive. [3]

The strength of a Raman transition can also be expressed by means of a Raman scattering cross-section σ (in m^2 per steradian per molecule) which relates the scattered intensity I_s to the irradiance \mathfrak{I} of the incident radiation,

$$I_s = \sigma \mathfrak{I}. \quad (3.13)$$

The irradiance can be written as a function of the amplitude $|\mathbf{E}_0|$ of the incident radiation $\mathfrak{I} = \frac{1}{2} c \epsilon_0 |\mathbf{E}_0|^2$, while the intensity of the scattered radiation equals

$$\begin{aligned} I_s &= \frac{\pi^2 c}{2 \epsilon_0 \lambda_d^4} |\mathbf{p}_d|^2 \\ &= \frac{\pi^2 c}{2 \epsilon_0 \lambda_d^4} \alpha_{tr}^2 |\mathbf{E}_0|^2 \end{aligned} \quad (3.14)$$

in which α_{tr} is a general transition polarizability α_{tr} and λ_d the emission wavelength of the scattered radiation. Substituting the irradiance in the former equation yields

$$\begin{aligned} I_s &= \frac{\pi^2 c}{2 \epsilon_0 \lambda_d^4} \alpha_{tr}^2 |\mathbf{E}_0|^2 \\ &= \frac{\pi^2 c}{2 \epsilon_0 \lambda_d^4} \alpha_{tr}^2 \frac{2}{c \epsilon_0} \mathfrak{I} \\ &= \frac{\pi^2 \alpha_{tr}^2}{\epsilon_0^2 \lambda_d^4} \mathfrak{I} = \sigma \mathfrak{I} \end{aligned} \quad (3.15)$$

such that the Raman cross-section is given by

$$\sigma = \frac{\pi^2 \alpha_{tr}^2}{\epsilon_0^2 \lambda_d^4}. \quad (3.16)$$

Usually such a cross-section has an extremely low value on the order of 10^{-35} m^2 per steradian per molecule (depending on the specific emission wavelength λ_d). [3] If N_s is the number of scattered photons in an element solid angle $d\Omega$ falling on the detector in unit time (i.e. the photon rate), then the scattered intensity also equals

$$I_s = \frac{N_s h c}{\lambda_d d\Omega} \quad (3.17)$$

The photon rate is then finally given by

$$N_s = \sigma d\Omega \frac{\mathcal{J} \lambda_d}{h c}. \quad (3.18)$$

It is instructive to calculate this photon rate for a realistic experimental situation in order to assess the expected strength of a typical Raman signal. Assume a laser beam with power 10 mW which is focused to a spot with area 10^{-12} m^2 (corresponding to a spot diameter of $1 \mu\text{m}$). The incident irradiance is then 10^{10} W/m^2 . For a single molecule with a cross-section of 10^{-35} m^2 per steradian per molecule emitting at $\lambda_d = 877 \text{ nm}$ (and $d\Omega = 1$), the photon rate $N_s = 44 \times 10^{-8}$ photons per second. So in order to generate 1 Raman photon per second, one would need about 2 million molecules. [3, 14] In this particular case, a 10 mW beam would produce a Raman signal of about $2.3 \times 10^{-4} \text{ fW}$. Due to the extreme weak nature of Raman signals, it is desirable to enhance them. Plasmonic structures, as introduced in the previous chapter, are ideal candidates for this purpose. This so-called Surface Enhanced Raman Spectroscopy (SERS) will be discussed in the next section.

3.5 Surface Enhanced Raman Spectroscopy

The field enhancement generated by metallic nanostructures has been extensively discussed in the previous chapter. Surface Enhanced Raman Spectroscopy (SERS) uses this effect to enhance Raman signals by several orders of magnitude. We will first give a qualitative treatment of the enhancement effects associated to molecules in the vicinity of a metal

surface and highlight the most important features that influence the field enhancement. In second instance, we will elaborate on a quantitative analysis of Raman enhancement and fluorescence quenching near metal nanoparticles.

3.5.1 Qualitative treatment of SERS

Assume a molecule is brought in the vicinity of a metallic nanoparticle which is resonant in the proper wavelength region. The enhanced field around the nanoparticle will excite the molecule such that its induced dipole moment is given by

$$\mathbf{p} = \alpha_R EF(\lambda_P) \mathbf{E}_0 \quad (3.19)$$

in which α_R is a general polarizability, $EF(\lambda_P)$ the field enhancement at the excitation (pump) wavelength λ_P (see Figure 3.6) and \mathbf{E}_0 the incident electric field. This induced dipole will generate an electric field \mathbf{E}_{Stokes} at a given Stokes wavelength λ_S , which in itself can also be enhanced by the nanoparticle

$$\mathbf{E}_{Stokes} \propto EF(\lambda_S) (\alpha_R EF(\lambda_P) \mathbf{E}_0), \quad (3.20)$$

where $EF(\lambda_S)$ is the field enhancement at the Stokes wavelength (see Figure 3.6). A similar argument holds for the anti-Stokes process. The overall intensity I_{SERS} of the scattered SERS light will then scale according to

$$|\mathbf{E}_{Stokes}|^2 \propto I_{SERS} \propto \alpha_R^2 EF(\lambda_P)^2 EF(\lambda_S)^2 I_0 \quad (3.21)$$

where I_0 is the incident intensity. The SERS enhancement factor EF , defined as the ratio of the Raman scattered intensity in the presence of the metal nanoparticle to its value in the absence of the particle, is then

$$EF = \left(\frac{\alpha_R}{\alpha_{R0}} \right)^2 EF(\lambda_P)^2 EF(\lambda_S)^2 \quad (3.22)$$

which takes into account the fact that the polarizability of the isolated molecule α_{R0} can change because of (chemical) interactions with the metal nanoparticle. For Raman modes with Stokes wavelengths close to the pump wavelength (i.e. low wavenumber modes), the SERS intensity scales with the fourth power of the field enhancement. Despite this non-linear scaling, SERS in itself is a linear optical effect which scales with the first power of I_0 . Moreover it is important to note that α_R is the polarizability of the molecule when the molecule is adsorbed on the metal surface. Hence it can contain contributions from the metal and may be greatly

altered in magnitude. This effect can significantly contribute to the so-called chemical enhancement of SERS. The electromagnetic enhancement of SERS on the other hand is included in the fourth power of the field enhancement. [15] Alternatively the SERS enhancement can be written in terms of the Local Field Intensity Enhancement Factor $LFIEF(\lambda) = EF(\lambda)^2$:

$$EF = \left(\frac{\alpha_R}{\alpha_{R0}} \right)^2 LFIEF(\lambda_P) LFIEF(\lambda_S). \quad (3.23)$$

Any optical technique that depends on the intensity of the light at a specific point will hence be linked to the $LFIEF$ and, in general, depending on whether the $LFIEF$ is > 1 or < 1 the optical process will be enhanced or quenched. [11]

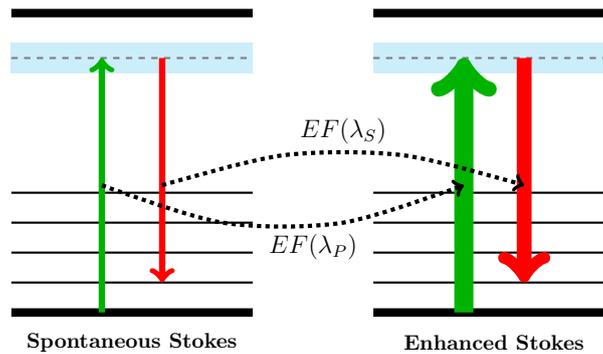


Figure 3.6: Schematic of Surface Enhanced Raman Spectroscopy. The excitation field (with pump wavelength λ_P) and emitted field (with Stokes wavelength $\lambda_S > \lambda_P$) are enhanced by $EF(\lambda_P)$ and $EF(\lambda_S)$ respectively.

3.5.2 Size and shape effects on field enhancement

Figure 3.7 depicts the extinction spectra of several silver nanoprisms and nanodiscs. Extinction spectra represent the sum of absorption and scattering losses due to the metallic nanoparticle and can be measured by simple transmission spectrometry. [17] One can see that the resonance wavelength can be tuned over a broad wavelength range by changing the size and shape of the metallic nanostructure. [16, 17] As a result, the previously introduced enhancement factor EF will also be strongly affected by these geometrical quantities.

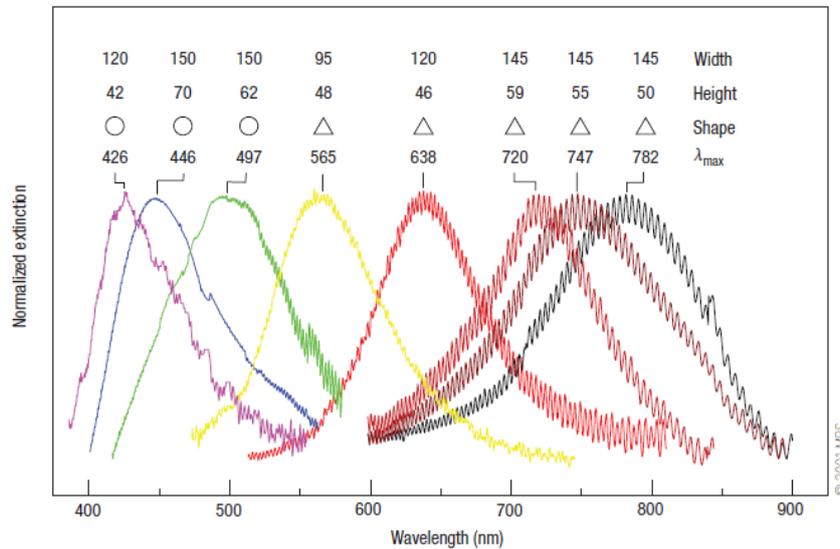


Figure 3.7: Effect of size and shape on the plasmon resonance of silver nanoparticles and nanodiscs. Taken from [17].

In the previous chapter we meticulously studied the plasmon resonances of a spherical nanoparticle ($\ll \lambda$) in the electrostatic approximation. Within this approximation the resonance wavelength is size independent and only depends on the dielectric environment and the mode order l . However, when the size of a plasmonic structure is a substantial fraction of the wavelength, size effects cannot be ignored anymore and the electrostatic approximation no longer holds. The study of size effects on the plasmon resonance is often very complicated and mostly relies on the numerical solution of Maxwell's equations. In a later chapter, such geometrical effects will be studied in more detail using FDTD simulations. Qualitatively one can however summarize size effects as follows [11]:

- The resonance wavelength redshifts if the size increases.
- Resonances are strongly damped as the size increases, mostly because of increased radiation losses. This results in a broadening of the resonance, and more importantly in a decrease of the associated field enhancement (see also previous chapter).

- Another typical consequence of size is the appearance of resonances that do not exist in the small size limit (where the electrostatic approximation was sufficient). These size-related resonances, for a fixed shape, are typically the result of the activation of multipolar resonances that do not couple to light very efficiently in the limit of small sizes.

The example of the spherical particle in the previous chapter clearly introduced the concept of geometry-induced resonance, because the resonance appears at a wavelength where there is no intrinsic feature (or peak) in the bulk dielectric function of the materials themselves. In other words, these resonances appear as purely geometrical aspects of the problem. It would therefore be interesting to know what happens with other geometries than spheres, i.e. how shape can affect the resonance. Unfortunately, the shape influence is usually only obtainable through numerical simulations. In order to highlight some important features that result from shape changes, we consider a triangular nanoparticle as shown in Figure 3.8. [11]

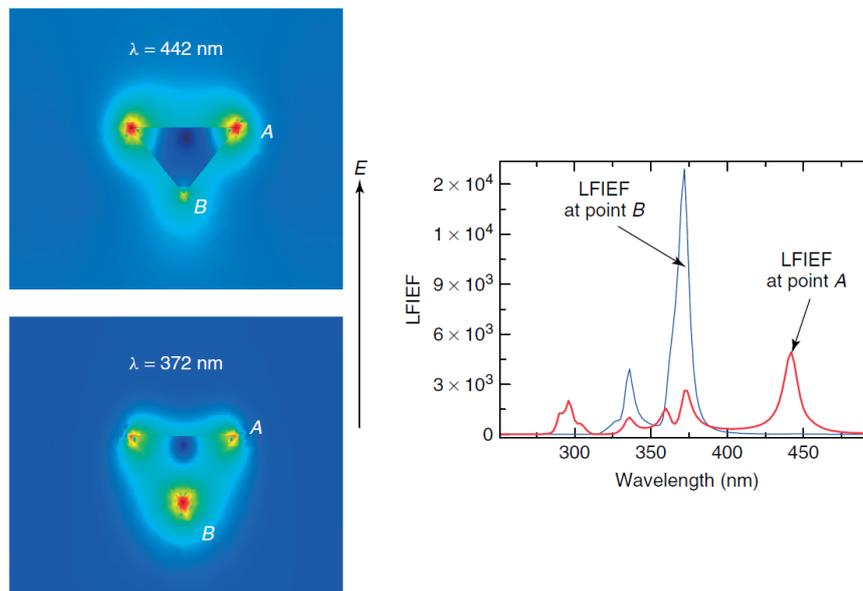


Figure 3.8: LFIEF at two different wavelengths for a triangular silver shape in the electrostatic approximation. The incident electric field \mathbf{E} is polarized vertically as shown. Taken from [11].

Some important conclusions can be drawn from this Figure [11]:

- In general, there will be more than one resonance condition associated with a given shape. Some of these resonances have complicated spatial enhancement distributions.
- Different points on the surface can have their maximum *LFIEFs* at different wavelengths. The *LFIEF* is strongly position dependent in most cases.
- The maxima of the *LFIEFs* shown in Figure 3.8 are examples of what is normally called the lightning rod effect (i.e. electric fields concentrating at sharp ends). The *LFIEF* can hence be very high for particular shapes, especially near sharp corners.
- In general, the resonances (in their spectral position and intensity) will not only depend on the shape but also on the orientation with respect to the excitation field. If the direction of the electric field in Figure 3.8 is changed while keeping the triangular shape in the same position, the *LFIEFs* at points A and B will change accordingly. So even when the triangular cross-section is equilateral, the electric field breaks the symmetry and points A and B are no longer equivalent. In the full (non-electrostatic) solution of the problem, there is also the direction imposed by the incident wavevector \mathbf{k} . While the latter has in general less importance than the direction of \mathbf{E} (in particular, for really small objects), there are many subtle details of the field distribution that do depend on it.

The strength of the *LFIEF* will furthermore depend on the type of metal used, and more specifically on the real and imaginary part of the dielectric permittivity, which determine the *Q*-factor. In Figure 3.9 the *LFIEF* of a silver and gold spherical particle is compared. The resonance wavelength is determined by $\text{Re}(\epsilon_M(\lambda)) = -2\text{Re}(\epsilon_D(\lambda))$, while the imaginary part of the metal dielectric function $\text{Im}(\epsilon_M(\lambda))$ determines the width and strength of the resonance. Silver has a stronger and narrower *LFIEF* as compared to gold because $\text{Re}(\epsilon_M(\lambda)) = -2\text{Re}(\epsilon_D(\lambda))$ is satisfied at a wavelength λ where the imaginary part of silver is much smaller (comparatively speaking) than that of gold at its corresponding resonance wavelength. This makes the resonance in gold more lossy and hence broader (lower *Q*-factor). [11] In the previous chapter we have indeed shown that the field enhancement is related to the *Q*-factor and the

exact relationship will depend on the configuration of the nanoparticles. In Figure 3.10 the quality factor of a metal/air nanoparticle for different metals is shown. [18] Below 650 nm, silver will clearly have a higher field enhancement (due to the larger Q -factor) as compared to gold. In the near-IR region (700 – 1000 nm) there is not much difference between the quality factors of silver and gold so both materials will exhibit similar field enhancement. Nevertheless, for biosensing applications gold is more preferred because of its chemical stability (silver can e.g. oxidize).

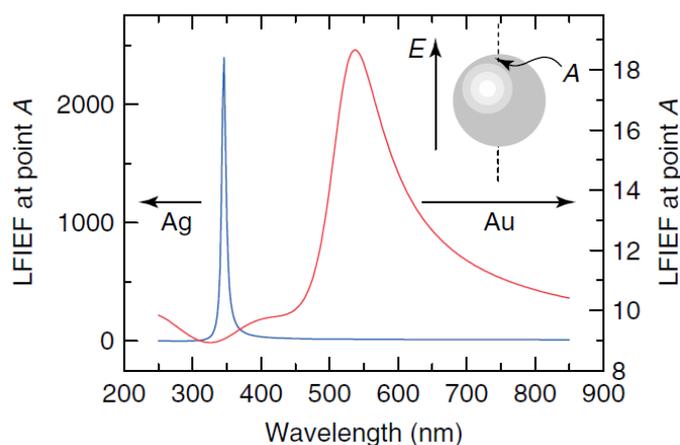


Figure 3.9: LFIEF at point A (inset) on a sphere of either gold (Au) or silver (Ag) in the electrostatic approximation. Taken from [11].

3.5.3 Fluorescence quenching versus Raman enhancement

Previously, we addressed the potential issue of fluorescence interference when acquiring Raman spectra. Fluorescence signals are typically much stronger as compared to spontaneous Raman signals such that the Raman signal can easily be drowned in a large fluorescence background. Fluorescence can be (partially) avoided by choosing a pump frequency which is much lower than the transition frequency to the first electronic state. By introducing a metallic nanoparticle near the analyte under study, there is however an additional channel by which fluorescence

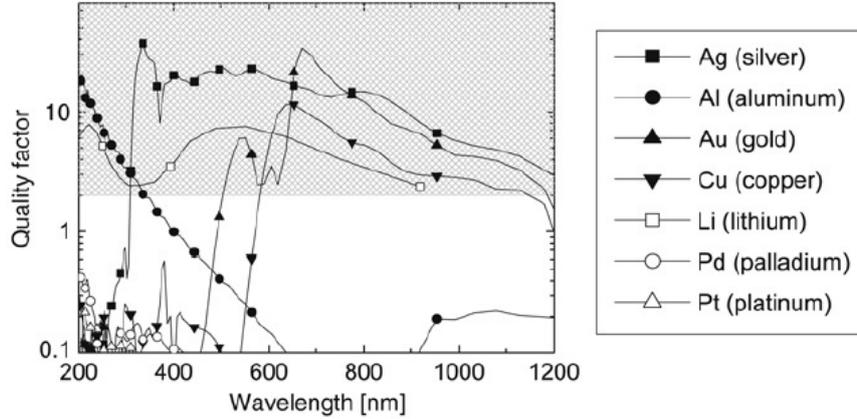


Figure 3.10: Quality factor of a metal/air nanoparticle for different metals. The shaded area is the area of interest for many plasmonic applications. Taken from [18].

can be mitigated while preserving the full enhancement of the Raman process. This simultaneous effect of fluorescence quenching and SERS will be quantitatively discussed in this section. Additionally, this analysis also provides a more thorough and quantitative understanding of the electromagnetic SERS enhancement.

For the theoretical analysis, assume a molecule at distance d from a spherical metallic nanoparticle with radius R (see Figure 3.11). Among all the plasmon modes of a single sphere, the $l = 1$ dipole mode (with resonance frequency ω_1) is the only one with non-zero dipole moment and hence the only one which can couple to external fields and decay radiatively with a decay rate γ_{rad} . All modes will furthermore decay nonradiatively at a damping rate γ due to metal losses. At first instance the fluorescence process will be considered. The energy level diagram and corresponding decay rates are depicted in the top part of Figure 3.11. A pump beam with frequency ω_P excites the molecule to a vibrational level $|3\rangle$ of the first excited electronic state. This level undergoes a quick vibrational relaxation at a rate γ_{32} to a level $|2\rangle$ (which typically has a relatively long lifetime). From level $|2\rangle$ the molecule can decay radiatively with an original radiative efficiency η_{rad} which is determined by the radiative $\gamma_{rad,2}$ and nonradiative $\gamma_{nrad,2}$ decay rate,

$$\eta_{rad} = \frac{\gamma_{rad,2}}{\gamma_{rad,2} + \gamma_{nrad,2}}. \quad (3.24)$$

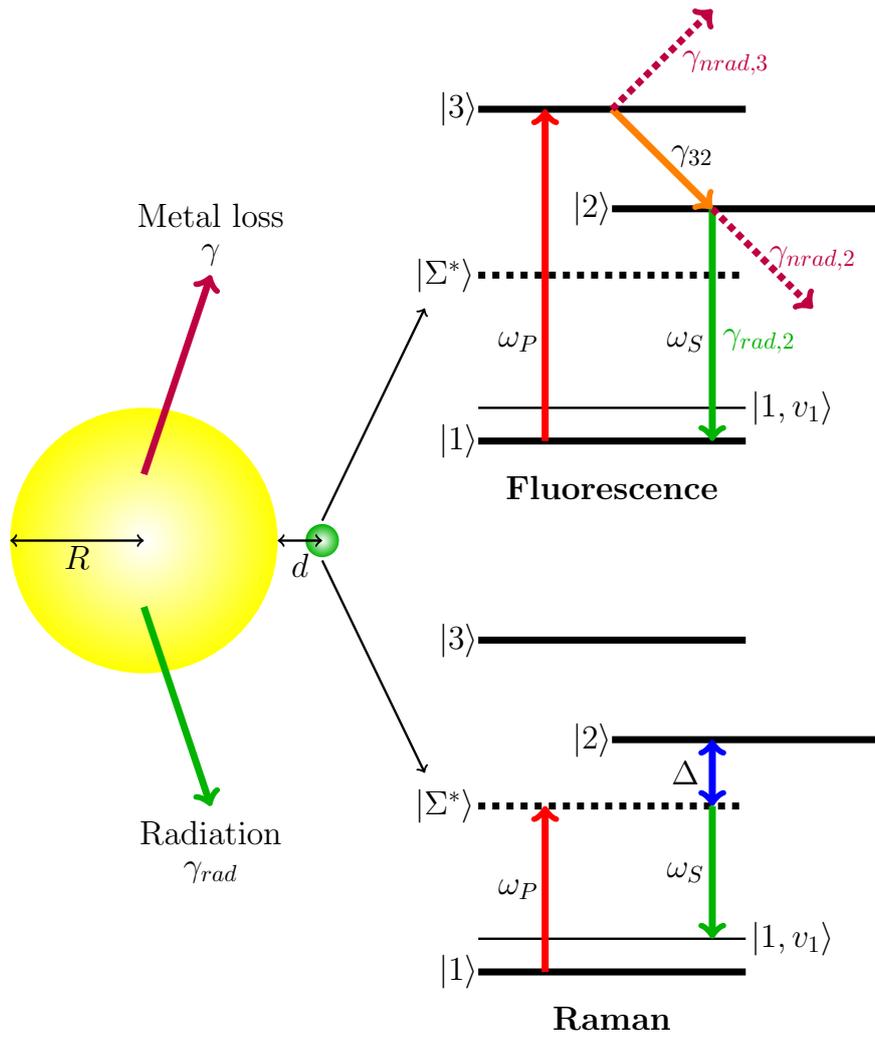


Figure 3.11: Energy level diagram for fluorescence and Raman emission near a metallic nanoparticle. Adapted from [19].

The fluorescence enhancement process is effectively a multistep process as depicted in Figure 3.12(a). [19] In the first step, the incident optical excitation at frequency ω_P couples into the dipolar $l = 1$ mode because all higher order modes have vanishing dipole moments and are hence not coupled with external fields. In the second step, the molecule makes a transition from the ground state $|1\rangle$ to the excited state $|3\rangle$, absorbing energy from the dipole mode with rate

$$\gamma_{abs} = \frac{c\sigma_{abs}}{V_{eff,1}} \left(\frac{R}{R+d} \right)^6, \quad (3.25)$$

where σ_{abs} is the original absorption cross section (at ω_P) and $V_{eff,1} = \pi R^3$ the effective dipole mode volume. Relative to the case in which no metal sphere is present and the molecule is situated in the diffraction-limited focal spot of a Gaussian beam with an electric field E_{foc} , the maximum enhancement during this optical absorption process (proportional to the square of the electric field) is given by

$$\begin{aligned} F_{abs}(\omega_P) &= \left| \frac{E_{max,1}}{E_{foc}} \right|^2 \left(\frac{R}{R+d} \right)^6 \\ &= \frac{2}{\left(\frac{\gamma_{abs}}{\omega_P} + \frac{\gamma_{rad} + \gamma}{\omega_P} \right)^2 + \delta^2} \left(\frac{R}{R+d} \right)^6 \end{aligned} \quad (3.26)$$

where $E_{max,1}$ is the maximum electric field of the dipole mode and $\delta = 2 \left(1 - \frac{\omega_1}{\omega_P} \right)$ the excitation detuning. [19] As opposed to the previous chapter, there is an additional channel by which the dipole mode can decay (see Figure 3.12(a)). Apart from the radiative γ_{rad} and nonradiative γ decay, part of the energy can now be transferred to the molecule itself (at a rate γ_{abs}). This explains the additional factor $\frac{\gamma_{abs}}{\omega_P}$ in the denominator. After absorption, the molecule quickly relaxes nonradiatively to level $|2\rangle$. All the energy stored in level $|2\rangle$ can now decay into multiple channels (see Figure 3.12(a)). First there is a nonradiative decay with rate $\gamma_{nrad,2}$. The original radiative decay rate $\gamma_{rad,2}$ (at the Stokes frequency ω_S) is enhanced by the Purcell factor $F_{P,l}$ of each of the corresponding plasmon modes ($l = 1 \dots \infty$), effectively changing the decay rate to $F_{P,l}\gamma_{rad,2}$ into each of these modes. The Purcell factor $F_{P,l}(\omega)$ itself is given by [19]

$$F_{P,l}(\omega) = \frac{3\pi\epsilon_D\omega}{8\chi^3} \frac{(2l+1)(l+1)^2}{(l+1)\epsilon_D+l} \left(\frac{\chi}{\chi+\chi_d} \right)^{2l+4} L_l(\omega) \quad (3.27)$$

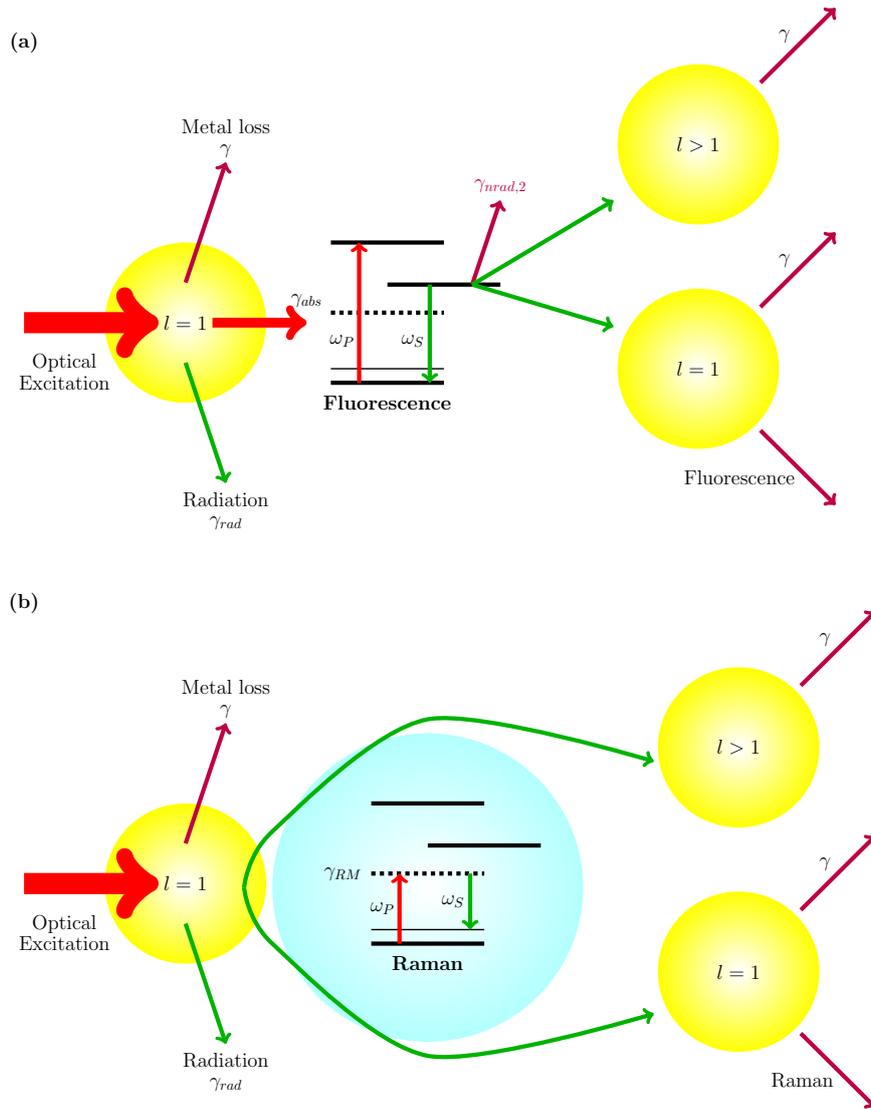


Figure 3.12: Decay channels of the optical excitation and subsequent fluorescence (a) and Raman (b) emission with corresponding decay rates. Adapted from [19].

where $\chi = \omega R/c$, $\chi_d = \omega d/c$. $L_l(\omega)$ is the normalized Lorentzian linewidth of the l th mode:

$$L_l(\omega) = \frac{\frac{\gamma_l}{2\pi}}{(\omega - \omega_l)^2 + \frac{\gamma_l^2}{4}}, \quad (3.28)$$

where the resonance frequency of the l th mode is ω_l and its decay rate is $\gamma_l = \gamma + \gamma_{rad}\delta_{l,1}$. Among all the modes, only the dipole mode will subsequently couple to external radiation modes with out-coupling efficiency $\eta_{dp} = \frac{\gamma_{rad}}{\gamma_{rad} + \gamma}$ and at a rate γ_{rad} . All the higher order ($l \geq 2$) modes dissipate energy in a nonradiative way. The large $F_{P,l}(\omega)$ of the higher order modes indicates that energy can be coupled rather efficiently into them for sufficiently small d . Because these modes are uncoupled to external fields, they will however not contribute to fluorescence emission. This phenomenon can be expressed by the fluorescence quenching ratio

$$f_q(\omega) = \frac{\sum_{l=2}^{\infty} F_{P,l}(\omega)}{F_{P,1}(\omega)}. \quad (3.29)$$

The efficiency by which the molecule then radiates in the presence of a metallic nanoparticle is given by

$$\eta_{MN}(\omega_S) = \frac{\gamma_{rad,2} + F_{P,1}(\omega_S) \gamma_{rad,2} \eta_{dp}}{\gamma_{rad,2} + \gamma_{nrad,2} + F_{P,1}(\omega_S) (1 + f_q(\omega_S)) \gamma_{rad,2}} \quad (3.30)$$

such that the overall emission enhancement eventually equals

$$F_{em}(\omega_S) = \frac{\eta_{MN}(\omega_S)}{\eta_{rad}} = \frac{1 + F_{P,1}(\omega_S) \eta_{dp}}{1 + F_{P,1}(\omega_S) (1 + f_q(\omega_S)) \eta_{rad}}. \quad (3.31)$$

This enhancement factor not only strongly depends on the radius R but also on the separation distance d . The optimal size R_{opt} will occur at a radius where the nanoparticle is small enough to yield a small effective mode volume (for an enhanced Purcell factor), yet is still sufficiently large to assure strong radiative out-coupling. If higher order modes could be neglected, it is always better to position the molecule as close as possible to the metal in order to take advantage of the strong Purcell factor of the dipole mode. However with higher order modes, the energy of a molecule which is too close to the metal surface will also couple to nonradiative (dark) modes and hence dissipates as metal loss. As a result, an optimized separation d_{opt} will exist that allows significant coupling into the dipole mode while adequately suppressing the fluorescence quenching by higher

order modes. [20] The overall fluorescence enhancement $F_{FL}(\omega_P, \omega_S)$ is then a combination of the absorption and emission enhancement,

$$F_{FL}(\omega_P, \omega_S) = F_{abs}(\omega_P)F_{em}(\omega_S). \quad (3.32)$$

This fluorescence enhancement will furtheron be compared with the Raman enhancement which we will now discuss in more detail. The same configuration of a molecule at distance d from a metallic sphere with radius R is considered. The energy diagram for the Raman process is shown in the bottom part of Figure 3.11. The pump beam (at ω_P) excites the molecule to a virtual level $|\Sigma^*\rangle$ which lies below the previously considered levels $|2\rangle$ and $|3\rangle$ of the first excited electronic state. Almost instantaneously the molecule relaxes to one of the vibrational states $|1, v_1\rangle$ of the electronic ground state $|1\rangle$. As long as the energy difference Δ is sufficiently large, i.e. as long as the excitation frequency is far from any absorption resonance ω_m of the molecule, the process is a pure spontaneous Raman scattering process. The overall rate of spontaneous Raman scattering equals

$$\gamma_{RM} = \frac{2\pi}{\hbar} \frac{|d_m E_{ex}|^2}{\hbar^2 (\omega_P - \omega_m)^2} \frac{|H_{ev}|^2}{\hbar^2 (\omega_S - \omega_m)^2} \frac{|d_m|^2 \omega_S}{\epsilon_0} \rho(\omega_S), \quad (3.33)$$

where d_m is the dipole moment of the molecular transition, E_{ex} the electric excitation field experienced by the molecule, H_{ev} the strength of the electron-vibrational coupling responsible for the Raman process and $\rho(\omega_S)$ the effective density of states (DOS) at the Stokes frequency. [19, 21] In the absence of a metallic nanoparticle, $E_{ex} = E_{foc}$ and $\rho(\omega_S)$ is the DOS of the radiation mode $\rho_{rad}(\omega_S)$. If the nanoparticle is introduced, the excitation field gets enhanced by the previously introduced field enhancement F_S to $E_{ex} = E_{foc} \rightarrow F_S E_{foc}$. Moreover it will induce a change in the DOS, which can now be written using the DOS for each plasmon mode

$$\rho(\omega_S) \approx \rho_{rad}(\omega_S) + \sum_{l=1}^{\infty} \rho_l(\omega_S). \quad (3.34)$$

From Figure 3.12(b) one can see that in the absence of a real molecular level, the energy which is stored in the dipolar $l = 1$ mode (due to optical excitation) now directly splits in three different path ways: radiative decay of the $l = 1$ mode, metal loss and Raman scattering into multiple plasmon modes at the Stokes frequency. The decay into the higher order plasmon modes is again lost as metal loss, while only the coupling to the $l = 1$ mode

will result in far field Raman scattering. The modified Raman decay rate is then

$$\begin{aligned}
\gamma_{RM}^{MN} &= \left(\frac{2\pi}{\hbar} \frac{|d_m F_S E_{foc}|^2}{\hbar^2 (\omega_P - \omega_m)^2} \frac{|H_{ev}|^2}{\hbar^2 (\omega_S - \omega_m)^2} \frac{|d_m|^2 \omega_S}{\epsilon_0} \right) \times \\
&\quad \left(\rho_{rad}(\omega_S) + \sum_{l=1}^{\infty} \rho_l(\omega_S) \right) \\
&= F_S^2 \left(1 + \sum_{l=1}^{\infty} \frac{\rho_l(\omega_S)}{\rho_{rad}(\omega_S)} \right) \times \\
&\quad \left(\frac{2\pi}{\hbar} \frac{|d_m E_{foc}|^2}{\hbar^2 (\omega_P - \omega_m)^2} \frac{|H_{ev}|^2}{\hbar^2 (\omega_S - \omega_m)^2} \frac{|d_m|^2 \omega_S}{\epsilon_0} \rho_{rad}(\omega_S) \right) \\
&= F_S^2 \left(1 + \sum_{l=1}^{\infty} F_{P,l}(\omega_S) \right) \gamma_{RM} \tag{3.35}
\end{aligned}$$

since the Purcell factor $F_{P,l}(\omega_S)$ can be estimated as the ratio of the effective density of the l th modes $\rho_l(\omega_S)$ to that of the radiation continuum $\rho_{rad}(\omega_S) = \frac{\omega_S^2}{3\pi^2 c^3}$. [20] Since only the coupling to the $l = 1$ mode will result in far field Raman scattering with a dipole out-coupling efficiency η_{dp} , the radiative Raman rate will be

$$\gamma_{RM,rad}^{MN} = F_S^2 (1 + F_{P,1}(\omega_S) \eta_{dp}) \gamma_{RM} \tag{3.36}$$

such that the overall Raman enhancement is then given by

$$\begin{aligned}
F_{RM} &= \frac{\gamma_{RM,rad}^{MN}}{\gamma_{RM}} = F_S^2 (1 + F_{P,1}(\omega_S) \eta_{dp}) \\
&= \left| \frac{E_{max,1}}{E_{foc}} \right|^2 \left(\frac{R}{R+d} \right)^6 (1 + F_{P,1}(\omega_S) \eta_{dp}). \tag{3.37}
\end{aligned}$$

In most cases, the Raman decay related to the $l = 1$ mode is much stronger than that of the radiation mode, i.e. $F_{P,1}(\omega_S) \eta_{dp} \gg 1$, hence $F_{RM} \propto \left| \frac{E_{max,1}}{E_{foc}} \right|^2 F_{P,1}(\omega_S)$. Because the Purcell factor of the dipole mode is proportional to the square of the field enhancement at the Stokes frequency, one obtains the E^4 dependence of the Raman enhancement (as in the qualitative treatment before). The main difference between F_{FL} and F_{RM} is the absence of the quenching denominator $1 + F_{P,1}(\omega_S) (1 + f_q(\omega_S)) \eta_{rad}$ in the formula for F_{RM} . In a non-resonant Raman process (where Δ is sufficiently large such that one is far from

a molecular resonance) the Raman cross-section is typically many orders of magnitude weaker than the absorption cross-section σ_{abs} . As a result, the non-resonant Raman enhancement F_{RM} can be obtained from the fluorescence enhancement in the limit of $\sigma_{abs} \rightarrow 0$ and $\eta_{rad} \rightarrow 0$ (this can be seen by comparing formulas (3.37) and (3.32)). [19]

In order to illustrate the difference between fluorescence and Raman enhancement, we consider the case of a molecule with $\sigma_{abs} = 0.1 \text{ nm}^2$ near a gold sphere with radius $R = 30 \text{ nm}$ for different separations d and radiative efficiencies η_{rad} (see Figure 3.13). The excitation for both processes is at the dipole mode resonance $\omega_P = \omega_1 = 2.562 \text{ eV}$, while fluorescence and Stokes emission are both at $\omega_S = 2.462 \text{ eV}$. [19]

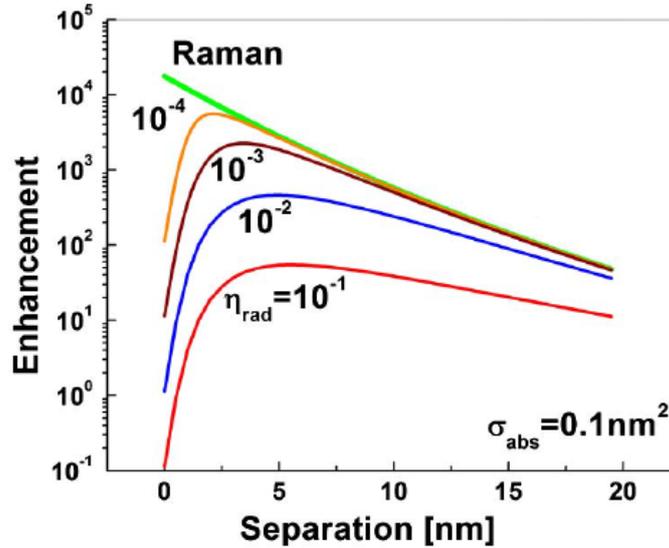


Figure 3.13: Comparison between Raman and fluorescence enhancement of a molecule with $\sigma_{abs} = 0.1 \text{ nm}^2$ near a gold sphere with radius $R = 30 \text{ nm}$ for different separations d and radiative efficiencies η_{rad} . Taken from [19].

The Raman enhancement always outperforms the fluorescence enhancement, since the latter undergoes various degrees of quenching depending on the original radiative efficiency. In the case of Raman, this quenching is absent and one can think of a Raman process as a fluorescence process with extremely low efficiency. The Raman enhancement will hence always be the upper limit of the enhancement one can obtain for fluorescence.

In the case of fluorescence there is an optimal separation d which will maximize F_{FL} , while for Raman it is desired that the molecule is placed as close as possible to the surface. [19] This immediately provides a scheme to suppress unwanted fluorescence interference. As long as the molecule is close enough to the metal surface, the Raman scattering will be enhanced while the fluorescence is quenched, effectively improving the Raman scattering Signal-to-Noise Ratio.

3.6 Conventional Raman spectroscopy setup

In this section we will discuss a typical conventional Raman spectroscopy setup (see Figure 3.14) and introduce its conversion efficiency in order to compare it later on with chip-based Raman systems. The excitation laser, at pump wavelength $\lambda_P = 2\pi c/\omega_P$, is focussed on the sample through an objective. Usually a pinhole will also be present in order to image a very small feature of the sample, in which case the system is denoted as a confocal Raman microscope. This sample can be a bare analyte, but also an analyte on top of a SERS substrate. The Raman (Stokes or anti-Stokes) radiation is then collected back through the same objective lens. The beamsplitter preferably filters out as much as possible of the pump and Rayleigh light because this light is much more intense than standard Raman scattering and can flood the detector, eventually masking the Raman spectrum. A notch filter behind the beamsplitter should assure that any remaining pump or Rayleigh light is completely filtered out before the beam is coupled to a spectrometer. Other interfering light sources, such as specular reflection from the surface of the sample, should be avoided as well. [1]

The spectrometer itself can come in a variety of configurations, however the development that is most responsible for the major increase in Raman applications is the introduction of multiplex and multichannel spectrometers (see Figure 3.14). A multichannel spectrometer monitors many wavelengths simultaneously, using many detectors in parallel (e.g. a CCD array). A common example is a grating-based dispersive spectrograph with a CCD at its focal plane. Since a multichannel system monitors all wavelengths simultaneously, it acquires a spectrum faster as compared to a scanning, single-channel system that has to monitor each wavelength in turn. In a multiplex spectrometer, the (broadband) spectrum is first sent

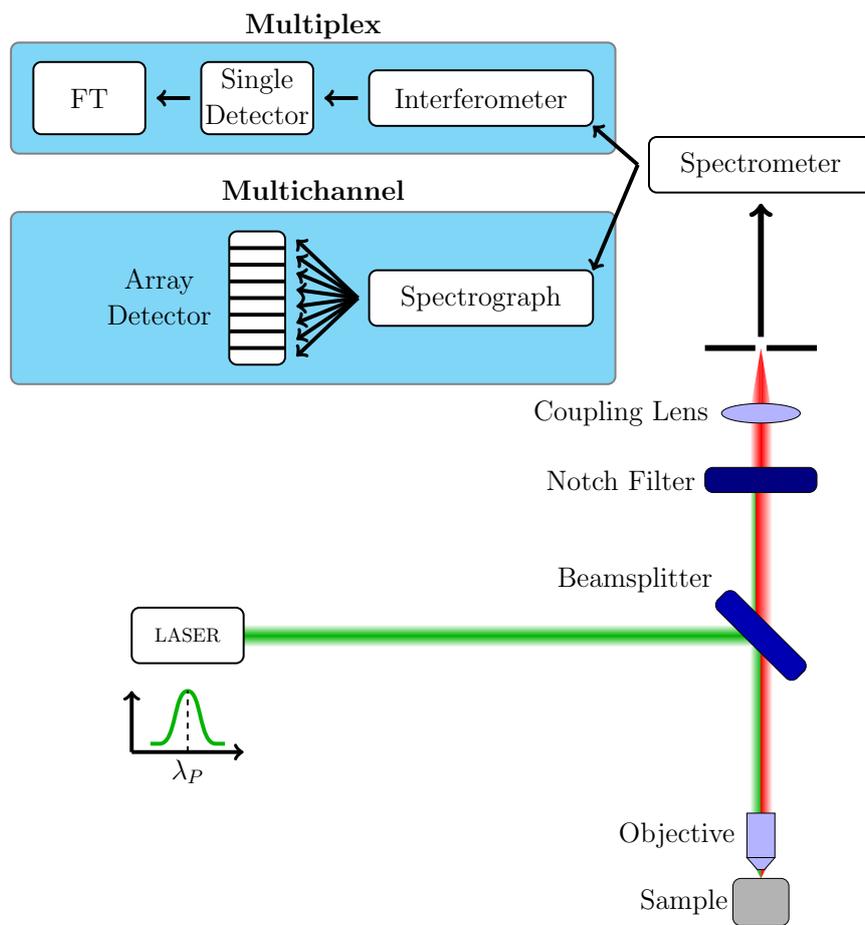


Figure 3.14: A typical Raman spectroscopy setup. The multiplex and multichannel drawings are adapted from [22].

through an interferometer (e.g a Michelson interferometer) after which the resulting signal is collected by a single photodetector. The single detector monitors the resulting signal S as a function of the optical path difference OPD created in the interferometer. Finally, the $S(OPD)$ is Fourier transformed (FT) to obtain the original spectrum. It is important to note that a multichannel spectrometer disperses the light for detection by many parallel detectors, while a multiplex spectrometer directs all of the scattered light onto a single detector. This difference between multichannel and multiplex approaches has major effects on the characteristics of the Raman spectrum, in terms of resolution, spectral coverage, signal magnitude, and Signal-to-Noise Ratio. [22] For more details we refer the reader to the book of McCreery [22].

Since the Raman scattering cross-section depends on the fourth power of the frequency, the obvious way of improving Raman sensitivity is to use the highest frequency possible, which usually means that a UV excitation source has to be used. UV excitation also has the advantage that there is less fluorescence than with visible excitation. However many compounds absorb UV radiation, which in combination with the high energy of the photons means that there is a high risk of sample degradation through burning. It also means that spectra may be rather different from normal Raman spectra due to resonance with any electronic transition that may cause absorption. This changes the relative intensities of the bands. Additionally, the required lasers can be quite expensive and the quality of the optics in the UV (detection systems, etc.) needs to be very high. Currently, most laboratories either choose visible or NIR laser excitation. For visible excitation, a dispersive spectrometer and a CCD are used for detection (such as the previously described multichannel system). NIR laser excitation typically requires an interferometer-based system and a FT program to produce the spectrum (i.e. the multiplex system). Both types have their advantages and disadvantages. For the visible, high-quality and low-cost lasers and detection systems are available, but there is always the potential issue of fluorescence (as described earlier). The main advantage of NIR excitation is that fluorescence can be avoided as much as possible. On the downside, the Raman scattering is inherently weaker because the radiated energy is lower since the fourth power law applies. However, since the exciting radiation does not absorb into most samples, the laser powers that can be used are relatively high (up to 2 Watts). Moreover, the interferometer detection system is usually very sensitive. On the other hand, detectors are typically more expensive for this wavelength region.

The major problem of visible source instruments is that the CCD chips (typically Si) lack sensitivity at wavelengths above 1000 nm. An inbetween solution is therefore the use of NIR lasers with excitation lines in the 790-850 nm wavelength region. Lasers that operate in this region will usually produce Raman signals that can still be detected by common CCD chips, however a drop in sensitivity for higher frequency peaks (larger wavenumber) is possible. [1]

From the above introductory discussion it is clear that depending on the application (and budget) one must make a careful choice of the excitation source (UV, visible or NIR) and corresponding detection unit (multichannel or multiplex). Furthermore it is interesting to calculate the amount of Raman power that can be captured by such a microscopy system for a given pump power. Later on this efficiency will be compared with the efficiency of waveguide-based Raman systems. For a diffraction limited confocal microscope, it is reasonable to assume that only particles located within the focal volume will contribute significantly to the Raman signal. In that case the conversion efficiency for a diffraction limited microscope η_{DL} , expressed as the ratio of collected Raman power P_{DL} to the pump power P_{pump} , is given by

$$\eta_{DL} = \frac{P_{DL}}{P_{pump}} = \Omega \rho \sigma b, \quad (3.38)$$

where Ω is the solid angle subtended by the optics, ρ and σ the respective concentration and Raman cross-section of the analyte under study and b the depth of focus. [23] The solid angle (corresponding to a half angle θ) can be expressed in terms of the numerical aperture NA of the objective lens and the refractive index n of the immersion medium,

$$\Omega = 2\pi (1 - \cos \theta) = 2\pi \left(1 - \sqrt{1 - \left(\frac{NA}{n} \right)^2} \right) \approx \pi \left(\frac{NA}{n} \right)^2. \quad (3.39)$$

For a Gaussian beam, the numerical aperture and depth of focus can be expressed in terms of the free-space excitation wavelength λ_P and the beam waist w_0 ,

$$NA = \frac{\lambda_P}{\pi w_0} \quad (3.40)$$

$$b = \frac{2\pi n w_0^2}{\lambda_P} \quad (3.41)$$

such that the conversion efficiency is eventually given by

$$\eta_{DL} = 2\rho\sigma\left(\frac{\lambda_P}{n}\right). \quad (3.42)$$

For a diffraction limited system it is hence impossible to improve on the conversion efficiency by changing the optics, as η_{DL} only depends on the analyte properties and the excitation wavelength. Later on it will be shown that waveguide-based Raman systems generate a Raman signal which is proportional to the length L_{wg} of the waveguide. This provides a major advantage for the overall collection of Raman light, since L_{wg}/λ_P is usually much larger than 1. [23]

3.7 Conclusion

In this chapter we outlined the rationale for conventional Raman and Surface Enhanced Raman Spectroscopy. A brief overview of IR absorption and fluorescence spectroscopy was included as well. Moreover, we discussed the mechanism of simultaneous Raman enhancement and fluorescence quenching by metallic nanoparticles. However, we only briefly touched upon the potential advantages of chip-based Raman platforms. An extensive discussion on this matter will be the subject of the next chapter.

References

- [1] Smith, E.; Dent, G. *Modern Raman Spectroscopy - A Practical Approach*; John Wiley & Sons Ltd, Chichester, UK, 2005.
- [2] Colthup, N.B.; Daly, L.H.; Wiberley, S.E *Introduction to Infrared and Raman Spectroscopy, 2nd Edition*; Academic Press Inc, USA 1975.
- [3] Long, D.A. *The Raman Effect*; John Wiley & Sons Ltd, Chichester, UK, 2002.
- [4] Stuart, B.H. *Infrared Spectroscopy: Fundamentals and Applications*; John Wiley & Sons Ltd, Chichester, UK, 2004.
- [5] Subramanian, A.Z.; Ryckeboer, E.M.P.; Dhakal, A.; Peyskens, F.; Malik, A.; Kuyken, B.; Zhao, H.; Pathak, S.; Ruocco, A.; De Groote, A.; Wuytens, P.C.; Martens, D.; Leo, F.; Xie, W.; Dave, U.D.; Muneeb, M.; Van Dorpe, P.; Van Campenhout, J.; Bogaerts,

- W.; Bienstman, P.; Le Thomas, N.; Van Thourhout, D.; Hens, Z.; Roelkens, G.; Baets, R. Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip (Invited). *Photon. Res.* **2015**, *3*(5), B47–B59.
- [6] Optical constants of H₂O. <http://refractiveindex.info/?shelf=main&book=H2O&page=Kedenburg> (accessed Feb 08, 2016).
- [7] *Technical Manual: Fluorescence Imaging - principles and methods*; amersham pharmacia biotech; 63-0035-28 Rev.A, 2000-12; <http://www.bu.edu/picf/files/2010/10/Fluor-ImagingPrinciples.pdf> (accessed Nov 25, 2015).
- [8] Sauer, M.; Hofkens, J.; Enderlein, J. *Basic Principles of Fluorescence Spectroscopy, in Handbook of Fluorescence Spectroscopy and Imaging: From Single Molecules to Ensembles*; Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2011.
- [9] Albani, J.-R. *Principles and Applications of Fluorescence Spectroscopy*; Blackwell Publishing company, Oxford, UK, 2007.
- [10] Raman, C.V.; Krishnan, K.S. A New Type of Secondary Radiation. *Nature* **1928**, *121*, 501–502.
- [11] Etchegoin, P.G.; Le Ru, E.C. *Basic Electromagnetic Theory of SERS, in Surface Enhanced Raman Spectroscopy: Analytical, Biophysical and Life Science Applications (ed S. Schlücker)*; Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010.
- [12] Lakowicz, J.R. *Principles of Fluorescence Spectroscopy, Second Edition*; Kluwer Academic Publishers, New York, USA, 1999.
- [13] Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press, New York, 2006.
- [14] Mohammed, A. Theoretical Studies of Raman Scattering. Ph.D. Thesis, Royal Institute of Technology, 2011.
- [15] Moskovits, M. Surface-enhanced Raman spectroscopy: a brief retrospective. *J. Raman Spectrosc.* **2005**, *36*, 485-496.

- [16] Willets, K.A.; Van Duyne, R.P. Localized Surface Plasmon Resonance Spectroscopy and Sensing *Annu. Rev. Phys. Chem.* **2007**, *58*, 267-297.
- [17] Anker, J.N.; Hall, W. P.; Lyandres, O.; Shah, N. C.; Zhao, J.; Van Duyne, R.P. Biosensing with plasmonic nanosensors. *Nat. Mater.* **2008**, *7*, 442-453.
- [18] Le Ru, E.; Etchegoin, P. *Principles of Surface Enhanced Raman Spectroscopy and related plasmonic effects*; Elsevier, Amsterdam, The Netherlands, 2009.
- [19] Sun, G.; Khurgin, J.B.; Ping Tsai, D. Comparative analysis of photoluminescence and Raman enhancement by metal nanoparticles. *Opt. Lett.* **2012** *37*(9), 1583-1585.
- [20] Sun, G.; Khurgin, J.B.; Yang, C. C. Impact of high-order surface plasmon modes of metal nanoparticles on enhancement of optical emission. *Appl. Phys. Lett.* **2009** *95*, 171103-1-171103-3.
- [21] Peticolas, W.L.; Nafie, L.; Stein, P.; Fanconi, B. Quantum Theory of the Intensities of Molecular Vibrational Spectra. *J. Chem. Phys.* **1970**, *52*(3), 1576-1584.
- [22] McCreery, R.L. *Raman Spectroscopy for Chemical Analysis*; John Wiley & Sons Inc., Hoboken, NJ, USA, 2000.
- [23] Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.; Skirtach, A.; Le Thomas, N.; Baets, R. In *Nanophotonic Lab-On-A-Chip Raman Sensors: a Sensitivity Comparison with Confocal Raman Microscope*, Proceedings of BioPhotonics 2015 Conference, Florence, Italy, May 20-22, 2015.

4

On-Chip Raman Spectroscopy

4.1 Introduction

In the previous chapter we have introduced several spectroscopic techniques and discussed their pros and cons. A common disadvantage for traditional spectroscopic systems lies in the fact that they usually rely on expensive and bulky instrumentation, prohibiting their widespread use, especially in non-laboratory environments. In recent years, there has been a large demand for hand-held devices that can be used in the field and that are capable of accurate, sensitive, and in situ spectroscopic detection. Such devices can serve applications ranging from physics and chemistry to biology and environmental science. Although optical sensing can be performed on different platforms ranging from bulk optics and optical fibers down to nanophotonic waveguides, an integrated photonics approach has significant advantages. First of all, it enables the integration of active and passive optical components. Second, the integration leads to devices that are robust, reliable, mass reproducible and cheap. Finally, integration allows for extensive parallelism and multiplexing. [1] After a short description of the basic concepts of optical waveguides and a brief introduction to the silicon nitride platform, we will give an in-depth analysis of spontaneous on-chip Raman spectroscopy.

The discussion on the coupling of dipole emission to the fundamental

waveguide mode [2] (see *Dipole emission near a waveguide* subsection) is the key theoretical cornerstone of both this PhD work [3], and the PhD work of Ashim Dhakal [4]. The measurement setup and procedures (see *Measurement setup* subsection) are also common to both PhD works. Nevertheless, the theoretical derivation of the amount of spontaneous Raman scattering coupling into the waveguide mode (see *Spontaneous on-chip Raman scattering* subsection) and the experimental demonstration of spontaneous on-chip Raman scattering (see *On-chip bulk Raman sensing* subsection) are at the heart of Ashim Dhakal's PhD work. In a later chapter we will describe how the key theoretical formula is used to describe on-chip SERS, which is the major topic of this PhD work.

4.2 Optical Waveguides

The fundamental idea behind integrated photonics is the manipulation of light by optical waveguides and not by free-space optical components like lenses and mirrors. A dielectric waveguide consists of a high refractive index core surrounded by lower refractive index cladding materials. The optical field in such structures can then be guided because light prefers to be concentrated in the area with the highest refractive index. A typical waveguide structure is shown in Figure 4.1(a). The higher index core is separated from the carrier substrate (which might have a higher refractive index as the core) by a lower index under- and uppercladding. [5] The discussion in this section is mainly taken from [5].

In order to discuss the modes of optical waveguides, an isotropic and non-magnetic medium and a harmonic time dependence of the electric and magnetic fields is assumed,

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r})e^{j\omega t}. \quad (4.1)$$

Moreover it is assumed that the following constitutive relations hold

$$\mathbf{D}(\mathbf{r}) = \epsilon\mathbf{E}(\mathbf{r}) = \epsilon_0 n(\mathbf{r})^2 \mathbf{E}(\mathbf{r}) \quad (4.2)$$

$$\mathbf{B}(\mathbf{r}) = \mu_0 \mathbf{H}(\mathbf{r}), \quad (4.3)$$

where the real part of the refractive index $n(\mathbf{r})$ determines the wavelength in that medium $\lambda = \lambda_0/n(\mathbf{r})$ ($\lambda_0 = 2\pi c/\omega = 2\pi/k_0$ is the vacuum wavelength for a field with angular frequency ω). The Maxwell equations

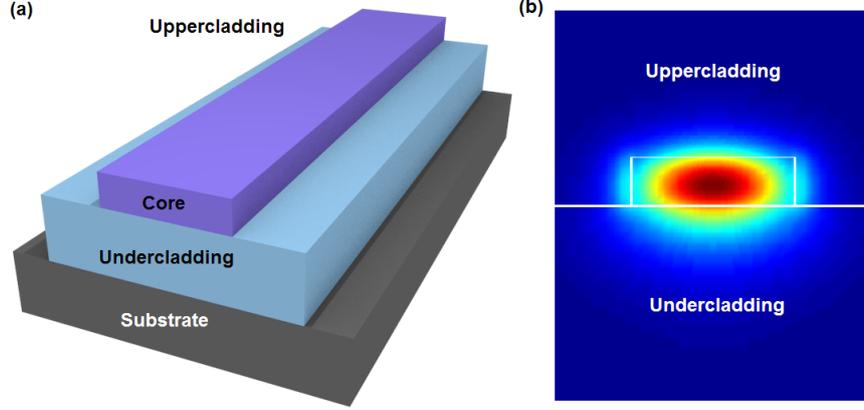


Figure 4.1: (a) A typical waveguide structure consisting of a substrate, a low refractive index under- and uppercladding and a high refractive index core. (b) Plot of the field intensity.

can now be used to derive the vectorial wave equations for the electric $\mathbf{E}(\mathbf{r})$ and magnetic $\mathbf{H}(\mathbf{r})$ field:

$$\nabla^2 \mathbf{E}(\mathbf{r}) + \nabla \left(\frac{\nabla n(\mathbf{r})^2}{n(\mathbf{r})^2} \mathbf{E}(\mathbf{r}) \right) + k_0^2 n(\mathbf{r})^2 \mathbf{E}(\mathbf{r}) = 0 \quad (4.4)$$

$$\nabla^2 \mathbf{H}(\mathbf{r}) + \left(\frac{\nabla n(\mathbf{r})^2}{n(\mathbf{r})^2} \right) \times (\nabla \times \mathbf{H}(\mathbf{r})) + k_0^2 n(\mathbf{r})^2 \mathbf{H}(\mathbf{r}) = 0. \quad (4.5)$$

In the above equations the gradient of the refractive index occurs, which couples the three components of the field vector. However, when the refractive index is piecewise constant or if the variation of the refractive index is small, we can neglect these gradients. In the latter case both vectorial equations decouple and reduce to the Helmholtz equation for every component of the electric and magnetic field vector:

$$\nabla^2 \psi(\mathbf{r}) + k_0^2 n(\mathbf{r})^2 \psi(\mathbf{r}) = 0. \quad (4.6)$$

Moreover, the electric and magnetic fields should satisfy the boundary conditions as introduced in an earlier chapter. In its simplest form a waveguide is invariant along the propagation direction of the optical power. Assuming the z -direction as the propagation direction, the refractive index profile can be written as $n(\mathbf{r}) = n(x, y)$. An eigenmode of the waveguide structure is a propagating or evanescent wave of which the transversal shape

does not change during propagation and can be represented by

$$\mathbf{E}(\mathbf{r}) = \mathbf{e}(x, y)e^{-j\beta z} \quad (4.7)$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{h}(x, y)e^{-j\beta z}. \quad (4.8)$$

An effective refractive index n_{eff} (and dielectric constant ϵ_{eff}) can be associated to the propagation constant β

$$n_{eff} = \frac{\beta}{k_0} \quad (4.9)$$

$$\epsilon_{eff} = n_{eff}^2. \quad (4.10)$$

For lossless optical waveguides ($\text{Im}(\epsilon(x, y)) = 0$) one can show that [5]

1. There are no eigenmodes with an eigenvalue larger than the maximum of the dielectric function.

$$\epsilon_{eff} < \max(\epsilon(x, y)) \quad (4.11)$$

2. Guided modes belong to a discrete set of eigenvalues which are in the range $\max(\epsilon_{clad}) < \epsilon_{eff} < \max(\epsilon(x, y))$, where ϵ_{clad} represents the dielectric constants of the different cladding materials, usually upper- and undercladding. For these modes

$$\lim_{|(x, y)| \rightarrow \infty} \psi(x, y) = 0. \quad (4.12)$$

3. The continuous part of the spectrum is formed by the radiating modes for which the eigenvalues satisfy $\epsilon_{eff} < \max(\epsilon_{clad})$. Radiating modes show an oscillating behaviour along at least one side of the waveguide structure. Depending on their effective refractive index they are classified as propagating or evanescent radiating modes. In the latter case the effective refractive index is purely imaginary.
4. Guided and radiating modes form a complete set of functions. This means that every field inside the waveguide can be represented by a superposition of these modes:

$$\mathbf{E}(\mathbf{r}) = \sum_m a_m \mathbf{e}_m(x, y)e^{-j\beta_m z} + \int a(k) \mathbf{e}_k(x, y)e^{-jkz} dk \quad (4.13)$$

The superposition consists of a discrete sum of guided modes and a continuous spectrum of radiating modes.

Usually numerical methods are required to solve for the complex field distributions in waveguide structures. Only a few cases can be calculated analytically and it is instructive to study one of such cases in order to show how discrete modes emerge and introduce the concept of TE and TM polarization. To this end, a symmetric slab waveguide depicted in Figure 4.2 is considered. A slab waveguide is not only invariant in the propagation direction (z), but also in one direction perpendicular to the propagation direction (here y). A symmetric slab has a high index core (index n_1 and thickness d) surrounded by two equal index ($n_2 < n_1$) materials. [5]

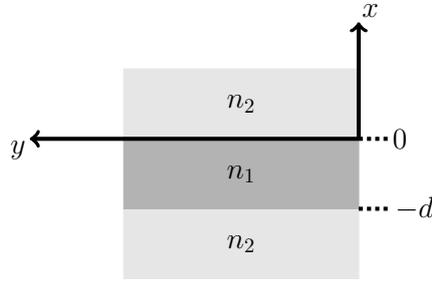


Figure 4.2: Geometry of a symmetric slab waveguide. [5]

The modal fields of a symmetric slab waveguide can then be written as

$$\mathbf{E}(\mathbf{r}) = \mathbf{e}(x)e^{-j\beta z} \quad (4.14)$$

$$\mathbf{H}(\mathbf{r}) = \mathbf{h}(x)e^{-j\beta z}. \quad (4.15)$$

Substituting these fields into Maxwell's equations leads to two sets of equations for the so called transverse electric (TE) and transverse magnetic (TM) eigenmodes, which are characterized by the field components $\mathbf{e}_y(x)$, $\mathbf{h}_x(x)$, $\mathbf{h}_z(x)$ and $\mathbf{h}_y(x)$, $\mathbf{e}_x(x)$, $\mathbf{e}_z(x)$ respectively. For the TE polarization, the electric field is parallel to the interface between the two media, while for the TM polarization the magnetic field is parallel to the interface. After applying Maxwell's equations and the appropriate boundary conditions, one can show that the solution for the TE case is given by

$$\mathbf{e}_y(x) = \begin{cases} Ae^{-\delta x}, & x \geq 0 \\ A \cos(\kappa x) + B \sin(\kappa x) & -d \leq x \leq 0 \\ (A \cos(\kappa d) - B \sin(\kappa d)) e^{\delta(x+d)} & x \leq -d \end{cases} \quad (4.16)$$

with

$$\delta = \sqrt{\beta^2 - n_2^2 k_0^2} \quad (4.17)$$

$$\kappa = \sqrt{n_1^2 k_0^2 - \beta^2}. \quad (4.18)$$

The eigenvalue equation for β , which will determine the discrete modes, is given by

$$\tan(\kappa d) = \frac{2\kappa\delta}{\kappa^2 - \delta^2}. \quad (4.19)$$

The constants A and B can be determined by imposing the proper boundary conditions. With each eigenvalue β , an eigenmode can be associated. Often an $\omega(\beta)$ diagram is used, being the graphical representation of the dispersion relation for the different eigenmodes. The associated effective index can be considered as some kind of average refractive index felt by the guided mode. In the specific case where there is only one guided mode, the waveguide is termed single mode. Outside the core, the field has the form

$$\mathbf{E}(\mathbf{r}) \propto e^{-\delta|x| - j\beta z}, \quad (4.20)$$

so while propagating along the z -direction, the field decays in the direction perpendicular to the core (x). This evanescent part of the field is prone to environmental variations around the core and is the basis for several on-chip sensing applications such as refractive index shift sensing and absorption spectroscopy. For a discussion of each of these applications we refer to the appropriate papers [6–9]. In the next section we will discuss the silicon nitride photonics platform and motivate why it is an appropriate choice for performing on-chip Raman spectroscopy.

4.3 Silicon nitride photonics

Of all the possible photonic integrated circuit (PIC) platforms that have been reported, those that are based on high index contrast (HIC) have gained the most attention in the past decade. HIC waveguides have a large difference between the refractive index of the core and that of the cladding and allow the fabrication of strongly miniaturized circuits, because one can make tight bends with very low loss. Furthermore HIC waveguides have tightly confined guided modes, implying that the electric field strength of the evanescent tail is very strong for a given optical power. This will lead to an intense interaction between light propagating in the waveguide and

its surroundings, which is of paramount importance for on-chip evanescent sensing. [1]

The most prominent example of a HIC waveguide platform is the silicon photonics waveguide platform. In this platform, Silicon-on-Insulator (SOI) wafers are used to form waveguides with a silicon core (refractive index of about 3.5 at 1550 nm) and a silica cladding (refractive index of 1.45 at 1550 nm). Silicon photonics has now reached a considerable level of maturity by using the same CMOS technologies and tools of the microelectronics industry to fabricate high-quality PICs on 200 and 300 mm SOI wafers with a high level of process control and yield. The major thrust for developing silicon photonics came from telecom applications, but increasingly silicon photonics is finding its way in important sensing and biosensing applications. Silicon is transparent between 1.1 and 8 μm , but the underlying oxide starts to absorb light from about 4 μm onwards. As a result SOI waveguides are only relevant in the 1.1 to 4 μm wavelength range. [1]

In order to extend silicon photonics to shorter wavelengths (below 1.1 μm), while maintaining the ability to make use of CMOS tools and preserving HIC, it is probably the most convenient to use silicon nitride (SiN). SiN has a refractive index of 1.89 around 780 nm. [10] Although SiN is transparent from 0.4 to 5 μm , it is particularly important for the visible and very near-IR wavelengths ($< 1 \mu\text{m}$) since this wavelength region covers the therapeutic window for biological media (750–930 nm). Within this region, there is minimal photo-damage to cells and negligible water absorption. Moreover there is a large availability of high-performance and low-cost light sources and silicon-based detectors in this wavelength range. [1] As we already discussed in a previous chapter, it is also beneficial to perform Raman spectroscopy in this near-IR window because fluorescence interference is largely mitigated due to the fact that the excitation energy is too low to excite electronic transitions.

Silicon nitride is predominantly deposited using low-pressure chemical vapor deposition (LPCVD) or plasma-enhanced chemical vapor deposition (PECVD). LPCVD provides an excellent control over the homogeneity of the refractive index and thickness. However, it remains a high-temperature process ($> 700^\circ\text{C}$) and induces high stress, particularly in thicker films ($> 300 \text{ nm}$). On the other hand, PECVD is a low-temperature process (200 – 400 $^\circ\text{C}$) and enables stressfree thicker film deposition, making it a better alternative for many photonic-based applications. The PECVD film homogeneity is however poorer than in the case of LPCVD films. There-

fore, an optimized process for low-loss PECVD SiN waveguides provides a very attractive route towards high-volume fabrication of integrated photonic devices. [10] Moreover people are trying to minimize the fluorescence from PECVD nitride, which is important for e.g. sensing applications. Cladded single mode rib waveguides with low waveguide (< 0.7 dB/cm) and bend (< 0.05 dB/90° for radii as small as $35 \mu\text{m}$) loss at 660 nm were fabricated within a CMOS pilot-line using PECVD technology. [11] A cross-section, perpendicular to the propagation direction, of a typical rib waveguide is shown in Figure 4.3. Such waveguides will also be used for the initial on-chip Raman work described furtheron. The high index core (refractive index n_{core}) has a width w_{rib} and height h_{rib} and is surrounded by a lower index top (index n_{tclad}) and bottom cladding (index n_{bclad}). A substrate with index n_{sub} is used as a carrier. Waveguide losses of < 1 dB/cm were also achieved for cladded photonic rib waveguides at 532 nm and 900 nm wavelength using PECVD SiN. [10] The fabrication process of such photonic wire waveguides, as described in [10], is briefly outlined below.

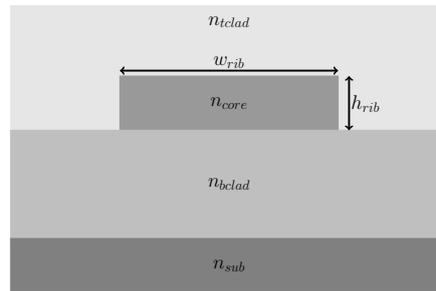


Figure 4.3: Cross-section of a photonic rib waveguide. The high index core (refractive index n_{core}) has a width w_{rib} and height h_{rib} and is surrounded by a lower index top (index n_{tclad}) and bottom cladding (index n_{bclad}). A substrate with index n_{sub} is used as a carrier.

To build a photonic circuit in SiN, one starts with a 200 mm bare silicon wafer. After cleaning the wafer, 2–2.4 μm of SiO_2 is deposited using a high-density plasma (HDP) CVD process. On top of the oxide, PECVD SiN is deposited using SiH_4 , N_2 and NH_3 at 400°C, ensuring CMOS back-end compatibility. After SiN deposition, waveguides are patterned with 193 nm optical lithography. This is followed by an inductive coupled plasma reactive ion etch process (ICP-RIE), using fluorine-based etch chemistry.

After dry etching, the wafers are cleaned with an oxygen plasma and a wet chemical process. A few typical waveguide dimensions and corresponding waveguide losses are summarized below.

Usually waveguides are designed to work in single mode operation, such that only the fundamental TE and TM mode propagate in the waveguide core and all other modes are in cut-off. If we assume a SiN refractive index of 1.89 (at 780 nm), a SiO₂ refractive index of 1.46 and $h_{rib} = 220$ nm, then the single mode width (at 780 nm) is 900 nm for an uncladded (i.e. an air top cladding) and about 630 nm for an SiO₂ top cladded waveguide. At 900 nm wavelength, single mode operation is obtained at ($h_{rib} = 220$ nm, $w_{rib} \approx 1100$ nm) for an uncladded and at ($h_{rib} = 220$ nm, $w_{rib} \approx 770$ nm) for a cladded waveguide respectively. In the relevant Raman window (780 – 950 nm) single mode waveguide operation can hence be obtained for uncladded rib waveguides with a height $h = 220$ nm and widths smaller than 900 nm. If the width gets too small, then also the fundamental modes will go into cut-off, which will result in a tremendous increase of the waveguide loss. At ($h_{rib} = 220$ nm, $w_{rib} = 700$ nm), the waveguide losses at 780 and 900 nm are in the range 0.5 – 1.5 dB/cm, hence low-loss single mode propagation in the relevant Raman window can be obtained using PECVD SiN. [10] In the next chapter we will discuss the performance of SiN waveguides which were fabricated with e-beam lithography in the context of the hybrid integration of SiN waveguides with nanoplasmonic antennas.

4.4 Raman radiation near a waveguide

In the above section it was shown that low-loss single mode SiN waveguides can be fabricated in the relevant Raman window using CMOS-compatible processes. Nevertheless, it remains to be shown whether the use of photonic integrated circuits can provide an advantage over conventional Raman spectroscopy systems. In this section we will outline the theoretical rationale for using on-chip Raman systems over conventional microscopy systems. The next sections will deal with the experimental on-chip Raman results.

A molecule emitting Raman radiation can in first instance be approximated by a dipole emitter (see previous chapter). The overall on-chip Raman scattering process then breaks down into the excitation of this dipole by the waveguide mode at the pump wavelength, and subsequent dipole emission, at the Stokes or anti-Stokes wavelength, into the waveguide

mode. The dipole moment of the emitter is then determined by the polarizability of the molecule and the strength of the waveguide mode at the pump wavelength. The emission process, which is modified due to the vicinity of the waveguide, will be discussed in the next subsection.

4.4.1 Dipole emission near a waveguide

As said, the theoretical treatment relies on the fact that an emitting molecule can be well approximated by a dipole source. The spontaneous decay rate of a dipole emitter in a homogeneous medium with index n is given by

$$\gamma_0 = \frac{n\omega^3 |\mathbf{d}|^2}{3\pi\hbar\epsilon_0 c^3} \quad (4.21)$$

where \mathbf{d} is the dipole moment of the emitter and ω the emission frequency. [12] In case of a conventional Raman microscope, the analyte can be considered as a collection of free-space radiating dipole emitters and the overall power collected by such a system has been derived in the previous chapter. In the waveguide case, the analyte will cover the core region, hence the top cladding consists of the analyte under study. Now, the molecules (dipole emitters) are not radiating in free-space, so the spontaneous emission rate will be modified due to the interaction with the environment (see Figure 4.4). If the waveguide is single mode, then part of the emission will radiate into the waveguide mode at a rate γ_{wg} (dotted lines), while the remaining part of the emission (solid lines) will decay into other radiating modes. This effect of spontaneous emission modification due to an emitter-environment interaction was established by Purcell. [13, 14] In general, one can calculate the spontaneous emission rate γ of a dipolar emitter, in the weak coupling regime, from Fermi's golden rule:

$$\gamma(\mathbf{r}_e, \omega) = \frac{2\pi}{\hbar^2} |g(\mathbf{r}_e, \omega)|^2 \rho(\omega) \quad (4.22)$$

where $|g(\mathbf{r}_e, \omega)|$ is the coupling strength between the dipole \mathbf{d} and the electromagnetic field $\mathbf{E}(\mathbf{r}_e)$ at the emitter position \mathbf{r}_e and $\rho(\omega)$ the electromagnetic density of states at frequency ω . [2]

For a rib waveguide with a sufficiently small cross-sectional area to only have a single TE and a single TM mode, a 1-dimensional (1D) density of states can be assumed. [2] The density of states for decay into this single TE

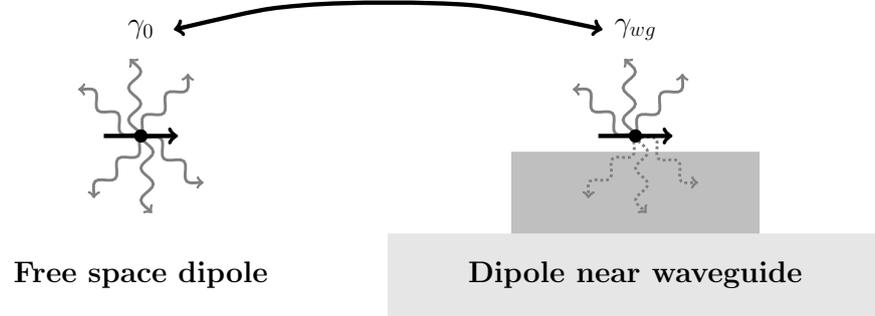


Figure 4.4: Radiation of a dipole in free-space (γ_0) and near a rib waveguide (γ_{wg}).

and TM mode can be calculated by considering a 1D quantization interval of arbitrary length L . Assume a particle which is confined to this interval, then it can be described by a wave function $\phi(\mathbf{r}) \propto \exp(i\mathbf{k} \cdot \mathbf{r})$. [15] The volume of this 1D quantization interval is simply L and $\phi(x) \propto \exp(ikx)$ with $x \in [0, L]$. By assuming an infinitely high potential barrier around the box, the wavefunction must have nodes at the walls, such that k must satisfy

$$k = m \frac{\pi}{L}, \quad m = 1, 2, 3, \dots \quad (4.23)$$

The various k -states are hence equally distributed over k -space with a spacing π/L such that every state needs a volume

$$V_k = \frac{\pi}{L} \quad (4.24)$$

in k -space. The number of states in an infinitesimal section $[k, k + dk]$ of 1D k -space is then

$$\rho(k)dk = \frac{L}{\pi} dk \quad (4.25)$$

Using the photon dispersion relation $E = \hbar\omega(k)$ one can furthermore write the density of states on the energy axis as

$$\rho(E)dE = \rho(E(k)) \frac{dk}{dE} dE = \rho(k) \frac{dk}{d\omega} d\omega = \frac{L}{\pi} \left(\frac{d\omega}{dk} \right)^{-1} d\omega. \quad (4.26)$$

For a waveguide mode with general dispersion relation $\omega(k)$, the group velocity is given by

$$v_g = \frac{d\omega}{dk} \quad (4.27)$$

such that the density of states at frequency ω is eventually given by

$$\rho(\omega) = \frac{L}{\pi v_g}. \quad (4.28)$$

The density of states is hence inversely proportional to the group velocity of the waveguide mode. [15] Apart from the density of states, we also need to calculate the coupling strength in order to evaluate the modified spontaneous emission rate. This coupling strength is given by

$$|g(\mathbf{r}_e, \omega)| = \zeta^2 |\mathbf{d} \cdot \mathbf{E}(\mathbf{r}_e, \omega)|^2 \quad (4.29)$$

where ζ is an energy normalization constant which relates the frequency of the radiation to the electromagnetic energy density (see also furtheron). A derivation of this constant is given below as it provides more insight into the different assumptions made to calculate the modified spontaneous emission rate. From the sourceless Maxwell equations it follows that

$$\begin{aligned} \nabla \cdot (\mathbf{E}(\mathbf{r}, t) \times \mathbf{H}(\mathbf{r}, t)) &= -\mathbf{E}(\mathbf{r}, t) \cdot \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} - \mathbf{H}(\mathbf{r}, t) \cdot \frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \\ &= -\frac{\partial W_E(\mathbf{r}, t)}{\partial t} - P_E(\mathbf{r}, t) - \frac{\partial W_M(\mathbf{r}, t)}{\partial t} \end{aligned} \quad (4.30)$$

where $W_E(\mathbf{r}, t)$, $W_M(\mathbf{r}, t)$ are the electric and magnetic field energy density respectively and $P_E(\mathbf{r}, t)$ the electric polarization loss density. [16] Taking into account that a general time-dependent electric field can be decomposed into its different Fourier components

$$\mathbf{E}(\mathbf{r}, t) = \int_{-\infty}^{+\infty} d\omega \left(\frac{\mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} + \mathbf{E}^*(\mathbf{r}, \omega) e^{i\omega t}}{2} \right), \quad (4.31)$$

it is possible to show that

$$\begin{aligned} &\mathbf{E}(\mathbf{r}, t) \cdot \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} \\ &= \frac{\epsilon_0}{2} \int d\omega \int d\omega' \mathbf{E}(\mathbf{r}, \omega) \cdot \mathbf{E}^*(\mathbf{r}, \omega') (i\omega' \epsilon^*(\mathbf{r}, \omega') - i\omega \epsilon(\mathbf{r}, \omega)) e^{-i(\omega - \omega')t}. \end{aligned} \quad (4.32)$$

Under the assumption of a real dielectric permittivity (i.e. a lossless medium) the latter expression can be written as

$$\underbrace{\frac{\partial}{\partial t} \left(\frac{\epsilon_0}{2} \int d\omega \int d\omega' \mathbf{E}(\mathbf{r}, \omega) \cdot \mathbf{E}^*(\mathbf{r}, \omega') \frac{(\omega' \epsilon(\mathbf{r}, \omega') - \omega \epsilon(\mathbf{r}, \omega))}{\omega' - \omega} e^{-i(\omega - \omega')t} \right)}_{W_E(\mathbf{r}, t)}. \quad (4.33)$$

For a narrow-band signal with frequency centered around ω_0 the electric field is

$$\mathbf{E}(\mathbf{r}, \omega) = \frac{1}{2} (\delta(\omega - \omega_0)\mathbf{E}(\mathbf{r}, \omega) + \delta(\omega + \omega_0)\mathbf{E}^*(\mathbf{r}, \omega)). \quad (4.34)$$

The time-averaged energy density for such a signal is

$$\langle W_E(\mathbf{r}, t) \rangle = \frac{1}{T} \int_0^T dt W_E(\mathbf{r}, t) \quad (4.35)$$

with $T = 2\pi/\omega_0$. After some calculation one can show that [16]

$$\langle W_E(\mathbf{r}, t) \rangle = \frac{\epsilon_0}{4} \left. \frac{\partial (\omega \epsilon(\mathbf{r}, \omega))}{\partial \omega} \right|_{\omega=\omega_0} |\mathbf{E}(\mathbf{r}, \omega_0)|^2. \quad (4.36)$$

Analogously, the magnetic field energy density is

$$\langle W_M(\mathbf{r}, t) \rangle = \frac{\mu_0}{4} |\mathbf{H}(\mathbf{r}, \omega_0)|^2, \quad (4.37)$$

such that the time-averaged field energy density for a dispersion-free material with real dielectric constant is given by [16]

$$\langle W_{EM} \rangle = \frac{\epsilon_0 \epsilon(\mathbf{r})}{4} |\mathbf{E}(\mathbf{r}, \omega_0)|^2 + \frac{\mu_0}{4} |\mathbf{H}(\mathbf{r}, \omega_0)|^2. \quad (4.38)$$

When the energy is carried by a single waveguide mode with frequency ω_0 , the normalization factor ζ^2 is determined by

$$\begin{aligned} \hbar\omega_0 &= L \int \int d\mathbf{r} \zeta^2 (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega_0)|^2 + \mu_0 |\mathbf{H}(\mathbf{r}, \omega_0)|^2) \\ &\approx 2L\zeta^2 \int \int d\mathbf{r} (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega_0)|^2) \end{aligned} \quad (4.39)$$

where the integration runs over the complete cross-sectional area of the waveguide. [2] Finally ζ^2 is then

$$\zeta^2 = \frac{\hbar\omega_0}{2L \int \int d\mathbf{r} (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega_0)|^2)}. \quad (4.40)$$

Please note that the arbitrary quantization length L also appears in this expression because the integration only runs over a two-dimensional

surface area while the energy density itself is contained in a three-dimensional volume. Assuming a dipole oriented along the direction \mathbf{e}_d and setting $\omega_0 \rightarrow \omega$, the coupling strength is given by

$$\begin{aligned} |g(\mathbf{r}_e, \omega)|^2 &= \zeta^2 |\mathbf{d} \cdot \mathbf{E}(\mathbf{r}_e, \omega)|^2 \\ &= |\mathbf{d}|^2 |\mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \omega)|^2 \frac{\hbar\omega}{2L \int \int d\mathbf{r} (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega)|^2)} \\ &= \frac{\hbar\omega |\mathbf{d}|^2}{2\epsilon_0 \epsilon(\mathbf{r}_e) V_{eff}(\mathbf{r}_e, \omega)}, \end{aligned} \quad (4.41)$$

where an effective mode volume $V_{eff} = A_{eff}L$ has been introduced, with

$$A_{eff}(\mathbf{r}_e, \omega) = \frac{\int \int d\mathbf{r} (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega)|^2)}{\epsilon_0 \epsilon(\mathbf{r}_e) |\mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \omega)|^2} \quad (4.42)$$

the effective mode area at the emitter position \mathbf{r}_e . The decay rate γ_{wg} of an emitter, positioned in a medium with refractive index n , into the fundamental mode $\mathbf{E}(\mathbf{r}, \omega)$ of the single mode waveguide is then

$$\gamma_{wg} = \frac{\omega |\mathbf{d}|^2}{\hbar v_g \epsilon_0 n^2 A_{eff}(\mathbf{r}_e, \omega)}. \quad (4.43)$$

Normalized with respect to the dipole emission in a homogeneous medium with index n one eventually gets

$$\frac{\gamma_{wg}}{\gamma_0} = \frac{3}{4\pi} \frac{(c/n)}{v_g} \left(\frac{\lambda}{n}\right)^2 \frac{1}{A_{eff}}, \quad (4.44)$$

where λ is the free space wavelength $\lambda = 2\pi c/\omega$. The normalized decay rate also equals the ratio of the power coupled into the waveguide mode P_{wg}^{tot} to the power radiated in a homogeneous medium P_0 [2],

$$\frac{\gamma_{wg}}{\gamma_0} = \frac{P_{wg}^{tot}}{P_0}. \quad (4.45)$$

4.4.2 Spontaneous on-chip Raman scattering

Formula (4.44) will now be used to calculate the overall spontaneous Raman signal coupling back into the waveguide mode as a result of a collection of incoherently radiating dipoles on top of the waveguide. Since

only half of the power is coupled into the forward propagating waveguide mode [12], the power at the output facet of the waveguide P_{wg} equals

$$\frac{P_{wg}}{P_0} = \frac{3}{8\pi} \frac{n_g}{n} \left(\frac{\lambda}{n}\right)^2 \frac{\epsilon_0 \epsilon(\mathbf{r}_e) |\mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \omega)|^2}{\int \int d\mathbf{r} (\epsilon_0 \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \omega)|^2)}, \quad (4.46)$$

where the group index $n_g = c/v_g$. The group index is also wavelength dependent. [2] So a Raman emitter which emits P_0 in a homogeneous medium will generate a power P_{wg} that is coupled to the fundamental mode of a single mode waveguide. When the emitter, at position \mathbf{r}_e , is excited by the waveguide mode with a guided pump power P_{pump} at pump wavelength λ_P , then the dipole strength can be written in terms of the pump power and the modal field $\mathbf{E}(\mathbf{r}, \lambda_P)$ [4]

$$|\mathbf{d}|^2 = \alpha^2 |\mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \lambda_P)|^2 \frac{n_g(\lambda_P) P_{pump}}{c \epsilon_0 \int \int \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \lambda_P)|^2 d\mathbf{r}}, \quad (4.47)$$

where α is the Raman polarizability (assumed scalar for simplicity). The integration in the denominator runs over the complete waveguide cross-section (uppercladding, core and undercladding). The emitter itself radiates at a wavelength λ_S , so

$$\frac{P_{wg}(\mathbf{r}_e)}{P_0} = \frac{3}{8\pi} \frac{n_g(\lambda_S)}{n} \left(\frac{\lambda_S}{n}\right)^2 \frac{\epsilon(\mathbf{r}_e) |\mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \lambda_S)|^2}{\int \int \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \lambda_S)|^2 d\mathbf{r}}. \quad (4.48)$$

Using the free space radiated power

$$P_0 = \frac{4\pi^3 c |\mathbf{d}|^2}{3\epsilon_0 \lambda_S^4} \quad (4.49)$$

and equation (4.47) one can show that [4]

$$\frac{P_{wg}(\mathbf{r}_e, \lambda_P, \lambda_S)}{P_{pump}} = \frac{\pi^2 \alpha^2 \eta(\mathbf{r}_e, \lambda_P, \lambda_S)}{\epsilon_0^2 \lambda_S^4} \quad (4.50)$$

where

$$\eta(\mathbf{r}_e, \lambda_P, \lambda_S) = \frac{n_g(\lambda_P) n_g(\lambda_S) \lambda_S^2 |\mathbf{E}_d(\mathbf{r}_e, \lambda_P)|^2 |\mathbf{E}_d(\mathbf{r}_e, \lambda_S)|^2}{n \left(\int \int \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \lambda_P)|^2 d\mathbf{r} \right) \left(\int \int \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r}, \lambda_S)|^2 d\mathbf{r} \right)} \quad (4.51)$$

and $\mathbf{E}_d(\mathbf{r}_e, \lambda) = \mathbf{e}_d \cdot \mathbf{E}(\mathbf{r}_e, \lambda)$. The factor $\frac{\pi^2 \alpha^2}{\epsilon_0^2 \lambda_S^4}$ is the earlier introduced Raman cross-section $\sigma = \sigma(\lambda_S)$ [17], so eventually one gets

$$\frac{P_{wg}(\mathbf{r}_e, \lambda_P, \lambda_S)}{P_{pump}} = \sigma(\lambda_S) \frac{\eta(\mathbf{r}_e, \lambda_P, \lambda_S)}{2}. \quad (4.52)$$

Formula (4.52) applies to a single emitter at a position \mathbf{r}_e around the waveguide core. When cladding the waveguide with an analyte in solution, the complete uppercladding consists of a collection of emitting molecules as shown in Figure 4.5 (blue dots). If the molecules are uniformly suspended with a density ρ , then a thin sheet of thickness dz will generate an overall guided Stokes power of

$$\frac{dP_{wg}}{P_{pump}} = dz \rho \sigma \int \int \frac{\eta(\mathbf{r}_e, \lambda_P, \lambda_S)}{2} d\mathbf{r}_e = dz \rho \sigma \frac{\eta_0(\lambda_P, \lambda_S)}{2} = dz \rho \sigma \frac{\eta_0}{2} \quad (4.53)$$

where the integration is evaluated over the complete Raman cladding. The specific conversion efficiency η_0 is a dimensionless parameter which only depends on the specific waveguide geometry and which can be evaluated using electromagnetic solvers. [4]

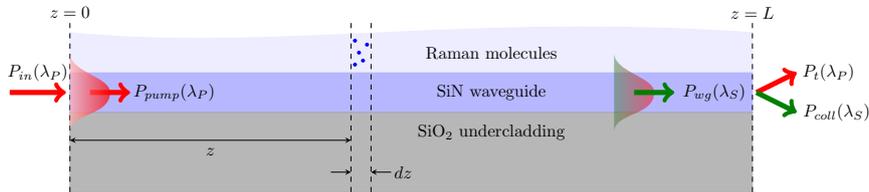


Figure 4.5: On-chip spontaneous Raman spectroscopy in forward scattering.

In a practical situation, a pump laser with initial power P_{in} will be coupled to the single mode waveguide with coupling efficiency γ_{in} such that $P_{pump} = \gamma_{in} P_{in}$. During propagation, molecules in the uppercladding are excited and the scattered spontaneous Raman signal is collected via evanescent coupling to the same waveguide. This Stokes light is subsequently coupled out with a coupling efficiency γ_{out} such that the collected Stokes power is $dP_{coll} = \gamma_{out} dP_{wg}$. At the same time, the pump and Stokes power will be attenuated by the inherent waveguide losses at the pump ($\alpha_{wg}(\lambda_P) = \alpha_P$) and Stokes ($\alpha_{wg}(\lambda_S) = \alpha_S$) wavelength

respectively. By integrating over the complete waveguide length L , one then finally obtains the total collected Stokes power $P_{coll} = \gamma_{out} P_{wg}$ at the output facet of the waveguide [4]

$$\frac{P_{coll}}{P_{in}} = \rho\sigma \frac{\eta_0}{2} \gamma_{in} \gamma_{out} \int_0^L e^{-\alpha_P z} e^{-\alpha_S (L-z)} dz. \quad (4.54)$$

Assuming that the outcoupling efficiency is the same for the pump and Stokes wavelength, the transmitted pump power P_t is then

$$P_t = \gamma_{in} \gamma_{out} e^{-\alpha_P L} P_{in} \quad (4.55)$$

such that

$$\frac{P_{coll}}{P_t} = \zeta(L) = \rho\sigma \frac{\eta_0}{2} \left(\frac{e^{L\Delta\alpha} - 1}{\Delta\alpha} \right) \quad (4.56)$$

and

$$\frac{P_{coll}}{P_{in}} = \xi(L) = \rho\sigma \frac{\eta_0}{2} \gamma_{in} \gamma_{out} e^{-\alpha_P L} \left(\frac{e^{L\Delta\alpha} - 1}{\Delta\alpha} \right) \quad (4.57)$$

with $\Delta\alpha = \alpha_P - \alpha_S$. For a given λ_P and λ_S , $\xi(L)$ is maximal for $L_{opt} = \frac{\ln(\alpha_P/\alpha_S)}{\Delta\alpha}$. Assuming that the waveguide losses at pump and Stokes wavelength are more or less the same, such that $\Delta\alpha \rightarrow 0$ and $\alpha_P = \alpha_S = \alpha$, this reduces to $L_{opt} = 1/\alpha$. In the $\Delta\alpha \rightarrow 0$ limit, $\zeta(L)$ equals

$$\zeta(L) = \frac{\rho\sigma\eta_0 L}{2}. \quad (4.58)$$

In the previous chapter we introduced the conversion efficiency of a diffraction limited confocal Raman microscope as

$$\eta_{DL} = 2\rho\sigma \frac{\lambda_P}{n} \quad (4.59)$$

where n is the refractive index of the immersion medium. This efficiency can now be compared to the efficiency ζ , eventually defining the enhancement ratio Φ of waveguide-based Raman systems

$$\Phi = \frac{\zeta}{\eta_{DL}} = \frac{n\eta_0}{4} \left(\frac{L}{\lambda_P} \right). \quad (4.60)$$

For the fundamental TE-mode of a SiN rib waveguide with $w_{rib} = 700$ nm and $h_{rib} = 220$ nm, $\eta_0 = 0.064$ if a pump wavelength of $\lambda_P = 785$ nm is used (and assuming $\eta_0(\lambda_P, \lambda_S) \approx \eta_0(\lambda_P, \lambda_P)$). The enhancement ratio of a waveguide with a length of 1 cm is then $\Phi \approx 200$ (for $n = 1$). So

the waveguide geometry offers a clear advantage as compared to a confocal Raman system. The main reason is that the ratio L/λ_P can be very high. [4, 18] Depending on the specific waveguide geometry, the factor η_0 can also be improved, providing a further boost in the overall enhancement ratio. In the next section we will outline the procedures to measure on-chip Raman signals and discuss some experimental results.

4.5 Evanescent excitation and collection of Raman spectra

4.5.1 Measurement setup

Spontaneous on-chip Raman spectra are measured with the setup depicted in Figure 4.6. A tunable Ti:saph laser is set to a pump wavelength of $\lambda_P = 785$ nm (red) after which the polarized beam passes through a half-wave plate ($\lambda/2$) in order to rotate the polarization to a TE-polarized beam. A beamsplitter BS1 then splits the beam into two parts (solid and dashed red line). The solid path is used to generate the forward propagating Raman beam and passes through a laser line filter (LLF) at λ_P for side-band suppression before it is coupled into the chip by an aspheric lens (ASPH). The output beam is then collected by an objective (OBJ) and passes through a polarizer P (set to TE) before it is filtered by a dichroic mirror which reflects the pump beam and transmits all Stokes wavelengths (green). The Stokes light is subsequently collected into a fiber using a parabolic mirror collimator (PMC) after which the fiber is split by a fiber splitter (FS) of which 1% goes to a power meter (PM) and 99% to a commercial spectrometer from Avantes (AvaSpec-ULS2048XL). Mirror M2 blocks the second (dashed) path during the measurement but can be removed for alignment purposes. The camera (CAM) and the 1% fiber tap are used during alignment and to measure the transmitted power P_t . Because the coupling between the small waveguide core and the external optics is easily lost, even for the smallest misalignments, it is critical to achieve a very precise alignment in order to allow a sufficiently high transmission of the Raman light. Below we will outline a possible alignment procedure in which only the sample and the position of the aspheric lens are changed while all other optical elements remain fixed. This procedure hence requires a one-time alignment of all other optical elements, while a re-alignment of the sample stage and the aspheric lens is required every time a new sample is mounted on the setup.

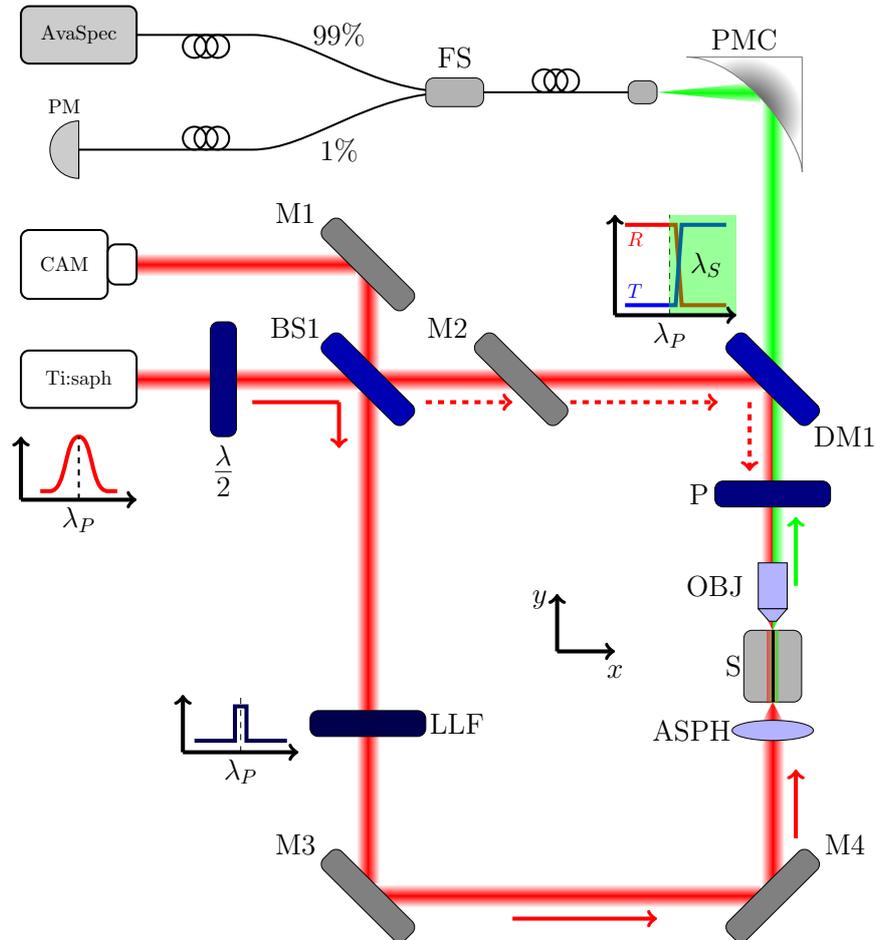


Figure 4.6: Measurement setup. Ti:saph: tunable Ti:sapphire laser emitting the pump beam at $\lambda_P = 785$ nm, CAM: camera, PM: power meter, AvaSpec: spectrometer from Avantes, BS1: beamsplitter, $\lambda/2$: half-wave plate, LLF: laser line filter for 785 nm, P: polarizer, M1/M3/M4: fixed mirrors, M2: removable mirror/beam block, OBJ: objective (50X, NA=0.9), ASPH: aspheric lens (NA=0.5), S: sample stage, DM1: dichroic mirror (reflection R and transmission T shown), PMC: parabolic mirror collimator (EFL=15 mm, NA=0.2), FS: fiber splitter.

Initially the sample stage is removed and a tip-to-tip alignment between the aspheric lens and the objective is carried out. To this end, the laser wavelength is set to $\lambda_T = 800$ nm such that it can be transmitted through DM1 and collected in the power meter. The transmitted power is maximized by moving the aspheric lens, which is mounted on a piezo-controlled XYZ-stage while the objective lens is fixed. By optimizing the transmitted power, a colinear beam between the aspheric lens and the objective is defined. After the sample stage has been installed again, the laser wavelength is set back to $\lambda_T = 785$ nm and mirror M2 is removed such that the laser is incident on the chip from the objective side (dashed lines). This side will be denoted as the backside furtheron, while the side adjacent to the aspheric lens is the frontside. At the same time the forward path (solid lines) is blocked. The objective itself is fixed, but the sample stage S is mounted on a piezo-controlled XYZ-stage. While the reflection from the backside of the chip is monitored by the camera CAM, the sample stage is moved until the waveguide core is properly imaged on the camera. Once properly focused, this will also be visible by a transmission of the light through the waveguide one is using to align (which will appear as a bright spot at the other side of the waveguide). This guarantees that the output facet of the chip is aligned with the optical path defined by the objective and all optics behind (such as the dichroic mirror and the parabolic mirror collimator). Once this is established, the forward path is opened (and the backward blocked) such that the laser is now incident on the frontside of the chip. The aspheric lens is also mounted on a piezo-controlled XYZ-stage and while the sample stage is now kept fixed, the aspheric lens is moved towards the input facet of the chip (in the y -direction, see Figure 4.6) until the beam is focused on the input facet. It is very important not to move the aspheric lens in the x - and z -direction such that the straight optical path along the y -direction (defined by the initial tip-to-tip alignment of the aspheric lens and the objective) remains fixed. In order to get the beam completely in focus (which can be monitored by the reflection of the input facet on the camera CAM) it is possible that the height of the sample stage has to be adapted such that the input facet of the chip is positioned at the proper height with respect to the aspheric lens. If such a height change is required, this means that the sample itself is tilted out of plane (so is not coplanar with the (x, y) -plane). Furthermore such a height change will result in a misalignment with respect to the backside (which was aligned before the height change). The tilt of the sample stage can be changed

along two angles and the movement is such that the backside is fixed upon tilt corrections, meaning that when we tilt up/down or left/right the frontside will move with respect to a fixed backside. At first, it is most convenient to check how the sample is tilted out of plane. This can be derived based on the previously described height change which is required to focus the aspheric lens on the input facet of the chip. Once this has been sorted out the sample should be tilted along the angle between the y - and z -axis. This tilt will require a re-alignment of the backside and frontside according to the procedure outlined above. This tilt change and re-alignment have to be repeated until the sample is coplanar with the (x, y) -plane. The sample can also be tilted in the (x, y) -plane, meaning that the straight waveguide used for alignment is not colinear with the initial optical path defined by the aspheric lens and the objective. Similarly as the tilt corrections described above, one can now tilt and re-align to make the waveguide colinear with the initial optical path. After this procedure, it should be possible to obtain transmission when the light is either coupled from the front- or the backside. For finetuning, the wavelength of the Ti:saph is then set to $\lambda_T = 800$ nm and mirror M2 is used again to block the backward path. Light coupled and transmitted from the frontside can then be transmitted through DM1 and collected in the power meter. First, the piezo of the sample stage is tuned to maximize the transmitted power P_t in the forward path. Only in the final step the aspheric lens is moved slightly along the x - and z -axis (with subsequent retuning of the sample stage) in order to optimize P_t . After optimizing the transmission, the laser is set back to $\lambda_P = 785$ nm and the measurement can start.

4.5.2 On-chip bulk Raman sensing

In [4] the first experimental demonstration of on-chip Raman sensing using the evanescent field of a single mode SiN photonic waveguide is reported. The work described in [4] (outlined in this section) is a key result of the PhD work of Ashim Dhakal. SiN waveguides with a cross-section of 220×700 nm² on top of a $2.4 \mu\text{m}$ SiO₂ cladding on a silicon substrate were used as the sensing platform. The waveguides (with different lengths 0.7, 1.6, 2.5, 4.4 and 8.1 cm) were wound up to form spirals as shown in Figure 4.7. A droplet of isopropyl alcohol (IPA) was then used as an analyte and applied on top of the waveguides. It was ensured that the IPA covers the entire waveguide by visually monitoring the chip with a camera on top. The analyte is excited by the evanescent tail of the fundamental waveguide mode, and the generated Raman signal is also collected back in the same

waveguide mode. [4] Since the evanescent tail is completely covered by IPA, this type of experiment is termed bulk Raman sensing.

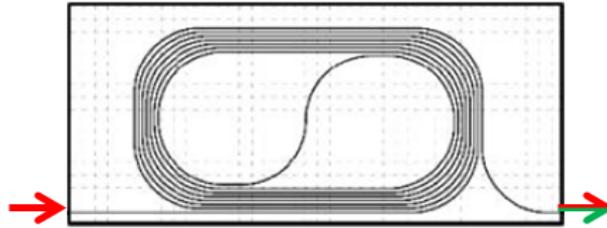


Figure 4.7: Top view of a spiral waveguide. The input (left arrow) and output (right arrow) of the waveguide are at the same height. Adapted from [4].

Measurements were performed using the setup described in the previous section. Spectra were then recorded before and after application of IPA, as shown in Figure 4.8. Before application of IPA, a broad Raman emission in the range $2100\text{-}2400\text{ cm}^{-1}$ with a peak around 2330 cm^{-1} (960 nm) is observed. Based on measurements with a confocal Raman microscope, this peak is identified as Raman emission from the SiN material. Another feature of the spectrum is the broad Raman background below 1200 cm^{-1} , which is also attributed to the SiN core. [4] This Raman background generates a substantial shot noise contribution (proportional to the square root of the background signal) that can potentially mask small signal peaks.

Immediately after application of IPA, an intense Raman peak at 819 cm^{-1} (839 nm) is readily observed. Other peaks associated to IPA (1450 and 2880 cm^{-1}) can also be distinguished and correspond well with IPA spectra reported in literature. [4] Figure 4.9(a) shows the measured $\xi(L)$ for the 819 cm^{-1} peak as a function of the waveguide length. Ten different sets of waveguide samples with different lengths were used for the measurements. [4] The average and standard deviation of the measured $\xi(L)$ and a least squared error fit of the averages to the equation for $\xi(L)$ are also shown. The results follow the expected $\xi(L)$ trend, despite the considerable standard deviation. This deviation can be ascribed to differences in the quality of the waveguide facets, particles that get stuck on the waveguides during cleaving and imperfections in waveguide processing. In Figure

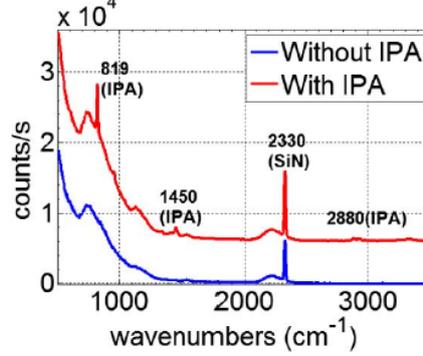


Figure 4.8: Raman spectra measured from a 1.6 cm waveguide without IPA (blue) and with IPA (red). The spectra are shifted vertically for clarity, with respective zeros at 3500 cm^{-1} . Taken from [4].

4.9(b) the measured $\zeta(L)$ values for the 819 cm^{-1} peak are depicted (as well as the mean, standard deviation and fit to the corresponding equation). For $\Delta\alpha \approx 0$ a quasi-linear trend is indeed observed. Unlike $\xi(L)$, which depends on coupling and waveguide losses and hence suffers from a large variance, $\zeta(L)$ has a smaller variance and is suitable for directly assessing the specific efficiency of on-chip Raman scattering as defined by $\zeta(L) = \rho\sigma\eta_0L$. This experiment clearly validates the developed theory for spontaneous on-chip Raman scattering and provides the first experimental demonstration thereof. [4] To achieve high-performance on-chip Raman sensors, it is crucial to optimize the various waveguide design parameters, such as geometry, dimensions, polarization of the mode and index contrast. By doing so, one can optimize η_0 . For more details we refer the reader to [19] in which such an optimization analysis is described.

4.6 Conclusion

In this chapter we motivated the use of a silicon nitride platform for on-chip Raman spectroscopy and briefly touched upon the fabrication process and the typical waveguide losses one obtains with this platform. Moreover we included an extensive theoretical treatment on dipole radiation near waveguides and outlined an analytical model for spontaneous on-chip Raman spectroscopy. Furthermore, we discussed the first experimental demonstration of such on-chip spontaneous Raman sensing using the

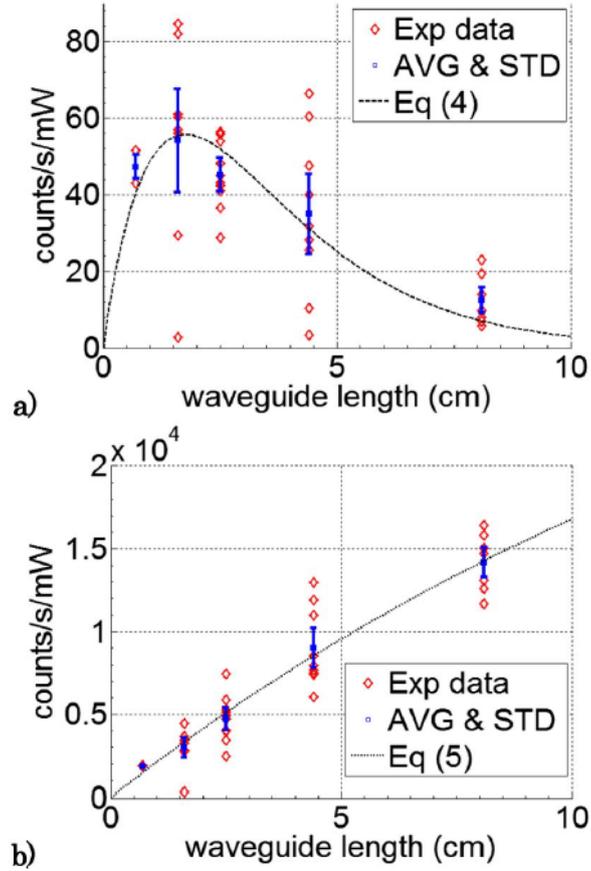


Figure 4.9: (a) $\xi(L)$ and (b) $\zeta(L)$ for the 819 cm^{-1} peak measured using photonic waveguides of several lengths. Red diamond markers: actual measured data. Blue solid squares with error bars: mean and standard deviation. Black dashed line: theoretical fit to the respective equations of $\xi(L)$ and $\zeta(L)$. Taken from [4].

evanescent field of a single mode SiN photonic waveguide. [4] Nevertheless, it is not convenient to start from the standard silicon nitride chips described in this chapter when one desires to integrate nanoplasmonic antennas (e.g. for on-chip SERS) on top of them. In the next chapter we will therefore outline alternative fabrication strategies for the hybrid integration of nanoplasmonic antennas on top of photonic waveguides.

References

- [1] Subramanian, A.Z.; Ryckeboer, E.M.P.; Dhakal, A.; Peyskens, F.; Malik, A.; Kuyken, B.; Zhao, H.; Pathak, S.; Ruocco, A.; De Groote, A.; Wuytens, P.C.; Martens, D.; Leo, F.; Xie, W.; Dave, U.D.; Muneeb, M.; Van Dorpe, P.; Van Campenhout, J.; Bogaerts, W.; Bienstman, P.; Le Thomas, N.; Van Thourhout, D.; Hens, Z.; Roelkens, G.; Baets, R. Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip (Invited). *Photon. Res.* **2015**, *3*(5), B47–B59.
- [2] Chul Jun, Y.; Briggs, R.M.; Atwater, H.A.; Brongersma, M.L. Broadband enhancement of light emission in silicon slot waveguides. *Opt. Express* **2009**, *17*(9), 7479–7490.
- [3] Peyskens, F.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform. *ACS Photonics* **2016**, *3*(1), 102–108.
- [4] Dhakal, A.; Subramanian, A.Z.; Wuytens, P.; Peyskens, F.; Le Thomas, N.; Baets, R. Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides. *Opt. Lett.* **2014**, *39*, 4025–4028.
- [5] Van Thourhout, D.; Baets, R.; Stijns, E.; Ottevaere, H. Course notes *Microphotonics*, 2011–2012.
- [6] Hoste, J-W.; Werquin, S.; Claes, T.; Bienstman, P. Conformational analysis of proteins with a dual polarisation silicon microring. *Opt. Express* **2014**, *22*(3), 2807–2820.
- [7] Claes, T.; Bogaerts, W.; Bienstman, P. Vernier-cascade label-free biosensor with integrated arrayed waveguide grating for wavelength

- interrogation with low-cost broadband source. *Opt. Lett.* **2011**, *36*(17), 3320–3322.
- [8] De Vos, K.; Bartolozzi, I.; Schacht, E.; Bienstman, P.; Baets, R. Silicon-on-Insulator microring resonator for sensitive and label-free biosensing. *Opt. Express* **2007**, *15*(12), 7610–7615.
- [9] Ryckeboer, E.M.P.; Bockstaele, R.; Vanslembrouck, M.; Baets, R. Glucose sensing by waveguide-based absorption spectroscopy on a silicon chip. *Biomed. Opt. Express* **2014**, *5*(5), 1636–1648.
- [10] Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Jansen, R.; Claes, T.; Rottenberg, X.; Peyskens, F.; Selvaraja, S.; Helin, P.; Du Bois, B.; Leysens, K.; Severi, S.; Deshpande, P.; Baets, R.; Van Dorpe, P. Low-Loss Singlemode PECVD Silicon Nitride Photonic Wire Waveguides for 532900 nm Wavelength Window Fabricated Within a CMOS Pilot Line. *IEEE Photon. J.* **2013**, *5*(6), 2202809.
- [11] Romero-García, S.; Merget, F.; Zhong, F.; Finkelstein, H.; Witzens, J. Silicon nitride CMOS-compatible platform for integrated photonics applications at visible wavelengths. *Opt. Express* **2013**, *21*(12), 14036–14046.
- [12] Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press, New York, 2006.
- [13] Purcell, E. M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **1946**, *69*, 681.
- [14] Pelton, M. Modified spontaneous emission in nanophotonic structures. *Nature Photon.* **2015**, *9*, 427-435.
- [15] Klingshirn, C.F. *Semiconductor Optics, 4th ed.*; Springer-Verlag Berlin, Heidelberg, 2012.
- [16] Kong, J.A. *Electromagnetic Wave Theory*; EMW Publishing, Cambridge, Massachusetts, USA, 2008.
- [17] Long, D.A. *The Raman Effect*; John Wiley & Sons Ltd, Chichester, UK, 2002.
- [18] Dhakal, A.; Wuytens, P.C.; Peyskens, F.; Subramanian, A.; Skirtach, A.; Le Thomas, N.; Baets, R. In *Nanophotonic Lab-On-A-Chip Raman Sensors: a Sensitivity Comparison with Confocal*

Raman Microscope, Proceedings of BioPhotonics 2015 Conference, Florence, Italy, May 20–22, 2015.

- [19] Dhakal, A.; Raza, A.; Peyskens, F.; Subramanian, A.Z.; Clemmen, S.; Le Thomas, N.; Baets, R. Efficiency of evanescent excitation and collection of spontaneous Raman scattering near high index contrast channel waveguides. *Opt. Express* **2015**, *23*(21), 27391–27404.

5

Fabrication and Characterization of Nanophotonic-Plasmonic Chips

5.1 Introduction

In a world where miniaturization is becoming increasingly important, there is a booming need for reliable tools to fabricate micro- and nanodevices. Well-established optical lithography techniques such as deep-UV lithography use a pre-defined mask in combination with 193 nm or 248 nm exposure wavelengths. [1] The advantages of such a technique consist of the large field size (many devices can be patterned on the mask), large amount of chips (thousands of samples can be fabricated in a short time) and the capability of handling complexity very well compared to other methods. [1] On the downside, optical lithography is diffraction limited by the exposure wavelength λ_{exp} and optics (e.g. the numerical aperture NA of the projection lens). The resolution is then roughly proportional to λ_{exp}/NA . The standard SiN chips, discussed in the previous chapter, were fabricated using such a deep-UV lithography process. Hybrid integration of SiN chips and nanoplasmonic antennas however requires to push the resolution towards the nanometer scale (1 – 10 nm range), because the typical dimension of nanoplasmonic antennas is smaller than 100 nm. Electron beam or e-beam lithography (EBL) is capable of such high resolution

(with electron beam spots of only a few nanometers). In this chapter we will first introduce some standard concepts in e-beam lithography and subsequently present an overview of the different fabrication schemes. Furthermore, we outline on the optical characterization of such devices. Such a characterization finally allows to propose an optimal fabrication strategy. The investigation of several e-beam strategies for the hybrid integration of SiN waveguides and nanoplasmonic antennas and the design of a setup to characterize on-chip plasmon resonances were key parts of this PhD work.

5.2 E-beam lithography

E-beam lithography allows the direct writing of structures down to sub-10 nm dimensions by exposing an electron sensitive resist with a highly focused electron beam such that the solubility of the resist is dramatically changed upon exposure. In the subsequent development step, the exposed areas will either dissolve or remain depending on whether it is a positive or negative resist respectively. There are two main EBL strategies: projection printing and direct writing. In projection printing, a large e-beam pattern is projected parallel through a mask onto a resist-coated substrate by using a high-precision lens system. In direct writing, a small e-beam spot is moved with respect to the wafer to expose the wafer one pixel at a time. This eliminates the expensive and timeconsuming production of masks [3], but also reveals the inherent drawback of direct writing, being the low throughput as a result of the serial way of writing structures. All fabrication described in this chapter relies on direct writing. Several distinct features of EBL, taken from [2–4], are outlined below in order to gain understanding in the effect they have on the final e-beam pattern.

5.2.1 Forward and backward scattering of an electron beam

When electrons enter the resist, they undergo a series of low energy elastic collisions, each of which will slightly deflect the electrons (see Figure 5.1). This forward scattering broadens the beam by an amount that increases when the resist gets thicker. Moreover it is more pronounced at low incident energies. Since beam broadening results in lower resolution, most e-beam exposures are performed with high-energy electrons (between 50 and 100 keV). In addition to forward scattering, there is also backscattering. Typically, most electrons pass entirely through the resist and penetrate

deeply into the substrate. A fraction of those electrons will eventually undergo sufficient large angle collisions to re-emerge into the resist at some distance from the point at which they left it (see Figure 5.1). At higher energies, these backscattered electrons may cause exposure which is microns away from where the beam entered. This leads to the so-called proximity effect, meaning that electrons which write a feature at one location can increase the exposure at a nearby feature, causing pattern distortion and overexposure. Backscattering can be minimized by exposing on a thin membrane substrate or using a low exposure energy (between 2 and 20 keV) as this limits the penetration depth of electrons. In the latter case, electrons will lose (almost) all their energy in the resist layer, which increases the exposure efficiency (or the sensitivity) and the throughput of the lithography tool. As a result, the proximity effect is also reduced due to the minimization of electron scattering from the substrate. This eliminates the need to implement expensive and time consuming computer algorithms to correct for the effect. Despite these advantages, there are several drawbacks when using low acceleration voltages. At low energy the beam size is increased, resulting in lower resolution. Furthermore one needs to use very thin resist layers, as for a thick resist layer it is possible that electrons expose only a fraction of the resist due to the reduced penetration depth at low voltage. During the development process, the unexposed underlying resist layer might be dissolved leading to a complete patterning failure. [2]

Secondary electrons are low energy electrons (a few to a few tens of eV) produced by ionizations resulting from inelastic collisions by the primary incident electrons. Because of their low energy, secondary electrons have a short travel range (several nanometers). Nevertheless, they may ultimately limit the best possible resolution due to unwanted exposures at locations away from the original writing path. [2]

A final issue is electrostatic charging, particularly when writing onto an insulating substrate. If there is no pathway for the absorbed electrons to dissipate, charge will build up and defocus the electron beam. In such cases, a thin metal or conductive polymer layer is required above or below the resist in order to avoid charging. [2] In order to achieve high resolution we hence require:

- high acceleration voltage
- very thin resist
- no electric charging of the sample

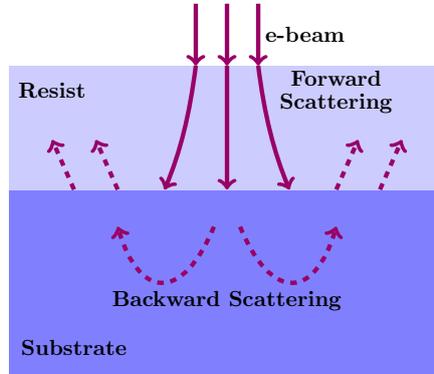


Figure 5.1: Electron scattering: forward scattering during propagation in the resist and backward scattering due to large angle collisions in the substrate, after which the electrons expose the resist again. Adapted from [2].

5.2.2 Electron Dose

Electron dose (usually expressed in $\mu\text{C}/\text{cm}^2$) is the number of electrons per unit area required to achieve the desired chemical response in the resist. [3] Each lithographic process has an optimum dose which represents the dose at which the measured linewidth after development is equal to the designed linewidth. However, for the highest resolution (i.e. sub-10 nm gaps) it is still very hard to achieve an exact correspondence between obtained and designed feature size as it depends on too many parameters. Exposing a pattern correctly usually requires a preliminary test exposure, referred to as a dose test. In this test, the same pattern is exposed at different doses and after development the optimum dose can be determined with a suitable inspection tool (such as a scanning electron microscope SEM, atomic force microscope AFM, optical microscope, etc). Exposing at lower doses compared to the optimum dose (underexposure) results in smaller pattern widths as compared to the design widths, but the probability for pattern irregularities also increases. Overexposure, which is obtained with a dose higher than the optimum dose, results in a widening of the pattern size. In general, the electron dose D should be low in order to achieve high sensitivity and throughput. [3] The estimated time T_{exp} it takes to write a pattern with surface area A (in cm^2) and a given current I (in Ampère) is

$$T_{exp} = \frac{AD}{I}, \quad (5.1)$$

hence the higher the dose D , the longer it takes to write the pattern. [4] The sensitivity of an e-beam resist is defined as the dose for which all the resist is removed after development. [3]

5.2.3 Shot pitch

The size of a pixel in an e-beam system is roughly determined by the square of the shot pitch a (see Figure 5.2). One can see that for a smaller shot pitch (i.e. increasing number of pixels for a given length), the feature definition will be better and the line edge roughness (LER), which is the fluctuation on the pattern linewidth, will be smaller. The shot pitch a depends on the scanning frequency f_{el} of the electronics, the current I and the dose D and can be calculated using the exposure time for one pixel, $t = Da^2/I$. The exposure time t moreover equals the inverse of the scanning frequency $t = 1/f_{el}$, hence

$$a \approx \sqrt{\frac{I}{f_{el}D}}. \quad (5.2)$$

So if the dose D is too low (for a given current I), then this will limit a and hence also the eventual roughness on the patterned features. [4]

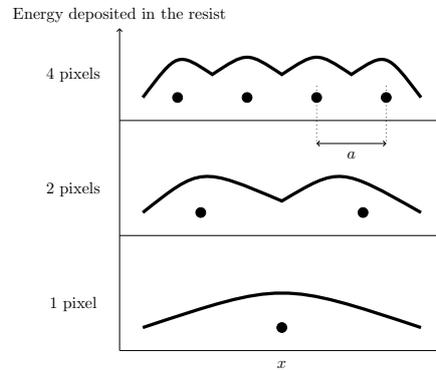


Figure 5.2: Effect of the shot pitch a on the line edge roughness. Adapted from [4].

5.2.4 Field stitching

Another important parameter is the size of the writing field, which is defined by the maximum deflection range of the electron beam. This size

is related to a . For the highest resolution, our e-beam system has a $128 \mu\text{m} \times 128 \mu\text{m}$ writing field for a 2 nm grid (so $a = 2 \text{ nm}$). For structures which don't require the highest resolution, the grid can be changed to 10 nm and the writing field is then $640 \mu\text{m} \times 640 \mu\text{m}$. Whenever we have patterns which extend over a range larger than the size of the writing field, the sample stage has to be moved mechanically. Such movements need to be very accurate in order to precisely stitch consecutive writing fields. In Figure 5.3(a) we depict a typical mask design, with several photonic components (in blue) while the writing fields are depicted by the green squares. At the border of each of these squares the stage has to move, which can result in a so-called stitching error due to imperfect matching between the two fields. When writing long waveguides (with lengths exceeding $640 \mu\text{m}$), this can result in kinks in the waveguide as depicted in Figure 5.3 (b). These stitching errors can influence the device performance; e.g. increased waveguide losses due to the kinks. When designing, it is hence important to limit the amount of stitching borders in order to avoid unwanted device deterioration as a result of stitching errors.

5.3 Scheme for nanoplasmonic-photonic patterning

In the previous section we outlined some important factors that one needs to take into account when fabricating chips with EBL. Since the highest SERS enhancements are expected for narrow-gap structures, it is important to achieve an as high as possible resolution when fabricating the nanoplasmonic antennas. High resolution typically requires the use of very thin e-beam resists ($< 100 \text{ nm}$). Since the height of all our single mode SiN waveguides is 220 nm, the required resist thickness is hence much smaller than the waveguides itself. Spincoating such thin resist layers on a sample with pre-defined waveguides of 220 nm high is not preferable due to possible non-uniform resist spreading which can distort the resulting antenna pattern. So instead of defining nanoplasmonic antennas on top of standard wafers which already contain dielectric waveguides produced by deep-UV lithography, the antennas are patterned first on a slab wafer after which the waveguides are etched around the antennas. In this way the very thin resist layer can be spun on a flat substrate, avoiding non-uniformities. Since also the waveguides need to be defined through EBL, the complete fabrication procedure consists of 2 consecutive e-beam steps. As discussed above, the plasmonic structures are defined in the first step.

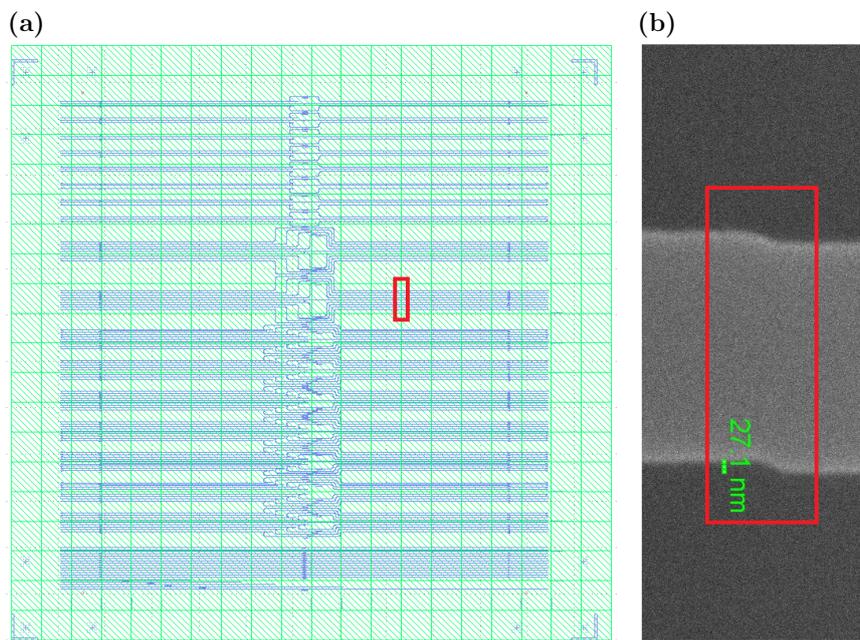


Figure 5.3: (a) Typical example of an e-beam mask. The green squares define the writing fields. At the border of each writing field (denoted by the red rectangle) there is a possibility to introduce stitching errors. (b) SEM image of a waveguide taken at the position of a stitching border. At this border, the waveguide exhibits an indentation of about 27 nm which will result in increased waveguide losses.

Alignment markers, used to align the second e-beam step, are also patterned in this first step. The waveguides are then defined around the antennas in the second e-beam step, whereby their relative position is fixed by the alignment markers. While different resist types were tested to define the nanoplasmonic structures, we repeatedly used the same procedure to pattern the SiN waveguides. We will first describe the different strategies to pattern the nanoplasmonic antennas and subsequently outline the SiN patterning.

5.3.1 Patterning plasmonic antennas

5.3.1.1 FOx-12

Initially, the negative tone hydrogen silsesquioxane (HSQ) resist FOx-12 (from Dow Corning Corporation) was used. HSQ is a serious candidate for high resolution EBL due to its small line edge roughness, high etching resistance and small molecular size. After e-beam exposure, HSQ has an amorphous structure similar to SiO₂, hence the name FOx or Flowable Oxide. [3] The processing steps are depicted in Figure 5.4.

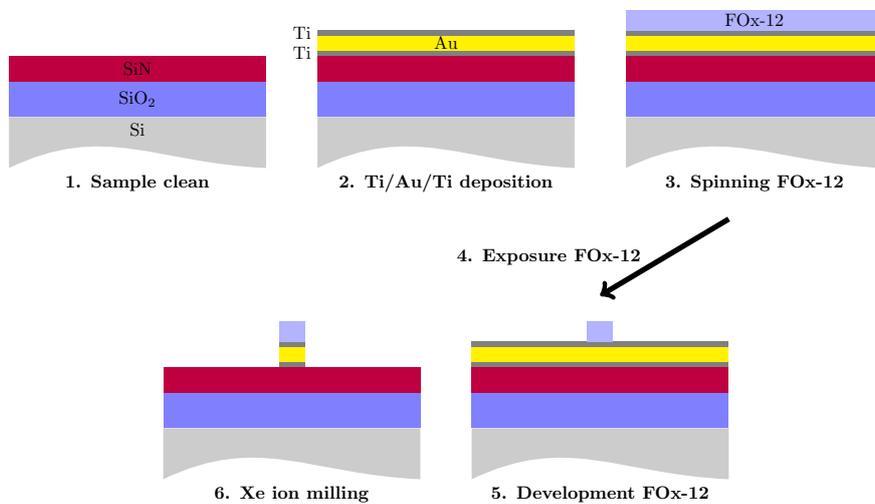


Figure 5.4: Fabrication of nanoplasmonic antennas using FOx-12 resist.

The initial step of every fabrication process consists of a thorough sample clean, which in our case consists of 3 consecutive steps. At first, the

samples are treated in an ultrasonic acetone bath for 30 min in order to remove tiny particles that might be stuck on the sample (e.g. resulting from cleaving). Subsequently, the samples are rinsed with acetone and IPA and dried with an N₂ gun. Alternatively one can immerse them in a hot acetone and IPA bath, although this was found to be unnecessary. Finally the samples are immersed for two to five minutes in a piranha solution (4:1 H₂SO₄:H₂O₂) to remove any remaining organic material. After piranha immersion, the samples are rinsed with DI water before they are dried with an N₂ gun.

Before deposition of the metal stack, a 10 min O₂ plasma is performed in the deposition machine itself, such that the vacuum is not broken between the plasma treatment and the metal deposition. The O₂ plasma serves as a final cleaning step and improves the adhesion between the SiN and the metal. Then a Ti/Au/Ti stack (with respective thicknesses of 2 nm, 30 nm and 2 nm) is sputtered with a low power recipe in a commercial Pfeiffer Spider 630. By using a low power, one can increase the control on the metal layer thickness. The first Ti layer serves as an adhesion layer between the SiN and the Au because the plasma treatment does not suffice for this purpose. Adhesion problems are frequently observed when this extra Ti layer is not included. Au is also known to have a bad adhesion to SiO₂ and since FOx resembles SiO₂, a second Ti adhesion layer is included at the top surface of the Au as well. Without this layer the resist easily detaches during development, resulting in a complete removal of both the exposed and unexposed regions.

After spincoating FOx-12, the samples are exposed on a 2 nm grid with 1300 μC dose. Subsequently, the samples are developed for 1 minute in TMAH 25 wt% (Silicon Etch) and 10 seconds in TMAH 2.38 wt% (OPD262). Then they are immersed for at least 30 seconds in an overflow bath of DI water, rinsed with IPA and dried with N₂. The effect of different doses for a fixed development process is depicted in Figure 5.5 for nine different doses ($900 \mu C + 50(n - 1) \mu C$, $n = 1 \dots 9$). Dose 9 (corresponding to 1300 μC) was determined to be the best dose for all considered structures. As expected, the smaller dose results in pattern irregularities and one can see that underexposed samples lead to cigar-like structures instead of the desired bowtie antenna geometry. Furthermore the gap is too large for the smaller doses.

The final step of the plasmonic processing consists of a Xe ion milling of the metal stack. The negative resist has defined the areas which have to stay, while the remaining metal has to be etched away without creating

an underetch. Since ion milling is a physical, anisotropic etching process, it perfectly suits this requirement. A 2.4 sccm Xe flow is combined with a 2 sccm Ar flow and a 0.4 sccm He flow. The average etching rate is about 0.283 nm/sec for the given metal stack and fixed machine parameters. The Ar flow counteracts electrostatic repulsions of the charged Xe atoms in order to gain a better control on the ion beam shape. The He flow is used for cooling because heating effects during the milling can deteriorate the resist. To this end, the milling is also performed in 12 steps of 10 seconds each ($12 \times 10 \times 0.2833 \approx 34 = 2 + 30 + 2nm$), with 1 minute waiting time in between to allow the substrate to cool down and avoid resist deterioration. A zoom on a typical antenna structure obtained using the above processing is shown in Figure 5.6(a). The excellent reproducibility using FOx is depicted in Figure 5.6(b). The white rim at the edge of the antennas is ascribed to a redeposition of metal during the ion milling. [5] Before spincoating ma-N 2403 for the SiN etching (as will be described later on), an additional O₂ plasma clean can be performed to reduce processing induced contaminations.

While this process allows high resolution and reproducibility, it also suffers from two major drawbacks. The first one is the residual FOx-layer which covers the metal stack at the end of the processing. Even after milling, there is still a thin resist layer left. For chemical or biological sensing applications it is however important to have full access to the plasmonic structure because the field enhancement is strongest near the metal surface. Since FOx chemically resembles SiO₂, it is not easy to remove without harming the other structures on the chip. Dry etching using SF₆ and O₂ or wet etching using HF will also etch the SiN that is not covered by metal. This can have major implications on the top surface roughness and hence on the waveguide loss. HF will moreover etch Ti very quickly. This often results in complete detachment of the metallic nanoantennas. Secondly, the top Ti layer between the Au and the FOx layer is not desired as this can increase the non-radiative losses and broaden the plasmon resonance. Furthermore the high dose also results in longer writing times. In summary, the FOx-processing is characterized by the following features:

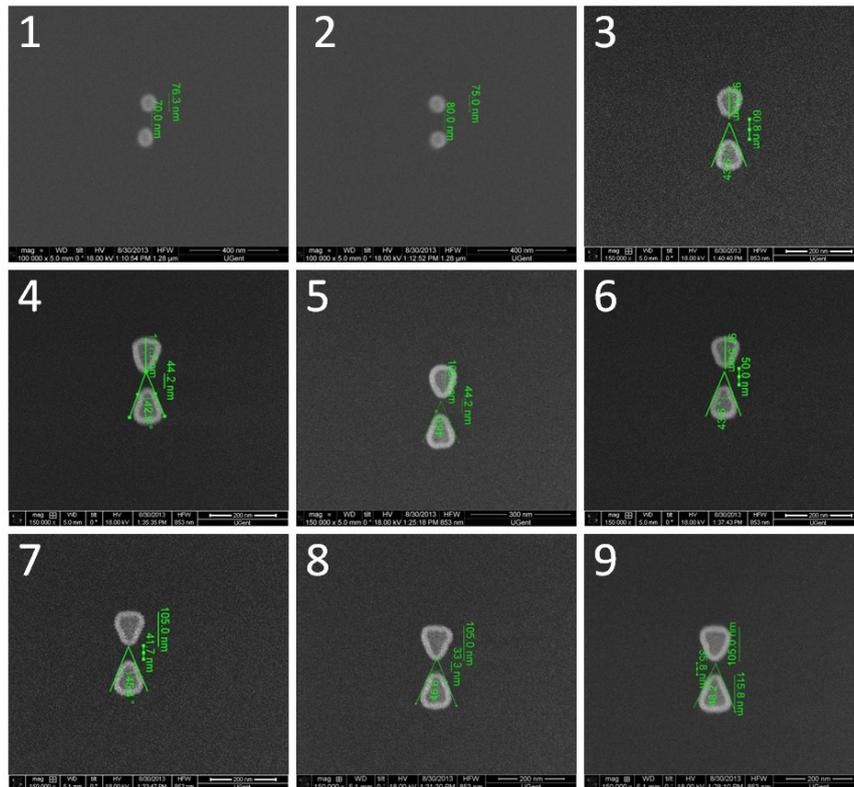


Figure 5.5: Dose test study: all samples are prepared in the same way, except for the exposure dose of the FOx-12. Sample n ($n = 1 \dots 9$) is exposed with a dose of $900 \mu\text{C} + 50(n - 1) \mu\text{C}$.

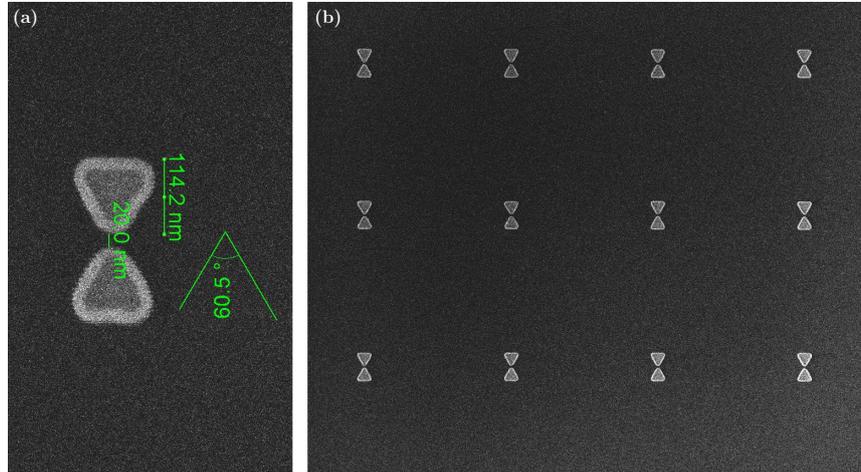


Figure 5.6: (a) Typical structure obtained through FOx-12 processing. (b) During a dosetest, large arrays of identical antennas are patterned in order to test the reproducibility. FOx-12 exhibits an excellent performance in this respect.

+	high resolution
+	high reproducibility and stable patterning
+	high etch resistance
-	remaining oxide-like layer on top of the antennas
-	top Ti adhesion layer
-	high exposure dose (so longer exposure time)

An alternative approach based on PMMA was subsequently tested to potentially circumvent the above drawbacks.

5.3.1.2 PMMA

PMMA is a positive tone e-beam resist which is frequently used for high-resolution patterning and cheaper as compared to FOx. The processing steps for the PMMA scheme are depicted in Figure 5.7. After the sample clean, which is identical to the clean described for the FOx-process, PMMA 950K (3% in chlorobenzene) is spincoated on top of the SiN and exposed on a 2 nm grid with a 280 μC dose. Note that due to the lower dose, the exposure time will also be shorter. A 1:1 mixture of MIBK:IPA is used as developer. The temperature of the developer ideally needs to stabilize for more than one hour in a temperature controlled bath before the actual development. The sample is then immersed in this thermally stable solution

(around 21° C) for 1 minute and is subsequently rinsed with IPA for about 30 seconds and dried with an N₂ gun. After development one needs to apply a short O₂ plasma (1 minute) to purify the surface because small remainders of PMMA can have a detrimental impact on the adhesion between the SiN and the metal stack. Nevertheless, if PMMA is exposed to an O₂ plasma for too long, it will be completely stripped. Immediately after this plasma treatment, a Ti/Au stack (with 2 nm/30 nm thickness) is sputtered on top. In the final step, the sample is immersed in acetone to lift-off the remaining PMMA, hence also removing the redundant metal. Since a lot of metal flakes are floating around during lift-off, there exists the possibility that metal particles sediment again on the sample. To partially avoid this, three beakers of acetone with decreasing concentration of metal flakes were used, such that the last beaker is almost flake-free. In this way, the first beaker removes the bulk part, after which the sample is immersed in the second and third (cleaner) beaker to avoid that metal flakes in the first beaker can sediment. After lift-off, the SiN waveguides are etched around the antennas as shown in Figure 5.3.

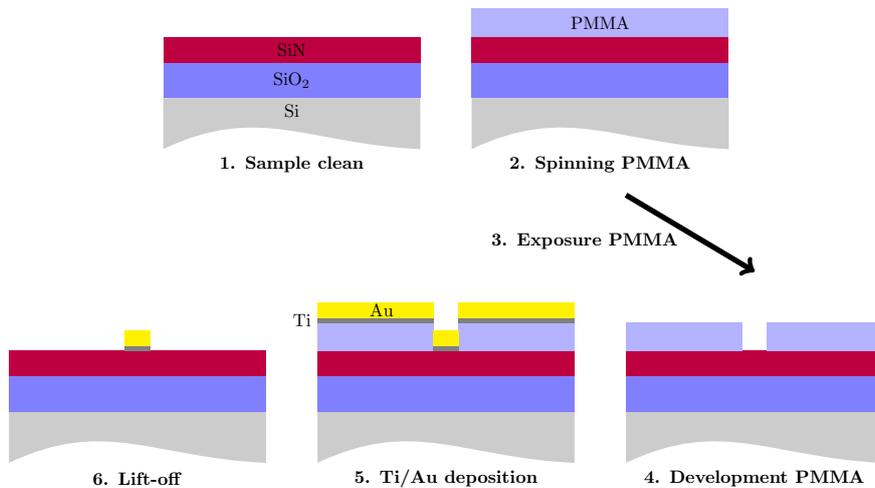


Figure 5.7: Fabrication of nanoplasmonic antennas using PMMA resist.

From the above process description it follows that the issues related to the FOx-processing can be avoided using PMMA. Since the metal structures are defined through lift-off there is no need for a top Ti adhesion layer such that the metal surface remains free of any other material. Furthermore, the

PMMA exposure can be performed at a much lower dose. Despite these advantages, some positive features of the FOx-process are lost as well. In Figure 5.8 we depict a typical bowtie antenna, obtained through the above processing. As compared to the bowtie antenna fabricated with FOx, the shape is much more irregular and the Au morphology is also more grainy. The irregular shape could be due to the short O₂ plasma used for the surface purification. The increased granularity results from the lift-off based metal deposition. As compared to the FOx-process, where the structures are milled in a large Au sheet, the PMMA structures are defined by depositing Au in nanometer size resist holes. This fill will never be perfect as becomes clear after lift-off. Furthermore it is very hard to achieve bowtie antennas with gaps smaller than 20 nm and the structures are also less reproducible as compared to the FOx-process. In conclusion, the PMMA-processing is characterized by the following features:

+	no top Ti adhesion layer
+	antenna surface is completely resist-free
+	low dose
±	lower etch resistance (however not required for this process)
±	moderate reproducibility
–	lower resolution as compared to FOx

Due to the loss of resolution and reproducibility, PMMA was not considered to be a proper alternative to FOx. Ultimately we would like to combine the high resolution and reproducibility of FOx with the easy removal and low dose of PMMA. To this end, a third (negative tone) resist ma-N 2400.6 was investigated, as described in the next subsection.

5.3.1.3 ma-N 2400.6

The negative tone resist ma-N 2400.6 is a thinner version of the ma-N 2403 resist. The fabrication process is almost identical to the FOx-processing, except for two important differences. While a top Ti layer is necessary to maintain a good adhesion between the metal stack and FOx, ma-N 2400.6 has a good adhesion to Au, eliminating the need for an additional adhesion layer. Secondly, ma-N 2400.6 can be easily removed with mr-Rem 400 remover. The full fabrication process is shown in Figure 5.9. After the sample clean, a Ti/Au stack is deposited and ma-N 2400.6 is spincoated on the sample. The optimal exposure dose was determined to be 300 μC . After

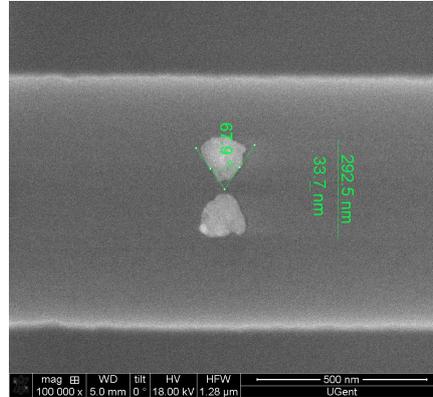


Figure 5.8: Typical structure obtained through PMMA processing. As compared to FOx-12 the shape is more irregular and the granularity of the gold is also increased.

exposure, the samples were developed in ma-D 525 for 30–35 seconds, subsequently immersed in a DI overflow bath for at least 40 seconds and then dried with an N₂ gun. The ion milling is identical to the FOx ion milling.

A typical bowtie antenna obtained via the above processing is depicted in Figure 5.10. Compared to PMMA, one can again achieve reproducible gaps of about 10 nm. The added advantages over FOx are the absence of a top Ti layer, the easy removal and the low exposure dose. Furthermore it should be noted that this process only requires one developer (ma-D 525) and one remover (mr-Rem 400) because the same developer and remover is used for the SiN patterning as well (see next section). This makes the overall process also quite cost-efficient. In conclusion, the ma-N 2400.6 processing has the following characteristics:

+	high resolution
+	high reproducibility and stable patterning
+	high etch resistance
+	antenna surface is completely resist-free
+	no top Ti adhesion layer
+	low dose

It is clear that ma-N 2400.6 combines all positive features of FOx and PMMA without suffering from their respective drawbacks. The ma-N

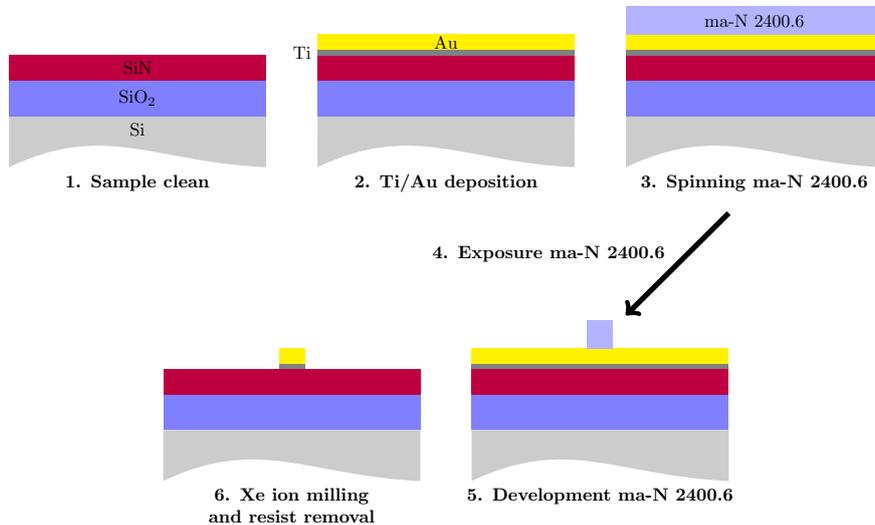


Figure 5.9: Fabrication of nanoplasmonic antennas using ma-N 2400.6 resist.

2400.6 processing is hence considered to be the optimal one for fabricating integrated nanoplasmonic antennas.

5.3.2 Patterning SiN waveguides

The procedure for patterning the SiN waveguides is shown in Figure 5.11. Mind that at this stage the metallic antenna (marked in yellow) is already patterned in the first e-beam step. The dielectric slab consists of a Si substrate, an SiO₂ undercladding and a 220 nm PECVD deposited SiN layer. A negative e-beam resist (ma-N 2403) is used in combination with a 10 nm grid exposure to define the 700 nm wide SiN waveguides. Since SiN is an insulating material, charging effects are possible. These can result in beam deflections as described above. In order to counteract charging, a thin (≈ 1 nm) e-spacer layer is spun on top of the ma-N 2403. This e-spacer is a conductive polymer which can be washed away with DI water after exposure. It is important to remove the e-spacer before the exposed resist is developed, as it can affect the developed profile. Positions where the e-spacer was not washed away, were found to have an increased roughness. After e-spacer removal, the sample is dried with N₂ before the resist is developed in ma-D 525 during 150 seconds. During development, the sample has to be gently stirred every 30 seconds. A 120 seconds immersion

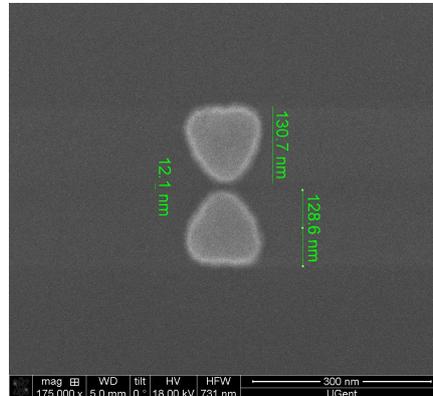


Figure 5.10: Typical structure obtained through *ma-N 2400.6* processing. As with *FOX-12*, the white rim originates from metal redeposition during the ion milling. These samples again have a reduced granularity as compared to PMMA.

in a DI overflow bath eventually stops the development. The sample is then etched in a commercial ICP plasma etcher (Oxford Plasmalab 100), using a mixture of 90 sccm C_4F_8 and 30 sccm SF_6 . In a final step the remaining resist is stripped with mr-Rem 400 remover in order to obtain a clean waveguide surface. An additional oxygen plasma clean can be useful as well.

5.4 Optical characterization

5.4.1 Waveguide loss

Up till now we only considered the structural and morphological aspects of the nanoantennas in determining the optimal fabrication strategy. It is however also important to minimize the waveguide losses. While the waveguides are always fabricated in the same way in all three schemes, the steps preceding the SiN etch are different for positive and negative resist. These differences could affect the eventual waveguide loss. It is important to note that we define waveguide loss as loss due to propagation in the SiN, hence it does not incorporate losses induced by the metallic nanostructures on top. Waveguide losses are typically measured using a vertical setup as shown in Figure 5.12. In a vertical coupling setup, optical fibers are placed above the chip under a certain angle. A grating coupler on the chip then

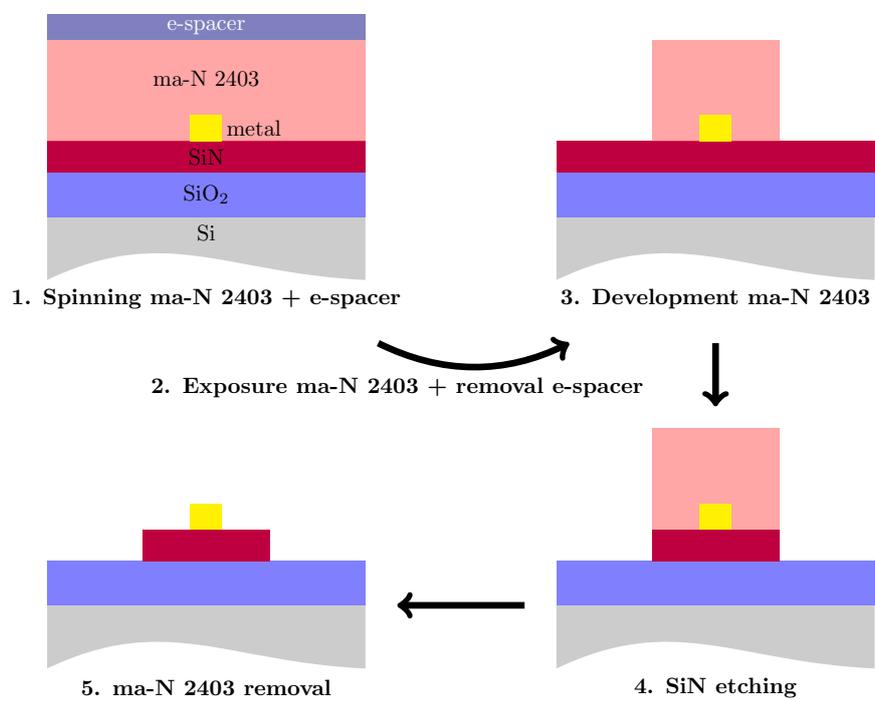


Figure 5.11: Fabrication steps for etching a SiN waveguide.

allows to re-direct the light from the fiber into the guided modes of the waveguide. The tunable Ti:saph laser emits polarized light at wavelength λ into the input optical fiber. A polarization rotator (PR) is then used to rotate the polarization and optimize the transmission since the grating couplers are designed for a specific polarization. Subsequently the fiber is aligned with the input grating coupler of the chip and launches light into the fundamental TE-mode of the waveguide. The light is finally coupled out of the chip and detected with a power meter PM.

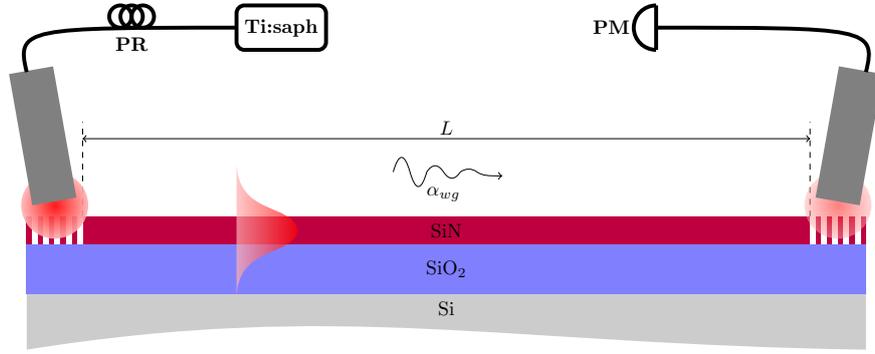


Figure 5.12: Vertical setup to measure waveguide losses. Ti:saph: laser; PR: Polarization rotator; α_{wg} : waveguide loss; L : length of the waveguide; PM: power meter.

For a given laser input power $P_{in}(\lambda)$, the detected power $P_{out}(\lambda)$ after propagating over a length L is

$$\frac{P_{out}(\lambda)}{P_{in}(\lambda)} = t_{in}(\lambda)t_{out}(\lambda)t_{sys}(\lambda)e^{-\alpha_{wg}(\lambda)L}, \quad (5.3)$$

where α_{wg} is the waveguide loss (in 1/cm), t_{in} and t_{out} the transmission of the input and output grating coupler respectively and t_{sys} the overall transmission of the fibers, PR and PM. All quantities are evaluated at the probing wavelength λ . In order to determine α_{wg} , each chip contains several reference waveguides of different length L . A reference waveguide is a bare SiN waveguide which contains no plasmonic antennas at all. Usually a logarithmic scale is used to express absolute power values (in dBm) or power differences (in dB). For notational simplicity we will omit the argument λ furtheron. Absolute power values in dBm P^{dBm} are

calculated using the following equation:

$$P^{dBm} = 10 \log_{10}(P). \quad (5.4)$$

In the above expression, P is always normalized to 1 mW. A power ratio on a linear scale translates to a power difference on a logarithmic scale since

$$10 \log_{10} \left(\frac{P_{out}}{P_{in}} \right) = 10 \log_{10}(P_{out}) - 10 \log_{10}(P_{in}) = P_{out}^{dBm} - P_{in}^{dBm}. \quad (5.5)$$

Hence equation (5.3) can be re-written according to

$$P_{out}^{dBm} - P_{in}^{dBm} = L_{in} + L_{out} + L_{sys} - \alpha_{wg} \frac{10}{\ln(10)} L, \quad (5.6)$$

where $L_{in} = 10 \log_{10}(t_{in})$, $L_{out} = 10 \log_{10}(t_{out})$ and $L_{sys} = 10 \log_{10}(t_{sys})$ are losses associated to the in-coupling, out-coupling and remaining system optics respectively. The input power P_{in}^{dBm} and system loss L_{sys} are constant and it is reasonable to assume that also the input and output coupling losses are constant over the chip area. As such, the output power only depends on the length L of the waveguide

$$P_{out}^{dBm}(L) = -\alpha_{wg} \frac{10}{\ln(10)} L + \underbrace{(P_{in}^{dBm} + L_{in} + L_{out} + L_{sys})}_{\text{constant}}. \quad (5.7)$$

By fitting a linear curve through the set of datapoints $(L, P_{out}^{dBm}(L))$, one can obtain the waveguide loss α_{wg} as the slope of the fitting line. Below we will list some important guidelines to minimize the waveguide losses. The loss is always calculated according to the above scheme at $\lambda = 785$ nm (as this will be the pumping wavelength for all our Raman experiments).

Several factors which might influence the loss were investigated for the FOx-12 scheme. During our initial e-beam runs, huge waveguide losses (close to 40 dB/cm) were reported. This was ascribed to a very thin metal layer remaining after the ion milling. In order to avoid this problem, it is recommended to slightly over-etch during the milling such that all metal is removed. This can be checked by measuring the resistance of the sample with a multi-meter. Since SiN is insulating, the multi-meter should have an Open Loop reading. If not, then there is still a residual metal layer which should be milled away before further processing. In itself, the over-etch could be a source of increased waveguide loss because it can potentially increase the top surface roughness of the SiN. In order

to investigate whether the milling introduces extra losses, we fabricated samples with and without ion milling. The waveguides of both samples were etched at the same time to avoid differences resulting from e.g. altered gas flows in the ICP etcher. Apart from loss characterization using a vertical setup, Atomic Force Microscopy (AFM) images were taken to measure the top surface roughness. The AFM data of the milled and non-milled sample are shown in Figure 5.13(a-b) and Figure 5.13(c-d) respectively. The left column ((a) and (c)) contains the roughness of the SiO₂ while the right column ((b) and (d)) depicts the roughness of the SiN. The SiO₂ roughness is only determined by the SiN etch and is the same for both cases (0.85 versus 0.87 nm). The top SiN roughness is 1.12 nm for the milled sample and 1.65 nm for the non-milled sample. So the milling has no significant influence on the morphology of the top surface. In this particular case, the milling even produces a slightly smoother surface. The waveguide loss of the milled sample was $\approx 14.55 \pm 0.38$ dB/cm, while the non-milled sample had a loss of $\approx 20.02 \pm 0.74$ dB/cm.

Secondly, the effect of the e-spacer layer was investigated. Experimentally it has been determined that stitching errors are sometimes (so not always) more prominent when no e-spacer is used. Larger stitching errors result in increased waveguide losses, hence it is important to ascertain whether the use of an e-spacer has a positive effect on the loss. In both the milled and non-milled case, we fabricated samples with and without an e-spacer. The above mentioned loss values were obtained from samples without e-spacer while the samples with e-spacer had a waveguide loss of $\approx 14.23 \pm 0.81$ dB/cm (milled) and $\approx 19.34 \pm 0.36$ dB/cm (non-milled). Hence, the e-spacer has no considerable effect on the waveguide losses (at least for this particular batch of samples). It has to be noted however that an e-spacer generally performs optimally only when the samples are loaded into the e-beam system immediately after spin coating the e-spacer. This was not the case for our samples, so increased performance could be possible by reducing the time between spincoating and loading. Nevertheless, it is desired to use an e-spacer to avoid pattern distortion on insulating substrates (as discussed earlier).

Waveguide losses of 15 to 20 dB/cm are still too high for practical applications. Despite the thorough ‘wet’ cleaning (ultrasonic acetone + acetone/IPA + piranha), we noticed that an oxygen plasma clean is essential after the piranha clean. This was observed when switching from FOx-12 to PMMA. In the PMMA scheme, the resist is spincoated before

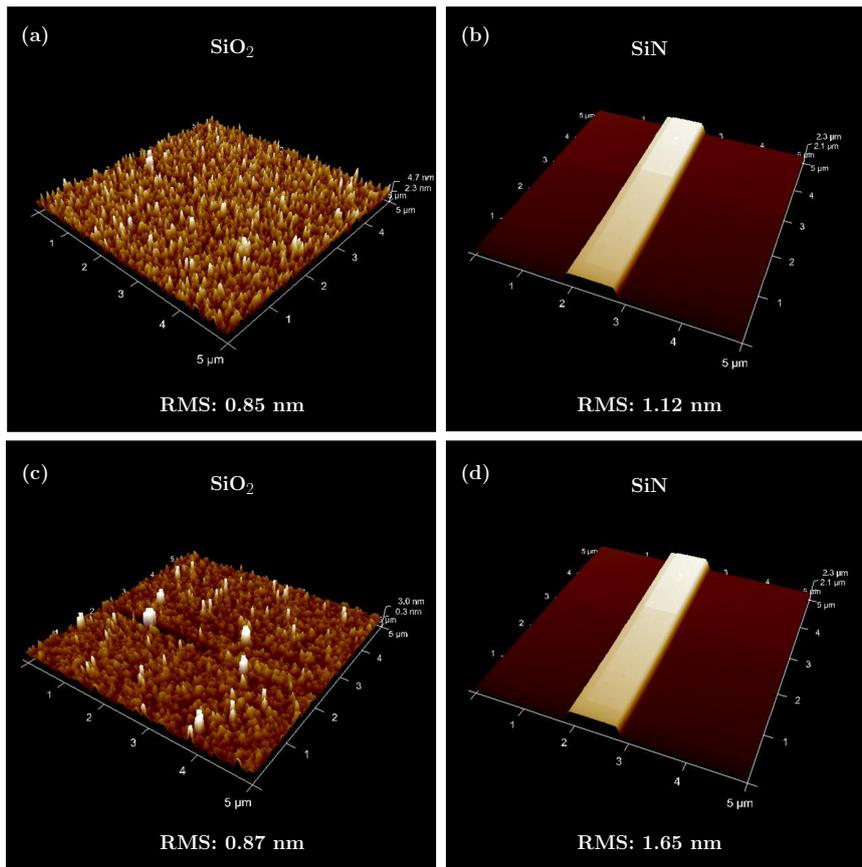


Figure 5.13: AFM roughness measurements of the SiO₂ and SiN. (a-b) RMS roughness with milling: (a) SiO₂ and (b) SiN. (c-d) RMS roughness without milling: (c) SiO₂ and (d) SiN. AFM measurements were performed by Ranjith Karuparambil Ramachandran (Ghent University).

metal deposition. Samples where PMMA was spincoated immediately after the piranha clean were compared to samples where PMMA was only spincoated after an additional oxygen plasma clean. The latter samples systematically showed lower waveguide losses, $\approx 5.08 \pm 0.81$ dB/cm compared to $\approx 20.28 \pm 1.78$ dB/cm; for the best sample a loss of $\approx 3.3 \pm 0.15$ dB/cm was measured. Just before spincoating ma-N 2403, a one to two minute oxygen plasma is applied as well in order to avoid contamination from previous steps. In the ma-N 2400.6 scheme, an oxygen plasma was applied just before metal deposition and also before spincoating ma-N 2403, resulting in samples with a waveguide loss of $\approx 3.18 \pm 1.6$ dB/cm. It seems that we cannot improve much more because the lowest measured waveguide losses were always around 3 dB/cm. The intrinsic SiN loss will however be lower, given the fact that there is a stitching border every $640 \mu\text{m}$ which contributes to local increases of the waveguide loss. Nevertheless, a value of 3 dB/cm is sufficiently low to perform on-chip SERS as will be shown in a later chapter.

5.4.2 Plasmon resonance measurement

Apart from low waveguide losses, it is also required that the nanoantennas exhibit a plasmon resonance in the proper wavelength region (700 – 1000 nm for our Raman-on-chip platform). The spectral behaviour of functionalized waveguides, i.e. waveguides that contain integrated nanoplasmonic antennas, is characterized with a horizontal coupling setup as shown in Figure 5.14. [6] As opposed to vertical coupling, where grating couplers are used to couple the light from the fiber to the chip, a horizontal coupling requires a cleaved chip facet to which a horizontally oriented fiber can be coupled.

An NKT EXR-4 Extreme Supercontinuum Source (SC) is coupled to a splitter module (NIRF), which transmits light in a wavelength band between 600 nm and 1120 nm. The fiber coupled (FC) light is then converted to a free space beam by an achromatic fiber collimator (C) and sent through a free space broadband polarizer (P) to convert the unpolarized beam into a TE-beam after which the light is fibercoupled again to a lensed fiber. The same configuration is used at the output in order to extract the TE component of the transmitted light. Since plasmon resonances are generally very broadband (> 150 nm), it is necessary to couple a broadband spectrum into the chip such that the spectral response can be measured at

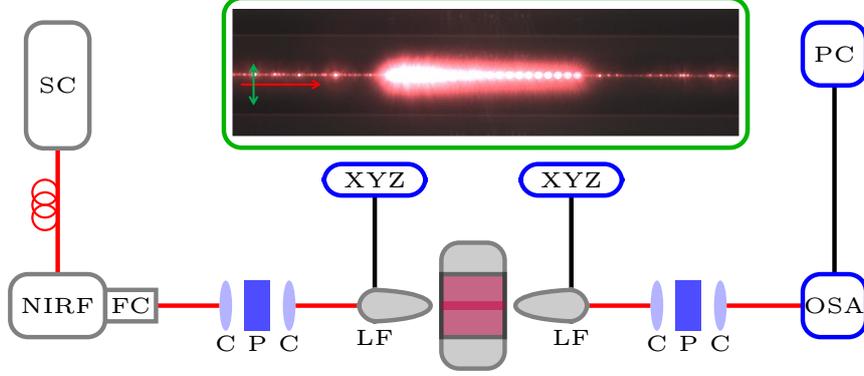


Figure 5.14: Horizontal setup to measure plasmon resonances. SC: supercontinuum source, NIRF: near-IR filter, FC: fiber coupling unit, C: achromatic fiber collimator, P: free-space broadband polarizer, LF: lensed fiber, XYZ: piezo controller stage, OSA: Optical Spectrum Analyzer, PC: OSA control using Python based measurement framework. Inset: scattering of an on-chip array of antennas. The red arrow depicts the propagation direction of the light while the green arrow represents the polarization state.

once. Contrary to vertical coupling through a grating coupler, which is quite narrowband (≈ 50 nm), a horizontally coupled lensed fiber can be used to focus broadband light onto the input facet of a cleaved chip. A piezocontroller (XYZ) is applied to achieve maximal coupling between the lensed fibers and the facets of the chip. Spectra are then recorded with an Advantest Optical Spectrum Analyzer (OSA) and read out with Python controlled (PC) software. The inset figure depicts the scattering of an array of antennas. Due to the loss induced by the metal, the scattering becomes less bright when light propagates along the array.

The total transmission loss at wavelength λ induced by one antenna, due to the plasmon resonance, is given by the antenna extinction $E(\lambda) \geq 0$ (in dB). This includes both absorption as well as scattering losses. The linear antenna extinction is then simply given by $e(\lambda) = 10^{E(\lambda)/10}$. By measuring the power transmission of a reference waveguide $T_{ref}(\lambda)$ [dBm] and of a waveguide functionalized with one antenna $T_{ant}(\lambda)$ [dBm], one can determine the extinction by

$$E(\lambda) = T_{ref}(\lambda) - T_{ant}(\lambda). \quad (5.8)$$

Usually the waveguide contains a periodic array of N antennas with spacing Λ , which is always large enough such that there is a negligible near-field

coupling between two consecutive antennas. In that case, the single antenna extinction is approximately given by

$$E(\lambda) \approx \frac{T_{ref}(\lambda) - T_{Nant}(\lambda)}{N} \quad (5.9)$$

where $T_{Nant}(\lambda)$ [dBm] is the power transmission of a waveguide functionalized with N antennas. The antenna array forms a periodic medium, so a diffractive coupling between the forward and backward propagating modes at discrete λ is expected. This will translate in a periodic ripple on the overall extinction. Since the light decays exponentially, the approximate formula will however give the correct experimental single antenna extinction. A typical transmission measurement is depicted in Figure 5.15. The plasmon resonance (blue) is clearly visible after applying equation (5.9) to the original data (red and green curve). The fringes superimposed on the broad envelope originate from the gold antennas because they do not appear in the reference waveguide spectrum. The distance between two fringes for this 1D grating with period Λ and group index n_g is given by $\lambda^2/(2n_g\Lambda)$ (around a given wavelength λ). In this particular case, the period equals $\Lambda = 10 \mu m$ and the group index is $n_g \approx 1.936$ (i.e. group index of the waveguide mode at $\lambda = 798$ nm). Theoretically a fringe spacing of 16.45 nm is expected, which is in close correspondence with the experimentally observed fringe spacing (≈ 17 nm for $\lambda = 798$ nm). A thorough analysis of plasmon resonance tuning will be included in the next chapter.

5.5 Conclusion

In this chapter we discussed several fabrication strategies for the hybrid integration of nanoplasmonic antennas and photonic waveguides. A two-step e-beam lithography process using ma-N 2400.6 and ma-N 2403 resist was found to optimal in terms of achieving high resolution and reproducibility, easy resist removal and low waveguide losses. Moreover we discussed the typical setup for measuring waveguide losses and outlined the procedure to determine on-chip plasmon resonances. In the next chapter we will analyze the resonance tuning of integrated nanoplasmonic antennas.

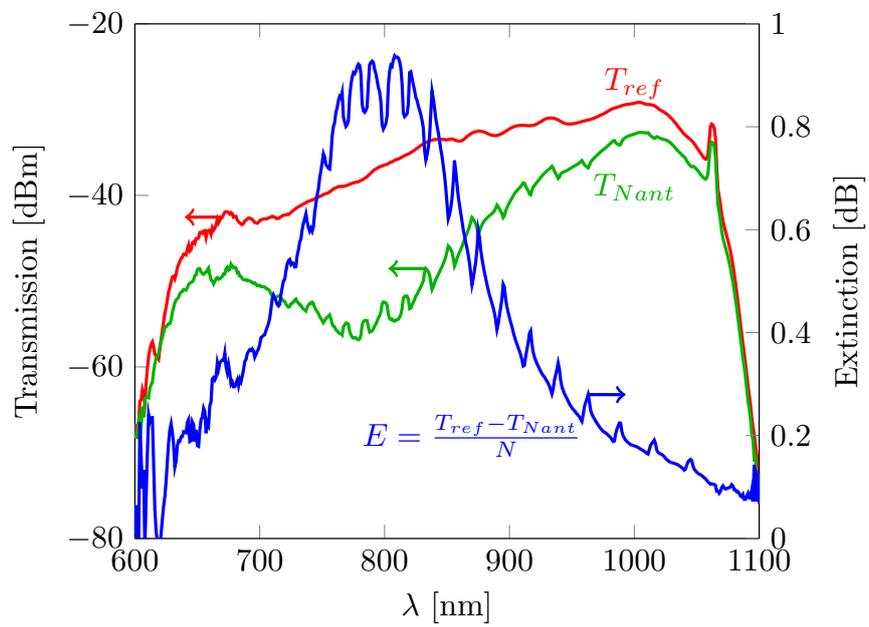


Figure 5.15: Typical transmission spectrum of a reference waveguide (red) and a waveguide functionalized with N antennas (green). From these two spectra, the single antenna extinction spectrum (blue) is extracted.

References

- [1] ePIXfab - The silicon photonics platform. http://www.epixfab.eu/index.php?option=com_content&view=article&id=57&Itemid=72 (accessed Sep 24, 2015).
- [2] Ali Mohammad, M.; Muhammad, M.; Dew, S. K.; Stepanova, M. Fundamentals of Electron Beam Exposure and Development. In *Nanofabrication - Techniques and Principles*; Springer-Verlag, Vienna, 2012; pp 11–41.
- [3] Grigorescu, A. E.; Hagen, C. W. Resists for sub-20-nm electron beam lithography with a focus on HSQ: state of the art. *Nanotechnology* **2009**, *20*, 292001.
- [4] JEOL JBX-9300FS Electron Beam Lithography System Training. http://nanolithography.gatech.edu/JEOL_JBX-9300FS_Training.pdf (accessed Sep 24, 2015).
- [5] Verellen, N.; Van Dorpe, P.; Huang, C.; Lodewijks, K.; Vandenbosch, G.A.E.; Lagae, L.; Moshchalkov, V.V. Plasmon Line Shaping Using Nanocrosses for High Sensitivity Localized Surface Plasmon Resonance Sensing. *Nano Lett.* **2011** *11*, 391-397.
- [6] Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Bright and dark plasmon resonances of nanoplasmonic antennas evanescently coupled with a silicon nitride waveguide. *Opt. Express* **2015**, *23*, 3088–3101.

6

Plasmon Resonances of Evanescently Coupled Nanoplasmonic Antennas

6.1 Introduction

In the previous chapters we have introduced conventional Raman spectroscopy and the rationale for using integrated Raman systems based on photonic integrated circuits. Despite the added value of integrated Raman systems, it is still difficult to detect very small amounts of analyte or single nanoparticles (or molecules) through spontaneous Raman spectroscopy. To this end it is desirable that the Raman signal can be further enhanced. The near-field coupling between closely spaced metallic nanoparticles can result in the formation of regions where the electric field is highly localized and intense (hotspots). These hotspots can be used to enhance extremely weak Raman signals through Surface Enhanced Raman Spectroscopy (SERS) as extensively described in a previous chapter. [1] Apart from their remarkable light enhancement capabilities, plasmonic nanostructures have attracted much interest in the world of biosensing due to their extreme sensitivity to local environmental changes. All these properties emerge as a result of the collective oscillation of conduction electrons (plasmons) upon light excitation. [1–4] The plasmon resonance wavelength can be tuned by varying the nanoparticle's size, shape, composition and

dielectric surroundings [1–12] and a broad range of biological and chemical applications stems from these spectral tuning capabilities. [13–18]

Most of the research has however focused on nanoplasmonic antennas on top of simple substrates (e.g. glass slides). In such cases, large and expensive free space optics is required for both excitation and collection of the signals. Recently, the integration of metallic nanoparticles on top of photonic integrated circuits (PICs) has been investigated. [19–24] In such a hybrid on-chip platform, the antennas are excited through the sub-wavelength evanescent tail of the waveguide mode. By fabrication, such antennas can be optimally aligned with the polarization of the input beam as well.

We further contributed to the emerging field of integrated nanoplasmonic antennas by conducting a comparative study of the resonance wavelength behaviour on different geometrical antenna parameters. Moreover, we demonstrated the integration of narrow-gap plasmonic nanostructures on a PIC. In [25] we present a theoretical and experimental study of three widely used antenna types (single rod, double rod and bowtie antenna), fabricated on top of a single mode silicon nitride SiN waveguide and resonant in the near-IR region (700-1000 nm). The near-IR region is of particular interest for on-chip Raman spectroscopy because pumping in the near-IR (e.g. at 785 nm) can reduce the unwanted fluorescence background and allows low-loss propagation of both pump and Stokes light in a SiN waveguide. [26–28] In this chapter we elaborate on how the antenna resonance is affected by geometrical tuning and how we can optimize the interaction between the evanescent excitation field and the nanoplasmonic antenna (see also [25]). Moreover we discuss the excitation of dark plasmon modes [25] and provide additional results that are relevant for on-chip SERS.

6.2 Hybrid nanophotonic-plasmonic chips

The platform used for the analysis in this chapter is depicted in Figure 6.1(a). It consists of a single mode SiN rib waveguide with an array of $N = 20$ nanoplasmonic antennas, organized in a 1D chain with period $\Lambda = 10 \mu m$, on top of the waveguide core. The bottom cladding is SiO₂ on a Si carrier substrate. The metal stack consists of a Ti adhesion layer (thickness t_{Ti}), a Au layer (thickness t_{Au}) and a top Ti layer (thickness t_{Ti}^{top}). On top of the metal stack, a residual FOX layer with thickness t_{FOX} remains after the processing. These chips were produced using a two-

step e-beam lithography procedure, consisting of a FOx-12 and ma-N 2403 step (for more details we refer to the previous chapter and to [25]). The antennas are excited with the fundamental TE-mode (green arrow) of the single mode waveguide. In [25] we investigated three different antenna geometries which are depicted in Figure 6.1(b-d). The single rod (SR) antenna with length H and width W , the double rod (DR) antenna with length H , gap Δ and width W and the bowtie (BT) antenna with length H , gap Δ and apex angle α . In the next sections we will investigate the plasmon resonance tuning of these three antenna types, but first we include some details on the simulation model.

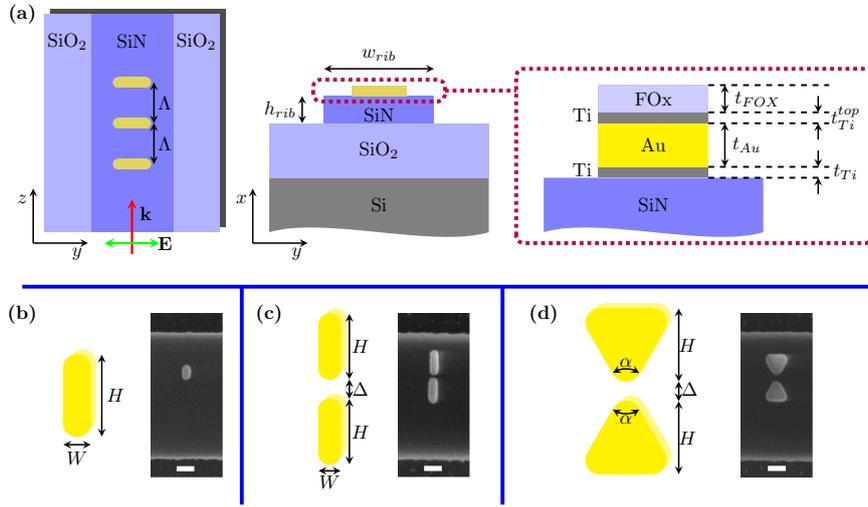


Figure 6.1: (a) Schematic top-view (left) and cross-section (right) of the waveguide geometry. Each waveguide is functionalized with an array of $N = 20$ antennas spaced by $\Lambda = 10 \mu\text{m}$. The antennas are excited with a TE-polarized (green arrow) waveguide mode with \mathbf{k} -vector along the z -direction (red arrow). The stack consists of a Si/SiO₂ substrate and a single mode SiN waveguide (height $h_{\text{rib}} = 220 \text{ nm}$ and width $w_{\text{rib}} = 700 \text{ nm}$). On top of the SiN, a Ti/Au/Ti layer is deposited (with thicknesses $t_{\text{Ti}}/t_{\text{Au}}/t_{\text{Ti}}^{\text{top}}$ respectively). A thin resist layer with thickness t_{FOX} remains at the end of the processing. (b-d) Investigated antenna geometries with their corresponding geometrical parameters and a SEM picture of the fabricated structure: (b) single rod (SR) antenna with length H and width W , (c) double rod (DR) antenna with length H , gap Δ and width W , (d) bowtie (BT) antenna with length H , gap Δ and apex angle α . The white scale bar in the SEM pictures equals 100 nm. [25].

6.3 Simulation model

In Figure 6.2 a 3D-view of the Lumerical simulation geometry is depicted. A SiN rib waveguide (darker blue) on an SiO₂ bottom cladding (light blue) is functionalized with one particular antenna geometry (yellow). The FOx layer on top of the antenna and a potential native oxide layer between the SiN and the nanoantenna [21] are also included. A TE modal source injects the fundamental TE mode. The transmitted power at the output of the waveguide is measured by a power monitor. The complete simulation region is surrounded by a PML boundary, which avoids unphysical reflections at the edge of the simulation region.

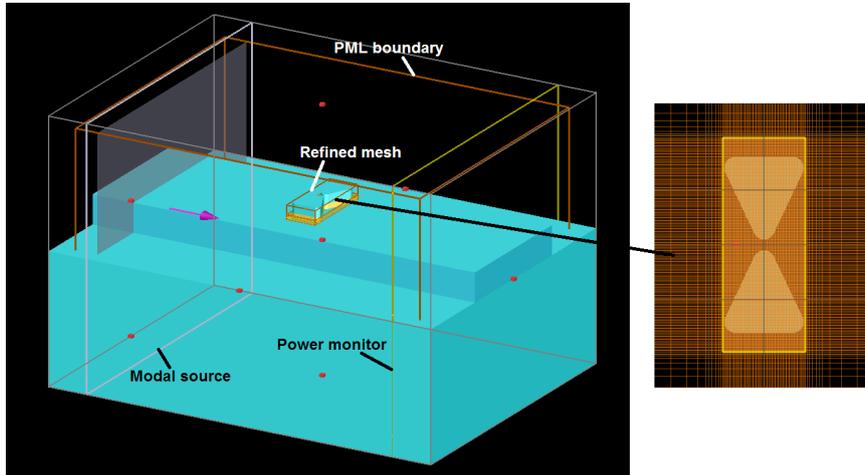


Figure 6.2: 3D-view of the Lumerical simulation geometry.

Usually, an electromagnetic FDTD simulation of a simple waveguide will give accurate results if there are at least 10 mesh cells per wavelength λ . For simulations in the near-IR (700–1000 nm) this means that the mesh size should not exceed 70 nm. However, if such waveguides are functionalized with nanoplasmonic antennas, then locally the 70 nm mesh will be too coarse since the gold structures (and the gaps inbetween them) have sizes of the order of 10 to 100 nm. Moreover, the electromagnetic fields near such nanoplasmonic structures are tightly confined to sub-wavelength regions much smaller than $\lambda/10$. As a result, it is necessary to locally refine the mesh around the antenna surface (see right part of Figure 6.2). Before

running long simulation sweeps, it is required to perform a convergence test in order to know for which meshing the physical quantities converge. Depending on the specific quantities, a mesh refinement between 0.5 and 2 nm is desired in the plane of the antenna. Furthermore the meshing always needs to resolve all refractive index changes. For example, the Ti layer is 2 nm in thickness so it has to be resolved with at least 1 nm meshing in the thickness direction.

Moreover it is important to select the FDTD simulation time T_{sim} judiciously. This time represents the physical time that the light would need to propagate over a certain distance. In order to capture all relevant physical effects (including possible reflections), T_{sim} hence needs to be large enough such that the light can travel along all structures within the simulation domain. Assume the waveguide has a length L and the waveguide mode an effective index n_{eff} , then a time Ln_{eff}/c will allow a single-pass forward propagation in a non-functionalized (without gold) reference waveguide. The simulation time hence needs to be at least $T_{sim} > Ln_{eff}/c$. In order to incorporate reflections, one needs to take integer multiples of Ln_{eff}/c .

All simulations are performed on a fixed waveguide structure (see Figure 6.1(a)) consisting of a SiN rib ($h_{rib} = 220$ nm, $w_{rib} = 700$ nm and $n_{rib} = 1.9$) on an SiO₂ bottomcladding ($n_{bclad} = 1.45$). The refractive index of the top cladding is $n_{tclad} = 1$ (air). The metal stack thicknesses ($t_{Ti} = t_{Ti}^{top} = 2$ nm and $t_{Au} = 35$ nm) are fixed for all antennas and a built-in refractive index model for Au (Johnson and Christy [29]) and Ti (CRC [30]) is used. The antenna region is meshed with a 2 nm grid and a mesh refinement to 1 nm is applied in the Ti region. The Si substrate is not taken into account in the simulation, but for the calculation of the excitation profile this is not necessary since the real oxide cladding is thick enough to avoid substantial power leakage to the Si. Single antenna extinction spectra $E(\lambda)$ (in dB) are calculated through $E(\lambda)[dB] = T_{ref}(\lambda)[dB] - T_{ant}(\lambda)[dB]$ in which $T_{ref}(\lambda) = 10 \log_{10}(t_{ref}(\lambda))$ is the power transmission of a non-functionalized reference waveguide while $T_{ant} = 10 \log_{10}(t_{ant}(\lambda))$ is the power transmission of a functionalized waveguide (see previous chapter). The interaction efficiency $\eta(\lambda)$ is defined as the fraction of the power taken away by a single nanoantenna, at wavelength λ , and is calculated through $\eta(\lambda) = 1 - t_{ant}(\lambda)/t_{ref}(\lambda)$. [24, 25]

6.4 Bright mode resonance tuning

In Figure 6.3 both the theoretical and experimental extinction spectra are shown. [25] Each spectrum is normalized with its own maximum and shifted with respect to the previous spectrum for improved visualization. Both the theoretical and experimental single rod (SR) spectra are shifted with steps of 0.5. The theoretical double rod (DR) and bowtie (BT) spectra for different H are shifted with 1.5 while for a fixed H the shift between different Δ curves is 0.1. For the experimental DR and BT spectra, the shift between different H curves is 2.5 and for a fixed H the shift between different Δ curves is 0.3.

The theoretical curves depict the response of a SR (a), a DR (b) and a BT antenna (c) respectively under changes in their geometrical parameters. In case of the SR and DR antennas, a fixed antenna width of $W = 55$ nm is chosen based on the SEM pictures of each antenna. For the BT antenna an apex angle $\alpha = 60^\circ$ is chosen. The resonance wavelength of a single rod antenna λ_{res}^{SR} redshifts with increasing rod length H with a sensitivity $S = d\lambda_{res}/dH$ given by $S^{SR} = 3.9$. The resonance wavelength of double rod λ_{res}^{DR} and bowtie λ_{res}^{BT} antennas shows a similar redshift with increasing length, but the sensitivity depends on the gap Δ . DR antennas are more prone to length changes as compared to SR antennas because they have an increased sensitivity, $S^{DR} = 4.9$ for $\Delta = 10$ nm and $S^{DR} = 4.1$ for $\Delta = 120$ nm. When the gap Δ increases, for a fixed length H , the resonance wavelength λ_{res} blueshifts and will converge for sufficiently large Δ since the individual gold particles decouple and the DR sensitivity will approach that of the SR antennas. BT antennas are less sensitive to length changes as compared to DR antennas: $S^{BT} = 4.4$ for $\Delta = 10$ nm and $S^{BT} = 3.6$ for $\Delta = 120$ nm. A similar blueshift for increasing Δ (and fixed H) is observed as well. The redshift for increasing H and blueshift for increasing Δ are also observed for non-integrated antennas. [12]

Figures 6.3(d-f) show the experimentally measured extinction spectra of SR (d), DR (e) and BT (f) antennas. The measured antennas were partially misaligned with respect to the center of the waveguide as can be seen in Figures 6.1(b-d). The measured SR sensitivity is $S^{SR} \approx 3.5$, which deviates about 10% from the theoretically calculated value. For the gapped structures we find $S^{DR} \approx 3.8$ and $S^{BT} \approx 3.34$ (both for a gap of ≈ 60 nm). These values deviate 12% and 12.4% respectively from their numerical

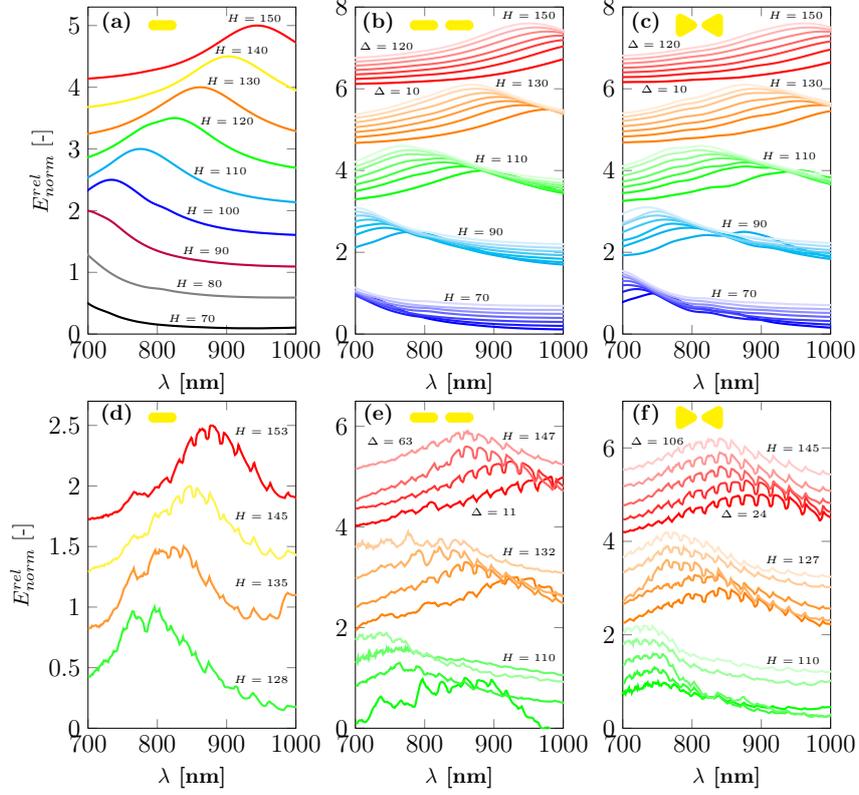


Figure 6.3: Relative Normalized Extinction Curves. The extinction spectra E_{norm}^{rel} are normalized with their own maximum and then shifted relatively with respect to each other for improved visualization of the resonance wavelength behaviour. **(a-c)** Theoretical spectra of single rod SR **(a)**, double rod DR **(b)** and bowtie BT **(c)** antennas. SR spectra are depicted for different lengths H (in nm) and fixed width $W = 55$ nm. For the DR and BT antennas, the spectra are plotted for five different lengths H (in nm). For each length, seven different gaps (10, 20, 30, 40, 60, 90, 120 nm) are simulated and depicted by the shaded curves in the corresponding length color. The outer right curve of each color corresponds with $\Delta = 10$ nm while the outer left curve corresponds with $\Delta = 120$ nm (see red set of curves). The width of the DR antennas is fixed to $W = 55$ nm and the apex angle of the BT antennas is $\alpha = 60^\circ$. **(d-f)** Experimental spectra of SR **(d)**, DR **(e)** and BT **(f)** antennas. Similar to the theoretical curves, the SR spectra are compared for different lengths while the DR and BT spectra are compared for different lengths and gaps. The lengths and gaps mentioned on the experimental subplots are averages calculated using the measured dimensions (with SEM). [25].

estimations. Given the uncertainty on the measured SEM dimensions, these values correspond well with the theoretically predicted trends.

For SR antennas we also (numerically) investigated the dependence on the width W , the gold thickness t_{Au} and the thickness of the Ti adhesion layer t_{Ti} , while keeping the other parameters constant. An increase in any of these parameters always results in a blueshift as can be seen in Figures 6.4(a-c). The effect of t_{Au} and W is much more pronounced as the effect of t_{Ti} .

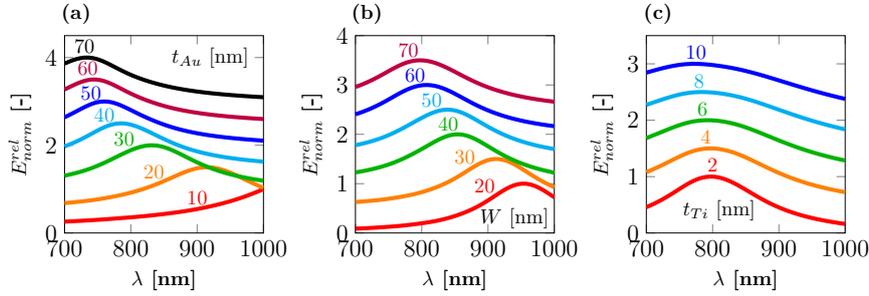


Figure 6.4: Theoretical normalized extinction spectra (shifted relative to each other for improved visualization) of a single rod antenna with fixed $H = 128$ nm. Spectra as a function of t_{Au} (a), W (b) and t_{Ti} (c).

In the above simulations we did not take into account a possible native oxide layer [21] between the SiN and the Ti adhesion layer. Such a layer can however have a serious impact on the exact position of the resonance wavelength. We numerically investigated this effect for a fixed SR antenna while the thickness of the native oxide layer t_{nox} was increased (see Figure 6.5(a)). One can see that the resonance wavelength blueshifts for an increasing t_{nox} . This can be understood because of the lower refractive index of oxide as compared to the index of SiN. Moreover, it is clear from the SEM pictures (see Figures 6.1(b-d)) that the antennas are not exactly centered on the waveguide. The effect of this misalignment Δ_y was also numerically calculated for SR antennas as shown in Figure 6.5(b). The misalignment clearly has a negligible impact on the precise position of the resonance wavelength.

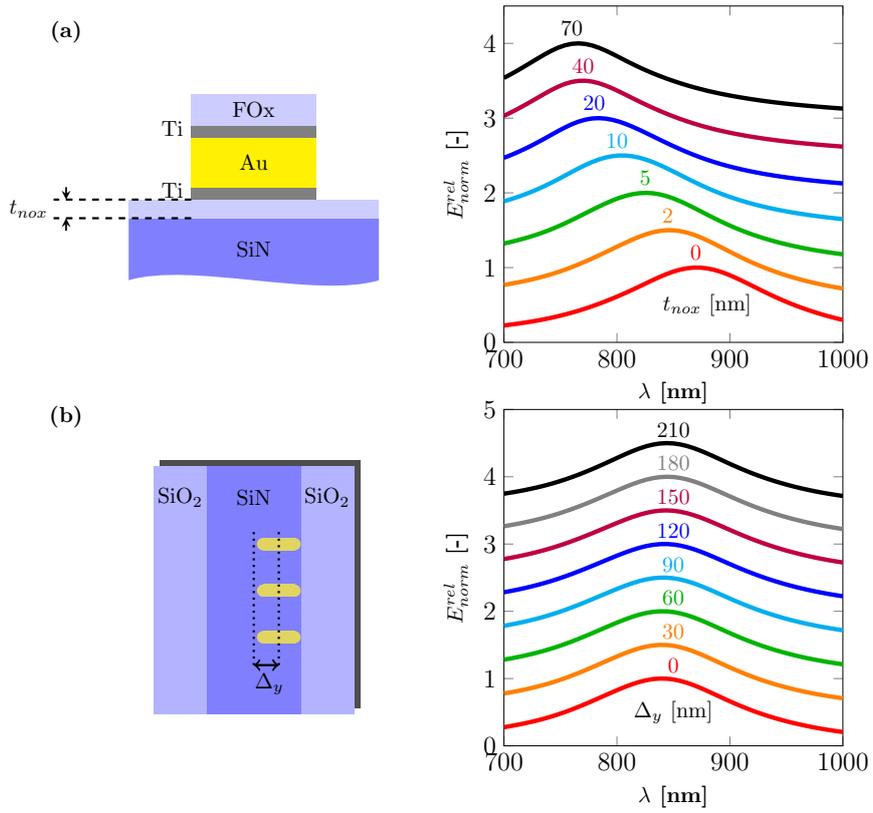


Figure 6.5: Theoretical normalized extinction spectra (shifted relative to each other for improved visualization) of a single rod antenna with fixed $H = 128$ nm and $W = 55$ nm. Spectra as a function of t_{nox} ((a)) and Δ_y ((b)).

6.4.1 Comparison between theory and experiment

In Figure 6.6 we present a comparison between the resonance wavelength tuning predicted by simulation and the experimentally measured resonance wavelengths for SR (Figure 6.6(a)), DR (Figure 6.6(b)) and BT (Figure 6.6(c)) antennas. In each subplot the resonance wavelength is plotted as a function of the antenna length, hence the slope of each curve represents the earlier defined sensitivity S . For the DR and BT antennas, each curve corresponds to a specific gap value. For a fixed length it is clear that the resonance wavelength blueshifts when the gap increases (for DR and BT antennas). The slope (sensitivity) of the experimental data points is in close correspondence with the numerically predicted values. The offset in the absolute position of the resonance wavelength could be ascribed to a variety of factors. As outlined before, the presence of a native oxide layer will blueshift the resonance. Since all numerical estimates are redshifted as compared to the experimental data, the incorporation of such a layer will already result in a better correspondence. Differences between the experimental refractive indices of Au and Ti (due to the specific metal deposition) and the indices used in the simulation or differences among the $N = 20$ antennas can also contribute to an increased deviation between the theoretical and experimental values. Moreover, an imprecise knowledge of the exact Au and Ti thickness also has an effect on the theoretical resonance wavelength (see Figures 6.4(a) and 6.4(c)). Finally, there will always be a SEM error on the measured antenna dimensions. An overestimation of the antenna length will e.g. associate a given resonance wavelength with a larger antenna length.

6.4.2 Q -factor of integrated nanoantennas

As shown in a previous chapter, the field enhancement of a nanoplasmonic antenna will increase when the Q -factor increases. The measured Q -factor, calculated using the 3dB-bandwidth of the extinction, of our fabricated devices ranges between ≈ 3.11 and ≈ 7.75 with a typical value of ≈ 5 at $\lambda_{res} = 780$ nm. For pure non-annealed Au antennas without Ti adhesion layers, Q -factors of ≈ 11 (around 780 nm) are experimentally measured. [31] Because we use a Ti adhesion layer on the bottom and top surface, the plasmon resonance will broaden and hence reduce the Q -factor as compared to a pure Au antenna. Our simulations show that a 2 nm Ti adhesion layer on the bottom and the top of the Au can easily

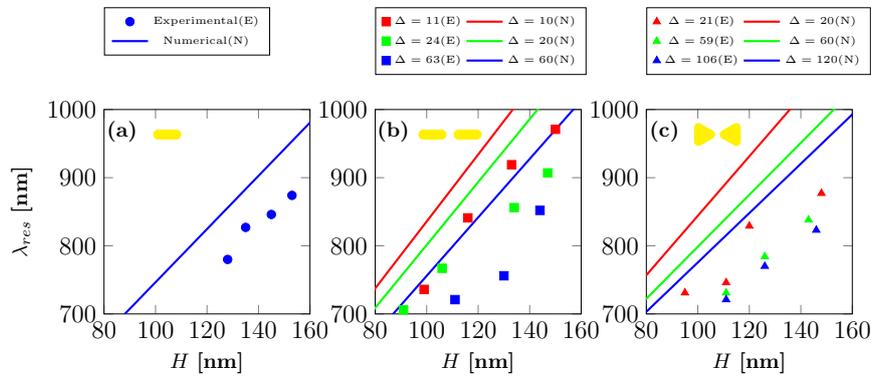


Figure 6.6: Comparison between numerical and experimental results. Resonance wavelength λ_{res} as a function of the antenna length H for the three antenna geometries. The solid lines represent simulation data while the markers represent experimental data. (a) Experimental SR data (circles) compared with simulated data (solid line). (b) Experimental DR data (squares) for different gaps Δ (11 nm (red), 24 nm (green), 63 nm (blue)) compared with the simulation data ($\Delta = 10$ nm (red), $\Delta = 20$ nm (green), $\Delta = 60$ nm (blue)). (c) Experimental BT data (triangles) for different gaps Δ (21 nm (red), 59 nm (green), 106 nm (blue)) compared with the simulation data ($\Delta = 20$ nm (red), $\Delta = 60$ nm (green), $\Delta = 120$ nm (blue)). [25].

reduce the Q -factor by 50% (e.g. $Q = 4.5$ versus $Q = 9.3$ for a SR antenna with and without Ti respectively). Local inhomogeneities (such as redeposited Au or small differences among the $N = 20$ antennas) will also contribute to a broadening of the experimental extinction. [24] Reducing the Ti thickness or using a molecular linker adhesion layer can increase the Q -factor again. [32] Another cause of low Q -factors in nanopatterned antennas is polycrystallinity. [33] Recently it has been shown that thermal annealing can increase the Q -factor of lithographically defined structures by reducing the grain boundaries. In that case, the antennas have to be encapsulated however to avoid shape changes. [31] The Xe ion milling in our FOx process scheme could, as opposed to procedures based on lift-off (such as PMMA), immediately be applied to annealed metal layers in which the nanoantennas can be milled instantly. On the downside, FOx is not easily removed and could hinder the formation of optimal structures for SERS applications. As described in the Fabrication chapter, ma-N 2400.6 can be used as an alternative to FOx and is easier to remove. Since we have shown that the Xe milling has no negative impact on the waveguide loss and allows the fabrication of integrated narrow gap antennas, the suggested scheme could hence be considered as a proper alternative for lift-off to produce potentially higher Q lithographic antennas.

6.4.3 Interaction efficiency

As we have established the possibility to produce gapped nanostructures on top of a single mode waveguide, it is important to compare the efficiency by which they interact with the evanescent field of the waveguide. This can be quantified by the earlier defined interaction efficiency $\eta(\lambda)$. In Figure 6.7 we plot the theoretical (a-b) and experimental (c) interaction efficiencies (at the resonance wavelength) for the different antenna types. Each marker represents a certain DR (Figure 6.7(a)) or BT (Figure 6.7(b)) geometry (so a certain (H, Δ) combination; the width W of the DR antenna is again 55 nm and the bowtie apex angle $\alpha = 60^\circ$). The gray line represents the SR interaction efficiency.

The theoretical curves (Figures 6.7(a-b)) predict that, for the same resonance wavelength λ_{res} , both DR and BT antennas generally have higher interaction efficiencies than SR antennas. The interaction efficiencies of DR antennas exceed the SR efficiencies $\eta_{SR}(\lambda_{res})$ by 30 – 50% depending on the specific λ_{res} in the investigated wavelength region. For BT antennas, a typical increase between 20 – 40% is observed. For a fixed λ_{res} , DR and BT antennas have an increased interaction efficiency when

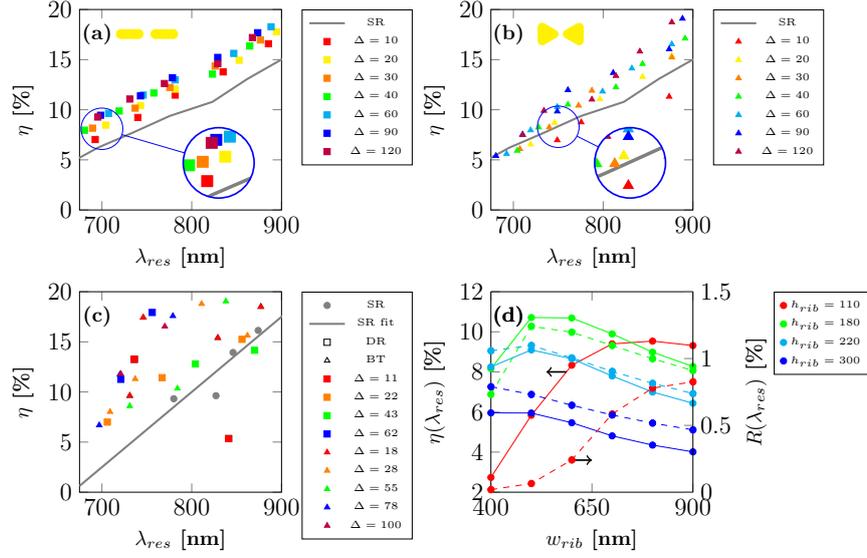


Figure 6.7: **(a-b)** Theoretical interaction efficiencies $\eta(\lambda_{res})$ of DR **(a)** and BT **(b)** antennas as a function of the resonance wavelength λ_{res} , compared with the SR interaction efficiency (gray line). Each color represents a certain gap: 10 nm (red), 20 nm (yellow), 30 nm (orange), 40 nm (green), 60 nm (cyan), 90 nm (blue), 120 nm (purple). For a given gap, an increasing λ_{res} corresponds with increasing H . **(c)** Experimental interaction efficiencies of DR (square markers) and BT (triangle markers) antennas compared with SR antennas (gray circles). The red, orange, green and blue square markers represent structures with gaps of 11 nm, 22 nm, 43 nm and 62 nm respectively while the red, orange, green, blue and purple triangle markers represent structures with gaps of 18 nm, 28 nm, 55 nm, 78 nm and 100 nm respectively. The gray line represents an interpolation of the SR interaction efficiencies. **(d)** Theoretical interaction efficiency η (solid curves, left vertical axis) of a SR antenna ($H = 100$ nm and $W = 55$ nm) and power fraction R (dashed curves, right vertical axis) as a function of the waveguide width w_{rib} and height h_{rib} (both evaluated at λ_{res}). [25].

the gap Δ is larger (compared to the SR η). This is highlighted in the magnified parts of Figs. 6.7(a) and 6.7(b) respectively. The red markers correspond to a gap of $\Delta = 10$ nm while the purple markers correspond to $\Delta = 120$ nm. Differences up to 32% and 66% for DR and BT antennas respectively are observed. In Figure 6.7(c) the experimental interaction efficiencies are plotted. The theoretically predicted trends are confirmed by our experiments. The interaction efficiencies of gapped nanoantennas generally exceed the SR efficiency while the larger gap structures generally have a higher η as compared to narrow gap structures (for the same λ_{res}). We report experimental interaction efficiencies as high as $\approx 19\%$ around 800 nm. Simulations show that the interaction efficiency decreases when the misalignment Δ_y of the antenna with respect to the center of the waveguide increases. For the centro-symmetric antennas we considered, it is therefore always desirable that the misalignment between the two e-beamsteps is kept to a minimum in order to maximize the interaction efficiency with the evanescent field.

Apart from changing the antenna parameters to improve the interaction with the evanescent field, the waveguide geometry can also be adapted to optimize the interaction. For a fixed SR geometry ($H = 100$ nm, $W = 55$ nm and thickness $2 + 35 + 2 = 39$ nm), we varied the waveguide width w_{rib} for four waveguide thicknesses ($h_{rib} = 110, 180, 220, 300$ nm) and calculated $\eta(\lambda_{res})$ for each (w_{rib}, h_{rib}) combination (see Figure 6.7(d)). For each h_{rib} , an optimal width w_{rib} can be found for which η is maximized. When w_{rib} is decreased below a certain value (for a given h_{rib}), the waveguide mode goes into cut-off and hence causes a drop in η . While the specific value of h_{rib} and w_{rib} has a negligible impact on the resonance wavelength λ_{res} , it does allow to change the spatial overlap between the excitation profile and the antenna cross-section by modifying the confinement of the waveguide mode. For an improved physical understanding we considered a non-functionalized waveguide with varying width w_{rib} and height h_{rib} . For each configuration we calculated the ratio $R(\lambda_{res})$ (again at λ_{res}) of the power flow through a cross-sectional area A_{ant} and the total modal power; A_{ant} represents the cross-sectional area, perpendicular to the propagation direction, that would be covered by the antenna in case the waveguide would be functionalized with it (for this particular SR case $A_{ant} = 100 \times 39$ nm²). For each configuration this ratio hence represents the fraction of the modal power that flows through a fixed cross-sectional area that the antenna would cover in a plane perpendicular to the propagation direction. We found that $R(\lambda_{res})$

is maximized for the same (w_{rib}, h_{rib}) combination as the (w_{rib}, h_{rib}) combination that maximizes the interaction efficiency (see Figure 6.7(d)). Physically this means that the optimal interaction efficiency is obtained when the fraction of the modal power that flows through the antenna cross-section is maximized.

To get a better physical understanding of the (improved) interaction between the antenna and the evanescent field, it would be instructive to compare the efficiency of free-space and waveguide excitation of nanoplasmonic antennas (i.e. the efficiency by which an external excitation source transfers energy to the plasmon mode). To this end, we will consider the coupled mode theory picture as described in Chapter 2. There it was shown that the temporal evolution of the dipolar plasmon mode amplitude A_1 can be described by

$$\frac{dA_1}{dt} = j(\omega - \omega_1)A_1 - \frac{\gamma_r + \gamma}{2}A_1 + \kappa_{in}s_+, \quad (6.1)$$

where ω_1 is the dipolar plasmon resonance frequency, γ_r and γ the radiative and non-radiative decay rate of the plasmon mode respectively, s_+ an external excitation and κ_{in} the coupling constant between the external excitation and the plasmon mode:

$$\kappa_{in} = \sqrt{\frac{A_c}{A_i}}\gamma_r, \quad (6.2)$$

where A_c is the effective aperture of the antenna and A_i the spot size of the excitation beam. [34] In steady-state and at resonance, the modal amplitude is given by

$$|A_1|^2 = \frac{4|\kappa_{in}|^2|s_+|^2}{(\gamma_r + \gamma)^2} = \frac{4\left(\frac{A_c}{A_i}\gamma_r\right)|s_+|^2}{(\gamma_r + \gamma)^2}. \quad (6.3)$$

In case of free-space excitation of a nanoplasmonic antenna in a uniform medium the radiative decay rate will be denoted by γ_r^{fs} , while for waveguide excitation the radiative decay rate is γ_r^{wg} . Due to the vicinity of the waveguide structure, the radiative decay rate will most likely differ from γ_{fs} . The presence of a layered medium can indeed modify the spontaneous decay rate of a nearby emitter (in this case the nanoplasmonic antenna). Moreover the spot size of the excitation beam will differ as well. For free-space excitation (e.g. illumination with a Gaussian beam) this will be denoted by A_i^{fs} , while for waveguide excitation (the evanescent field) the spot size is denoted by A_i^{wg} . In a similar fashion the effective antenna

aperture can be different as well (e.g. different overlap) and will be denoted A_c^{fs} and A_c^{wg} respectively. The intrinsic damping γ and the strength of the external excitation $|s_+|^2$ is assumed to be the same for both cases. The ratio of energy stored in the plasmon mode with waveguide and free-space excitation respectively is then given by

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} = \frac{A_c^{wg}}{A_c^{fs}} \frac{A_i^{fs}}{A_i^{wg}} \frac{\gamma_r^{wg}}{\gamma_r^{fs}} \left(\frac{\gamma_r^{fs} + \gamma}{\gamma_r^{wg} + \gamma} \right)^2. \quad (6.4)$$

Define $\psi = \frac{A_c^{wg}}{A_c^{fs}} \frac{A_i^{fs}}{A_i^{wg}}$, $\chi = \frac{\gamma_r^{wg}}{\gamma_r^{fs}}$ and $x = \frac{\gamma}{\gamma_r^{fs}}$ such that

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} = \frac{\psi}{\chi} \left(\frac{1+x}{1+\frac{x}{\chi}} \right)^2. \quad (6.5)$$

We will now make some assumptions in order to assess the relative strength of waveguide and free-space excitation. At first we approximate $A_c^{fs} \approx A_c^{wg}$. A_c is dependent on the overlap between the excitation field and the antenna, so it is reasonable to assume that A_c^{wg} will never be smaller than A_c^{fs} , as one would expect that the overlap between a sub-wavelength excitation field and a nanoplasmonic antenna will be better as compared to the overlap with a diffraction limited free-space field. The spot size of the free-space excitation beam is estimated according to $A_i^{fs} \approx (\lambda/NA)^2$, where λ is the excitation wavelength and NA the numerical aperture of the objective. For $\lambda = 785$ nm and $NA = 0.7$ we get $A_i^{fs} \approx (1120)^2$ nm². A single mode waveguide operating at 785 nm has a typical evanescent field size of $A_i^{wg} \approx 700 \times 200$ nm². Eventually $A_i^{fs}/A_i^{wg} \approx 9$, so $\psi \approx 9$. Moreover we can assume that γ_r and γ are of similar order such that $x \approx 1$ and

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} \approx \frac{4\chi\psi}{(1+\chi)^2}. \quad (6.6)$$

Now we can consider three regimes

1. $\chi \ll 1$, i.e. radiative quenching due to the waveguide:

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} \approx 4\chi\psi \quad (6.7)$$

If the quenching is really strong (i.e. $\chi \ll 1(4\psi)$), then waveguide excitation will be less efficient. For moderate quenching (i.e. $1/(4\psi) < \chi \ll 1$), waveguide excitation will be more efficient.

2. $\chi \approx 1$, i.e. no radiative quenching or enhancement:

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} \approx \psi \quad (6.8)$$

In this case the waveguide excitation will always be more efficient.

3. $\chi \gg 1$, i.e. radiative enhancement:

$$\frac{|A_1|_{wg}^2}{|A_1|_{fs}^2} \approx \frac{4\psi}{\chi} \quad (6.9)$$

If the radiative decay is increased too much (i.e. $\chi \gg 4\psi$), then waveguide excitation is less efficient. For moderate enhancement $1 \ll \chi < 4\psi$, waveguide excitation will be more efficient.

In most practical cases, the radiative decay rate will be changed by the addition of a layered waveguide medium, but not in such extremes that the quenching or enhancement is huge. Therefore, one could argue that in most cases the waveguide excitation will be ψ times more efficient as compared to free-space excitation. This coupled-mode theory picture therefore qualitatively explains the more efficient excitation by evanescent fields.

6.4.4 Raman enhancement

In Figure 6.8 we plot the Raman enhancement in a plane parallel to the waveguide surface and halfway the gold antenna. For this particular case, a bowtie antenna with fixed $\alpha = 60^\circ$ and $\Delta = 10$ nm and varying H was considered. The Raman enhancement factor $EF_R = EF(\lambda_P)^2 EF(\lambda_S)^2$ is evaluated for $\lambda_P = 785$ nm and $\lambda_S = 877$ nm; $EF(\lambda)$ is the field enhancement factor relative to a non-functionalized reference waveguide (i.e. without antenna) at wavelength λ . By varying the length, the antenna extinction is tuned across the 700 – 1000 nm (see Figures 6.8(a-d)). The pump and Stokes wavelength are highlighted by a red and cyan line respectively. One can see that an increased overlap of the plasmon resonance with the pump and Stokes wavelength increases the maximum Raman enhancement EF_R^{max} . Nevertheless, it should be noted that the Raman enhancement is only one part of the story since we are working with integrated nanoplasmonic antennas. Stokes light that is generated by an antenna is now propagating in a functionalized waveguide, as a result

of which it will also attenuate due to the antenna extinction. In the next chapter we will consider this interplay in close detail.

Moreover, the bowtie antennas exhibit the highest Raman enhancement factor (among the three studied geometries) as can be seen in Figure 6.9. There we compare single rod (a), double rod (b) and bowtie (c) antennas. In all three cases the length is varied to tune the resonance, while the other antenna parameters are kept fixed (see figure caption). An optimum EF_R^{max} is found for all geometries, but the bowtie optimum exhibits the highest EF_R^{max} . However, the difference between double rod and bowtie antennas is not that large. The single rod on the other hand has an EF_R^{max} which is almost two orders of magnitude lower, so the use of integrated gapped nanostructures clearly has an advantage over single rod structures. In the next chapter we will investigate the on-chip SERS capabilities of these same structures as well.

6.5 Excitation of dark plasmon modes

Apart from the dipolar resonances considered so far, on-chip bowtie antennas also exhibit higher order plasmon modes. These can be excited when the apex angle is changed (Figures 6.10(a) and 6.10(b)). In Figure 6.10(a) the theoretical extinction spectra of a bowtie antenna with fixed length $H = 160$ nm and gap $\Delta = 20$ nm but varying apex angle α are depicted. When α increases, the dipolar resonance redshifts, while a second blueshifted resonance appears for sufficiently large α . In order to reveal the nature of this resonance, the charge densities at the top surface of the $\alpha = 90^\circ$ BT antenna were simulated with Lumerical FDTD Solutions. The charge density is evaluated at the resonance of each peak, corresponding to $\lambda \approx 700$ nm (red dot in Figure 6.10(a)) and $\lambda \approx 1125$ nm (green dot in Figure 6.10(a)). Figures 6.10(c) and 6.10(d) represent the normalized charge density (the density is normalized with its maximum value) of the $\lambda = 700$ nm and $\lambda = 1125$ nm peak respectively. The $\lambda = 700$ nm resonance exhibits a quadrupolar nature while the $\lambda = 1125$ nm resonance corresponds to a dipolar charge distribution. The effective wavelength of the propagating waveguide mode is $\lambda/n_{eff} \approx 433$ nm (for $\lambda = 700$ nm) while the base of the BT antenna is $2H \tan(\alpha/2) = 320$ nm. In this

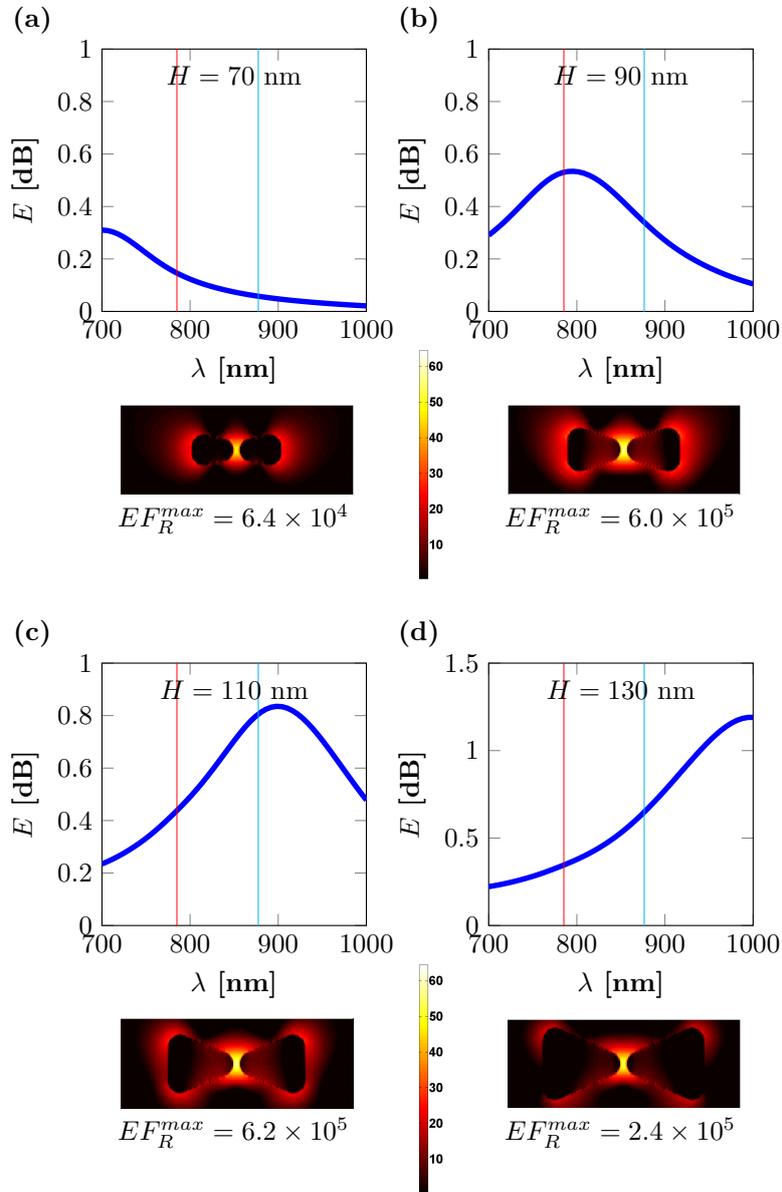


Figure 6.8: Extinction spectra and Raman enhancement factor (for $\lambda_P = 785$ nm and $\lambda_S = 877$ nm, highlighted by the red and cyan lines respectively) of different bowtie antennas with fixed $\alpha = 60^\circ$ and $\Delta = 10$ nm. The colorbar is on a logarithmic scale. The maximum Raman enhancement factor EF_R^{max} is also shown. (a) $H = 70$ nm. (b) $H = 90$ nm. (c) $H = 110$ nm. (d) $H = 130$ nm.

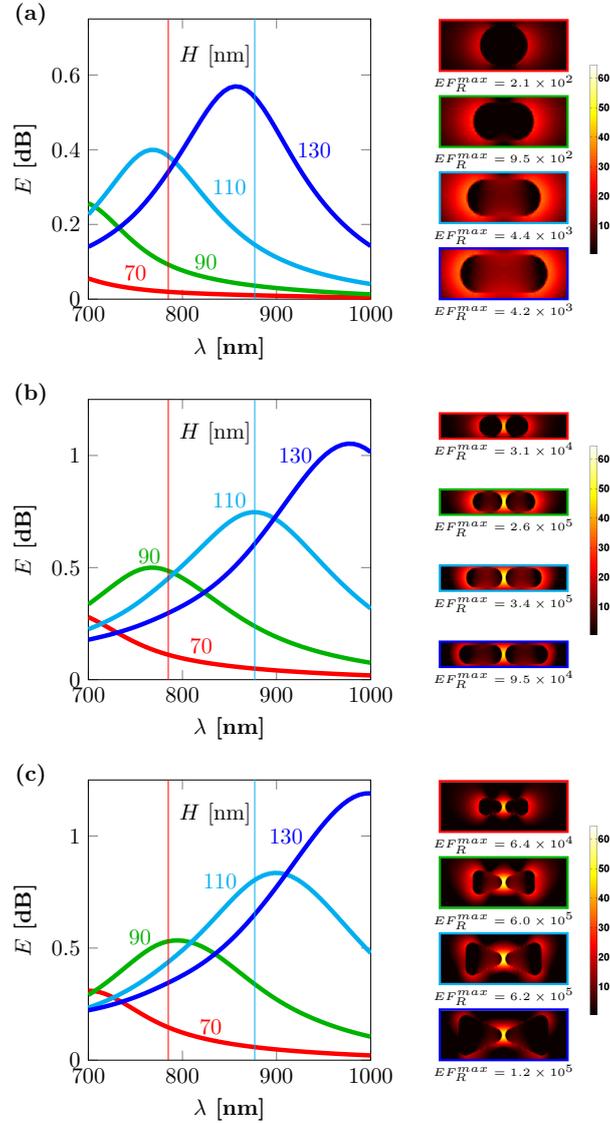


Figure 6.9: Comparison of the extinction spectra and Raman enhancement factors of single rod (a), double rod (b) and bowtie (c) antennas. In all cases the length is varied while the other parameters are kept constant. For the single rod we considered a fixed $W = 55$ nm, for the double rod a fixed $W = 55$ nm and $\Delta = 10$ nm and for the bowtie antenna a fixed $\alpha = 60^\circ$ and $\Delta = 10$ nm. The colorbar is on a logarithmic scale.

case, the quasi-static approximation no longer holds and phase delay effects become important. Since the antenna is not illuminated from the top but from the side, the phase of the excitation field changes considerably while the wave propagates along the antenna. The phase of the y -component of the excitation field (along the TE direction) is shown in Figure 6.10(e) and plotted as a function of the propagation distance. The blue dashed lines correspond to the outer tips of the bowtie antenna. The phase difference along the bowtie antenna is $< \pi$ for the dipole mode and $> \pi$ for the quadrupole mode. For the quadrupole mode, the excitation field hence changes sign when propagating along the antenna. So the waveguide mode can excite dark plasmon modes, as soon as the apex angle α is large enough, with an interaction efficiency of $\approx 15.62\%$ (for this particular BT antenna). When the length H (for fixed α) or the apex angle α (for fixed H) is increased, the dark resonance redshifts with concomitant increase in the interaction efficiency. It is interesting to note that waveguide excitation allows an equally efficient interaction of the evanescent field with the bright and the dark plasmon mode because the interaction efficiencies of both modes are of the same order (for a given λ_{res}).

The excitation of a multipolar plasmon resonance is also experimentally verified and shown in Figure 6.10(b), where extinction spectra of BT antennas with different apex angles α are shown. The broad dipolar peak redshifts and a second resonance appears when α is increased. Please note that the simulations were performed for a fixed gap and length, while the experimental gaps and lengths differ for different α . As a result, there will be a difference between the absolute positions of the theoretical and experimental resonance. Nevertheless, the qualitative trends are predicted well by the simulation and for the $\alpha = 90^\circ$ bowtie, for which the experimental dimensions correspond well with the simulated ones, the predicted resonance wavelengths are also in reasonable correspondence with the experimental ones. While the dipolar peak is again superimposed with a fringe pattern, the quadrupolar peak is much more smooth. Apart from a (partial) quenching of the radiative plasmon decay into the waveguide, the lack of fringes could also be explained by an anisotropic radiation pattern of the quadrupolar mode (e.g. predominant forward scattering and no backscattering along the waveguide mode propagation direction) or radiation along directions which don't couple back into the waveguide. Other antenna geometries, such as the nanocross [35], could be excited

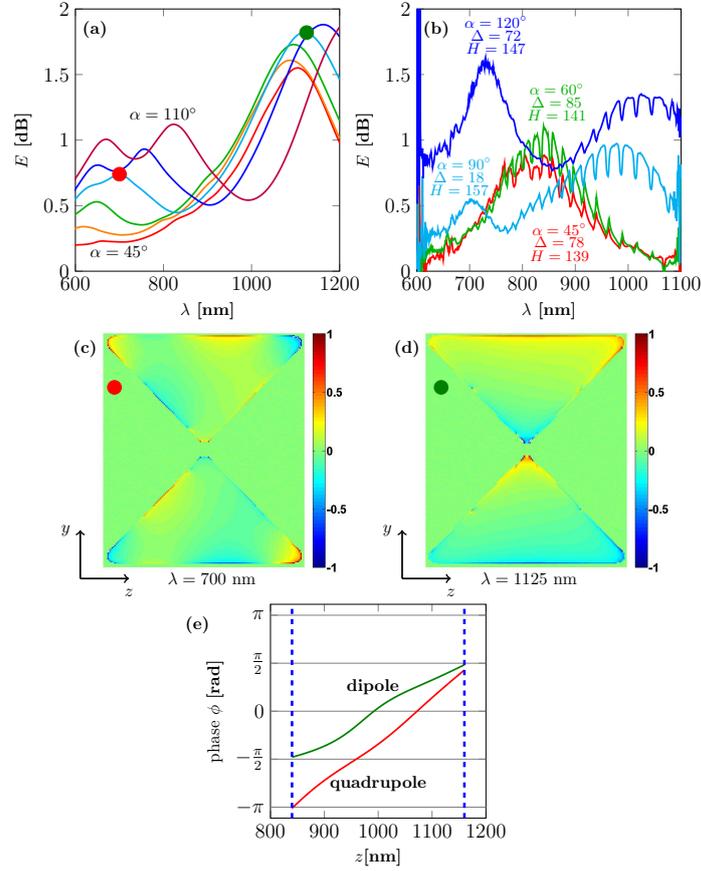


Figure 6.10: (a) Theoretical extinction spectra of BT antennas with varying apex angle (45° (red), 60° (orange), 75° (green), 90° (cyan), 100° (blue), 110° (purple)) and fixed length $H = 160$ nm and gap $\Delta = 20$ nm. (b) Experimental extinction spectra of BT antennas with increasing apex angle: $\alpha = 45^\circ$ (red), $\alpha = 60^\circ$ (green), $\alpha = 90^\circ$ (cyan) and $\alpha = 120^\circ$ (blue). Gaps Δ and lengths H are in [nm]. (c-d) Simulated charge density plots at the top surface of a BT antenna with length $H = 160$ nm, gap $\Delta = 20$ nm and apex angle $\alpha = 90^\circ$. (c) Normalized charge density of the quadrupolar plasmon mode at $\lambda = 700$ nm (corresponding with the red dot in (a)). (d) Normalized charge density of the dipolar plasmon mode at $\lambda = 1125$ nm (corresponding with the green dot in (a)). (e) Phase of the y -component of the electric field as a function of the propagation distance (evaluated in the center of the waveguide cross-section). The blue dashed lines represent the outer tips of the bowtie antenna ($\alpha = 90^\circ$, $H = 160$ nm, $\Delta = 20$ nm). The red line corresponds to the quadrupole mode (red dot in (a)) and the green line corresponds to the dipolar mode (green dot in (a)). [25].

through an evanescent waveguide field in order to tailor specific and efficient on-chip multipolar plasmonic resonances which can be used for e.g. improved LSPR sensing. Furthermore, the multipole character of these gapped antennas could be used for on-chip enhanced Raman spectroscopy by tuning the resonance peaks such that they optimally match with both the pump and Stokes (or anti-Stokes) wavelength.

6.6 Conclusion

In this chapter we discussed the resonance tuning of integrated nanoplasmonic antennas. The influence of several geometrical parameters was investigated for three different antenna structures (single rod, double rod, bowtie antenna). For each of these three antenna types, we moreover performed a numerical analysis to assess their maximum Raman enhancement factor. Apart from the broad dipolar resonance, we also discussed the excitation of quadrupolar plasmon modes in large apex bowtie antennas. This hybrid nanophotonic-plasmonic platform, with several resonance tuning possibilities, will be applied in the next chapter to perform on-chip SERS.

References

- [1] Anker, J.N.; Hall, W.P.; Lyandres, O.; Shah, N.C.; Zhao, J.; Van Duyne, R.P. Biosensing with plasmonic nanosensors. *Nat. Mater.* **2008**, *7*, 442–453.
- [2] Willets, K. A. and Van Duyne, R.P. Localized surface plasmon resonance spectroscopy and sensing. *Annu. Rev. Phys. Chem.* **2007**, *58*, 267–297.
- [3] Halas, N. J.; Lal, S.; Chang, W-S.; Link, S.; Nordlander, P. Plasmons in strongly coupled metallic nanostructures. *Chem. Rev.* **2011**, *111*, 3913–3961.
- [4] Mayer, K. M. and Hafner, J.H. Localized surface plasmon resonance sensors. *Chem. Rev.* **2011**, *111*, 3828–3857.
- [5] Kelly, K. L.; Coronado, E.; Zhao, L. L.; Schatz, G. C. The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment. *J. Phys. Chem. B* **2003**, *111*(3), 668–677.

- [6] Sundaramurthy, A.; Crozier, K. B.; Kino, G. S.; Fromm, D. P.; Schuck, P. J.; Moerner, W. E. Field enhancement and gap-dependent resonance in a system of two opposing tip-to-tip Au nanotriangles. *Phys. Rev. B* **2005**, *72*, 165409-1–165409-6.
- [7] Fischer, H. and Martin, O. J. F. Engineering the optical response of plasmonic nanoantennas. *Opt. Express* **2008**, *16*(12), 9144–9154.
- [8] Giannini, V.; Fernández-Domínguez, A. I.; Heck, S. C.; Maier, S. A. Plasmonic nanoantennas: fundamentals and their use in controlling the radiative properties of nanoemitters. *Chem. Rev.* **2011**, *111*, 3888–3912.
- [9] Mock, J.J.; Smith, D.R.; Schultz, S. Local refractive index dependence of plasmon resonance spectra from individual nanoparticles. *Nano Lett.* **2003**, *3*(4), 485–491.
- [10] Jensen, T.R.; Duval Malinsky, M.; Haynes, C.L.; Van Duyne, R.P. Nanosphere lithography: tunable localized surface plasmon resonance spectra of silver nanoparticles. *J. Phys. Chem. B* **2000**, *104*, 10549–10556.
- [11] Su, K-H.; Wei, Q-H.; Zhang, X.; Mock, J.J.; Smith, D.R.; Schultz, S. Interparticle coupling effects on plasmon resonances of nanogold particles. *Nano Lett.* **2003**, *3*(8), 1087–1090.
- [12] Muskens, O.L.; Giannini, V.; Sánchez-Gil, J.A.; Gómez Rivas, J. Optical scattering resonances of single and coupled dimer plasmonic nanoantennas. *Opt. Express* **2007**, *15*(26), 17736–17746.
- [13] Haes, A.J.; Zou, S.; Schatz, G.C.; Van Duyne, R.P. Nanoscale optical biosensor: short range distance dependence of the localized surface plasmon resonance of noble metal nanoparticles. *J. Phys. Chem. B* **2004**, *108*, 6961–6968.
- [14] Haes, A.J.; Hall, W.P.; Chang, L.; Klein, W.L.; Van Duyne, R.P. A localized surface plasmon resonance biosensor: first steps toward an assay for Alzheimers disease. *Nano Lett.* **2004**, *4*(6), 1029–1034.
- [15] Verellen, N.; Van Dorpe, P.; Huang, C.; Lodewijks, K.; Vandenbosch, G.A.E.; Lagae, L.; Moshchalkov, V.V. Plasmon line shaping using nanocrosses for high sensitivity localized surface plasmon resonance sensing. *Nano Lett.* **2011**, *11*, 391–397.

- [16] Kinkhabwala, A. ; Yu, Z.; Fan, S.; Avlasevich, Y.; Müllen, K.; Moerner, W. E. Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna. *Nature Photon.* **2009**, *3*, 654–657.
- [17] Chu, Y.; Banaee, M.G.; Crozier, K.B. Double-resonance plasmon substrates for surface-enhanced Raman scattering with enhancement at excitation and Stokes frequencies. *ACS Nano* **2010**, *4*(5), 2804–2810.
- [18] McFarland, A.D.; Young, M.A.; Dieringer, J.A.; Van Duyne, R.P. Wavelength-scanned surface-enhanced Raman excitation spectroscopy. *J. Phys. Chem. B* **2005**, *109*, 11279–11285.
- [19] Arnaud, L.; Bruyant, A.; Renault, M.; Hadjar, Y.; Salas-Montiel, R.; Apuzzo, A.; Léron del, G.; Morand, A.; Benech, P.; Le Coarer, E.; Blaize, S. Waveguide-coupled nanowire as an optical antenna. *J. Opt. Soc. Am. A* **2013**, *30*, 2347–2355.
- [20] Arango, F. B.; Kwadrin, A.; Koenderink, A.F. Plasmonic Antennas Hybridized with Dielectric Waveguides. *ACS Nano* **2012**, *6*, 10156–10167.
- [21] Février, M.; Gogol, P.; Aassime, A.; Mégy, R.; Delacour, C.; Chelnokov, A.; Apuzzo, A.; Blaize, S.; Lourtioz, J-M.; Dagens, B. Giant Coupling Effect between Metal Nanoparticle Chain and Optical Waveguide. *Nano Lett.* **2012**, *12*, 1032–1037.
- [22] Février, M.; Gogol, P.; Lourtioz, J-M.; Dagens, B. Metallic nanoparticle chains on dielectric waveguides: coupled and uncoupled situations compared. *Opt. Express* **2013**, *21*, 24504–24513.
- [23] Février, M.; Gogol, P.; Barbillon, G.; Aassime, A.; Mégy, R.; Bartenlian, B.; Lourtioz, J-M.; Dagens, B. Integration of short gold nanoparticles chain on SOI waveguide toward compact integrated biosensors. *Opt. Express* **2012**, *20*, 17402–17410.
- [24] Chamanzar, M.; Xia, Z.; Yegnanarayanan, S.; Adibi, A. Hybrid integrated plasmonic-photonic waveguides for on-chip localized surface plasmon resonance (LSPR) sensing and spectroscopy. *Opt. Express* **2013**, *21*, 32086–32098.
- [25] Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Bright and dark plasmon resonances of

- nanoplasmonic antennas evanescently coupled with a silicon nitride waveguide. *Opt. Express* **2015**, *23*(3), 3088–3101.
- [26] Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Jansen, R.; Claes, T.; Rottenberg, X.; Peyskens, F.; Selvaraja, S.; Helin, P.; Du Bois, B.; Leyssens, K.; Severi, S.; Deshpande, P.; Baets, R.; Van Dorpe, P. Low-Loss Singlemode PECVD Silicon Nitride Photonic Wire Waveguides for 532900 nm Wavelength Window Fabricated Within a CMOS Pilot Line. *IEEE Photon. J.* **2013**, *5*(6), 2202809.
- [27] Romero-García, S.; Merget, F.; Zhong, F.; Finkelstein, H.; Witzens, J. Silicon nitride CMOS-compatible platform for integrated photonics applications at visible wavelengths. *Opt. Express* **2013**, *21*(12), 14036–14046.
- [28] Dhakal, A.; Subramanian, A.Z.; Wuytens, P.; Peyskens, F.; Le Thomas, N.; Baets, R. Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides. *Opt. Lett.* **2014**, *39*, 4025–4028.
- [29] Johnson, P.B. and Christy, R.W. Optical constants of the noble metals. *Phys. Rev. B* **1972**, *6*(12), 4370–4379.
- [30] Haynes, W. M., ed., *CRC Handbook of Chemistry and Physics* 95th Edition (Internet Version 2015); CRC/Taylor and Francis, 2015.
- [31] Bosman, M.; Zhang, L.; Duan, H.; Fen Tan, S.; Nijhuis, C.A.; Qiu, C-W.; Yang, J.K.W. Encapsulated annealing: enhancing the plasmon quality factor in lithographically-defined nanostructures. *Sci. Rep.* **2014**, *4*, 5537.
- [32] Siegfried, T.; Ekinici, Y.; Martin, O.J.F.; Sigg, H. Engineering metal adhesion layers that do not deteriorate plasmon resonances. *ACS Nano* **2013**, *7*(3), 2751–2757.
- [33] Novotny, L. and van Hulst, N. Antennas for light. *Nature Photon.* **2011**, *5*, 83–90.
- [34] Joon Seok, T.; Jamshidi, A.; Kim, M.; Dhuey, S.; Lakhani, A.; Choo, H.; James Schuck, P.; Cabrini, S.; Schwartzberg, A.M.; Bokor, J.; Yablonovitch, E.; Wu, M.C. Radiation Engineering of Optical Antennas for Maximum Field Enhancement. *Nano Lett.* **2011**, *11*, 2606–2610.

-
- [35] Verellen, N.; Van Dorpe, P.; Vercruyse, D.; Vandenbosch, G.A.E.; Moshchalkov, V.V. Dark and bright localized surface plasmons in nanocrosses. *Opt. Express* **2011**, *19*(12), 11034–11051.

7

Integrated Surface Enhanced Raman Spectroscopy

7.1 Introduction

As outlined in an earlier chapter, Surface Enhanced Raman Spectroscopy (SERS) is a promising technique for enhancing inherently weak Raman signals by introducing a plasmonic nanostructure in the vicinity of the analyte under study. [1–12] SERS spectra have however mainly been generated using non-integrated plasmonic substrates (e.g. on a glass slide) or colloidal solutions of metallic nanoparticles, whereby the signal is excited and collected by a conventional microscopy system. Our group established that photonic integrated circuits (PICs) offer a promising alternative to standard confocal microscopy to probe spontaneous Raman spectra. [13] The fundamental building block of such circuits is a single mode waveguide in which the pump and Stokes light is guided in a high-index core material surrounded by lower-index cladding materials. While researchers also started integrating nanoplasmonic antennas on top of such waveguides [14–21], PICs have only been used to probe SERS signals from external, non-integrated, metallic nanoparticles [22–24]. Such an approach is however poor in terms of quantitative results owing to the large uncertainty on the Raman enhancement and coupling between the

excitation beam and the metallic nanoparticles. [25] In order to obtain quantitative SERS spectra, a complete control of the plasmonic enhancement and coupling with the underlying waveguide is necessary. This challenging problem forms the central goal of this PhD work. In the previous chapters we already described the fabrication challenges and thoroughly discussed the resonance tuning of integrated nanoplasmonic antennas in the near-IR region (covering 700 – 1000 nm). As previously mentioned, this region is of particular interest for on-chip Raman spectroscopy since pumping in the near-IR (e.g. at 785 nm) can reduce the unwanted fluorescence background and allows low-loss propagation of both pump and Stokes light in a silicon nitride (SiN) waveguide. [26, 27]

In this chapter we present a new class of SERS probes which overcomes the above issues. We demonstrate for the first time the generation of SERS signals from integrated bowtie nanoantennas, excited and collected by a single mode SiN waveguide. Due to the fully integrated nature of this single mode SERS probe it is possible to rigorously quantify the complete enhancement and coupling process. We show that the SERS spectra can be attributed to a genuine plasmon resonance effect and experimentally determine the enhanced Stokes power, generated by a single nanoantenna, that is coupled into the fundamental TE-mode. Furthermore, an analytical model identifying the relevant design parameters and figure of merit for waveguide-based SERS is developed as well. We also discuss the impact of the shot noise resulting from the inherent SiN Raman background and outline the interplay between signal optimization and noise reduction. Moreover, a comparison between the theoretical model and the experimental data is included as well. Finally, we give a comparative overview of Spontaneous Raman On-Chip (ROC) versus Surface Enhanced Raman On-Chip (SEROc). Large parts of the work described in this chapter are also published in [28].

7.2 On-chip SERS platform

A schematic of the device under study is shown in Figure 7.1(a). [28] The fundamental TE-mode of a SiN rib waveguide is used to excite a periodic array of N gold bowtie antennas. In order to investigate the SERS properties of this platform, we need a probe molecule which selectively binds to gold and does not stick to SiN in order to avoid contributions from evanescently generated spontaneous Raman scattering. To this end, the samples are immersed overnight in a 1 mM 4-nitrothiophenol:ethanol

solution and subsequently rinsed with pure ethanol to remove the residual 4-nitrothiophenol (NTP). A self-assembled monolayer of NTP is assumed to form on the Au surface through a Au-S bond. [29] As a result, the collected NTP Stokes light should only originate from molecules near the Au surface, hence probing the SERS performance. The pump wavelength for all experiments is set to $\lambda_P = 785$ nm and NTP Stokes light (at λ_S) is subsequently collected back into the same waveguide mode. The samples for our key experiment were fabricated using a two-step e-beam lithography process consisting of a PMMA and ma-N 2403 process (the specific details can be found in the fabrication chapter). As opposed to FOx-12, PMMA can be easily removed, allowing a larger contact surface between the NTP molecules and the Au. Apart from a different spectrometer, we use the same measurement setup as the one used for spontaneous on-chip Raman spectroscopy (Chapter 4, Figure 4.6). Because the SERS signals originate from a very small number of molecules as compared to the huge number of molecules in the bulk IPA experiments, the overall signal will be very weak (despite the enhancement). Therefore a much better detector is necessary. For our on-chip SERS experiments we replaced the Avantes spectrometer by a spectrometer system from ANDOR, consisting of a Shamrock 303i spectrometer and iDus 416 cooled CCD camera. Alignment procedures are identical as the ones described before.

A scanning electron microscope image of the functionalized waveguide, i.e. a waveguide patterned with plasmonic nanoantennas, is depicted in Figure 7.1(b). All waveguides have a cross-sectional area of 220×700 nm². Raman spectra of an uncoated (without NTP) and coated (with NTP) waveguide functionalized with 40 antennas are shown in Figure 7.1(c). The spectral regions where an NTP Stokes peak is expected (1080 , 1110 , 1340 and 1575 cm⁻¹) [29] are highlighted by the cyan shaded areas. Before coating, no NTP peaks can be distinguished from the inherent SiN background. After coating, four additional peaks appear and coincide with the expected NTP Stokes peaks. This demonstrates that SERS signals from single monolayer coated antennas can indeed be excited and collected by the same fundamental waveguide mode. The peaks at 1250 and 1518 cm⁻¹ (marked by the black dashed lines) are attributed to interference effects of the Au array which act on the scattered background light. These peaks are also observed on the extinction curves of the functionalized waveguides. On one hand, the array forms a multiple Fabry-Perot resonator of which the

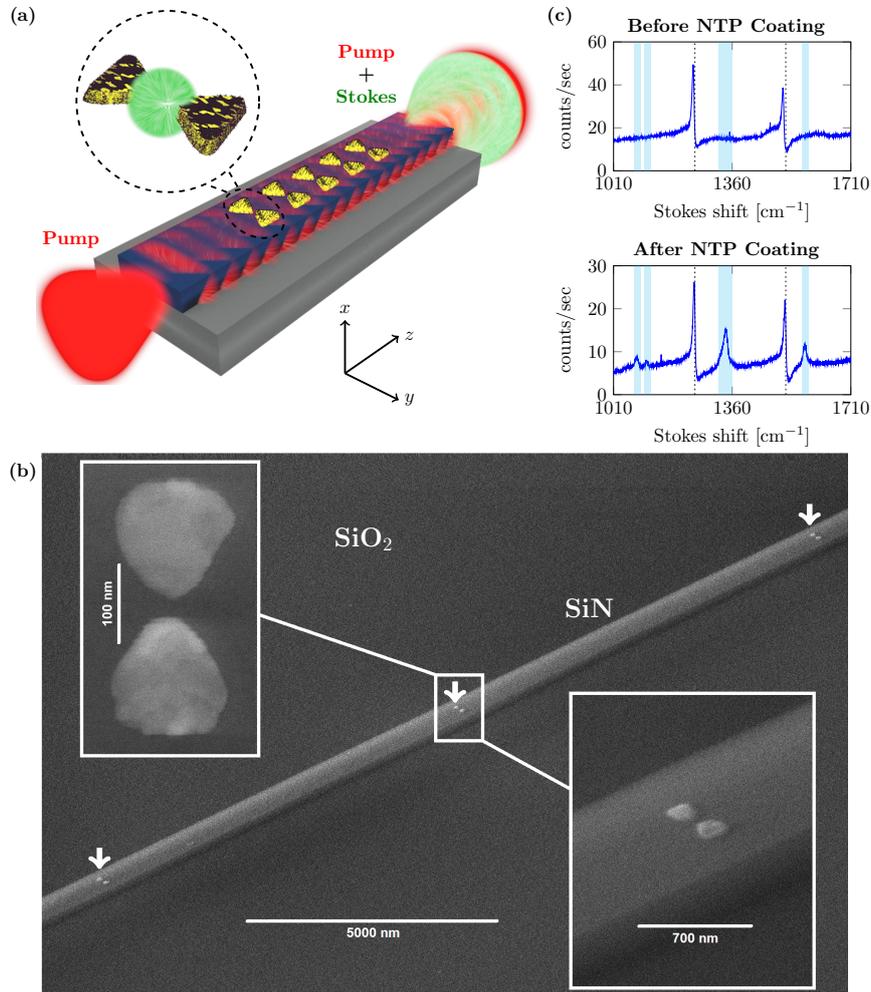


Figure 7.1: (a) Schematic of the chip consisting of single mode SiN waveguides (blue) on an SiO₂ undercladding (gray) and functionalized with an array of gold bowtie antennas (yellow). All antennas are coated with an NTP monolayer (purple dots) and evanescently excited by the fundamental TE-mode (red). The NTP Stokes signal (green) is collected by the same mode. (b) Scanning electron microscope image of a functionalized waveguide. The white arrows indicate antenna positions. The insets show a zoom of a typical antenna. (c) Raman spectra of a waveguide functionalized with 40 antennas. The cyan shaded areas mark the NTP Stokes peaks while the black dashed lines represent peaks attributed to interference effects of the plasmonic array. Before coating, the waveguide already generates a Raman background in itself (top). After coating, NTP peaks emerge (bottom). Taken from [28].

expected free spectral range FSR of 19.9 nm around 1250 cm^{-1} (using the fabricated array period of $\Lambda = 10 \mu\text{m}$) matches well with the experimentally obtained value ($\approx 20.8 \text{ nm}$). On the other hand, the waveguide mode interferes with the radiative decay of the plasmon mode. This will affect the specific lineshape and strength of the fringes. The spectral positions at which we see sudden changes in the Raman background coincide with the fringes observed on the extinction curves. Hence these features are not attributed to specific Raman lines.

In all our experiments, each spectrum was acquired $N_{avg} = 10$ times. This averaging results in a shot noise reduction by $\sqrt{N_{avg}}$ and allows to test whether the spectra remain stable over a certain measurement time. A coupling loss between the external excitation beam and the chip or between the chip and the outcoupling optics would result in a complete loss of the spectrum. The stability of the spectrum with time (see Figure 7.2) proves that the setup is stable for a sufficiently long integration time and hence allows reliable extraction of SERS spectra. After a very long time (> 20 minutes) the stage unavoidably drifts, which results in a loss of the spectrum.

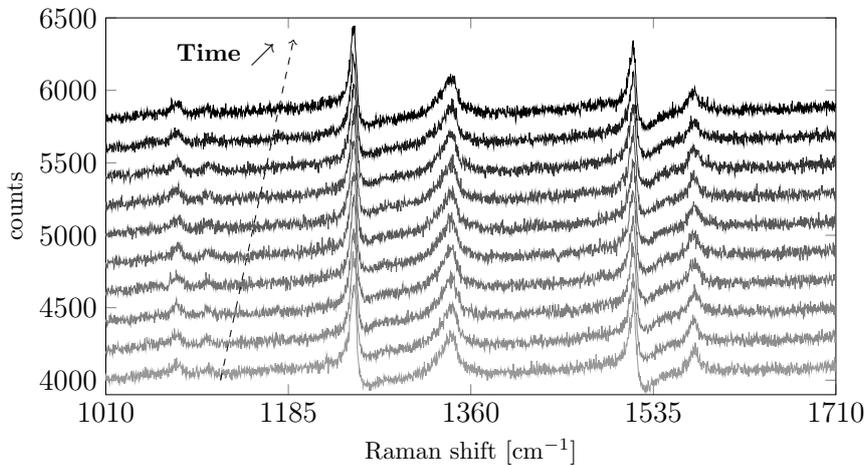


Figure 7.2: Kinetic series of $N_{avg} = 10$ SERS spectra of a waveguide coated with $N = 40$ antennas. The integration time for each spectrum was 30 seconds. Each spectrum is shifted by 200 counts for improved visualization.

7.3 Influence of the plasmon resonance

Subsequently the dependence of the SERS signal on the position of the plasmon resonance was investigated to verify that it can be attributed to a resonance effect and not to coincidental surface roughness. To this end, waveguides functionalized with a fixed number of antennas ($N = 20$) but varying bowtie geometries were considered. The relevant bowtie parameters are its length L , gap Δ and apex angle α (Figure 7.3(a)) and by changing the length one can tune the antenna resonance (extinction). The single antenna extinction spectra $E(\lambda)$ (in dB) are calculated through $E(\lambda) = (T_{ref}(\lambda) - T_{Nant}(\lambda))/N$ in which $T_{ref}(\lambda)$ is the power transmission through a reference waveguide and $T_{Nant}(\lambda)$ the power transmission of a waveguide functionalized with N antennas (see [14] and previous chapters). Figure 7.3(b) depicts such single antenna extinction spectra for three different $\alpha \approx 60^\circ$ bowties ($L_A \approx 95 \pm 8$ nm, $L_B \approx 110 \pm 12$ nm and $L_C \approx 125 \pm 13$ nm with fixed $\Delta \approx 51 \pm 13$ nm), while the corresponding Raman spectra are shown in Figure 7.3(c). The Raman spectrum of a reference waveguide without any Au functionalization is also shown. Even after coating the reference waveguide does not generate NTP peaks, so any Raman signal indeed originates from the antenna region and does not contain contributions from spontaneous Raman scattering along the waveguide. [13] The L_A resonance is detuned from the relevant pump and Stokes region, resulting in a poor Raman spectrum. By increasing the length (L_B and L_C bowties), the resonance redshifts and lines up with the pump and Stokes wavelengths. For these bowties the NTP spectrum starts to emerge. The reported SERS spectra can hence be attributed to a plasmon resonance effect such that a stable and reproducible enhancement factor can be associated with them, in contrast to SERS events originating from random surface defects. The increased overlap with the plasmon resonance, and hence extinction, also results in a decreased background.

Background reduction is an important aspect since the background associated shot noise can potentially mask the weakest SERS (or spontaneous Raman) signals. Due to the lower background, the 1110 cm^{-1} peak is visible for the L_C antenna array in Figure 7.4, but not for the L_B array (Figure 7.4 depicts the L_B and L_C spectra on a linear scale). Despite this advantage of the L_C array, the absolute SERS enhancement (and hence output signal strength) of the Stokes peaks itself is lower for the L_C array as compared to the overall SERS enhancement of the L_B array. This means that if one wants to resolve the total spectrum by reducing the background

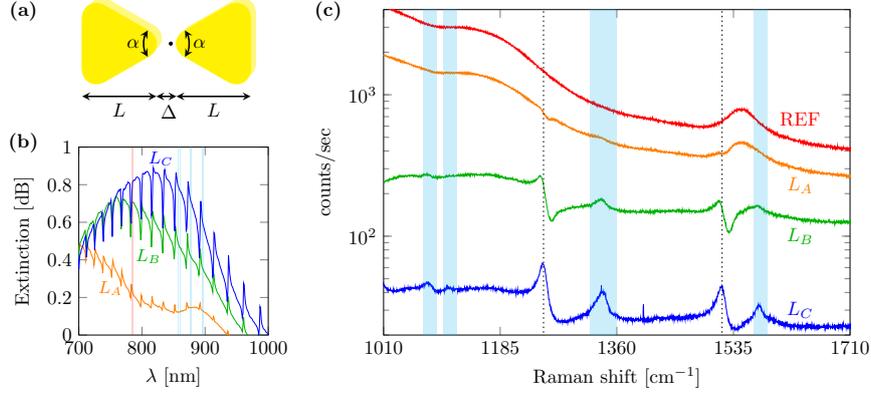


Figure 7.3: **(a)** Geometrical parameters of a bowtie antenna: length L , gap Δ and apex angle α . **(b)** Single antenna extinction spectra for 3 different $\alpha \approx 60^\circ$ bowtie antennas with fixed gap $\Delta \approx 51 \pm 13$ nm but varying length: $L_A \approx 95 \pm 8$ nm (orange), $L_B \approx 110 \pm 12$ nm (green), $L_C \approx 125 \pm 13$ nm (blue). The red and cyan shaded lines correspond to the pump and Stokes wavelengths respectively. **(c)** Corresponding Raman spectra of the waveguides functionalized with these 3 bowtie antennas (orange, green, blue) and Raman spectrum of a reference waveguide (red). Taken from [28].

through antenna attenuation (so e.g. replacing the L_B array by an L_C array), one might have to select a nanoplasmonic antenna array with a non-optimal performance in terms of absolute SERS enhancement. In a later section we will come back to this trade-off.

7.4 Influence of the number of antennas

Due to the metal induced loss, there will exist an optimum number N_{opt} of patterned antennas such that the SERS signal is maximized. Such an optimum is investigated in Figure 7.5 for a fixed $\alpha \approx 60^\circ$ bowtie geometry ($L \approx 106 \pm 8$ nm and $\Delta \approx 48 \pm 13$ nm) but varying N : $N = 10, 20, 30, 40, 70$ and $N = 0$ which is a reference waveguide. Each waveguide is measured 10 times and the averaged Raman spectra are reported in Figure 7.5(a). The $N = 70$ signal is not shown since it could not be distinguished from the inherent offset signal of the detector. The reference waveguide generates a considerable background signal, resulting in background associated shot noise which will again mask the smaller peaks (as described earlier). Functionalizing the waveguide with an increasing N reduces this unwanted background due to the attenuation caused by the nanoantennas. In addition,

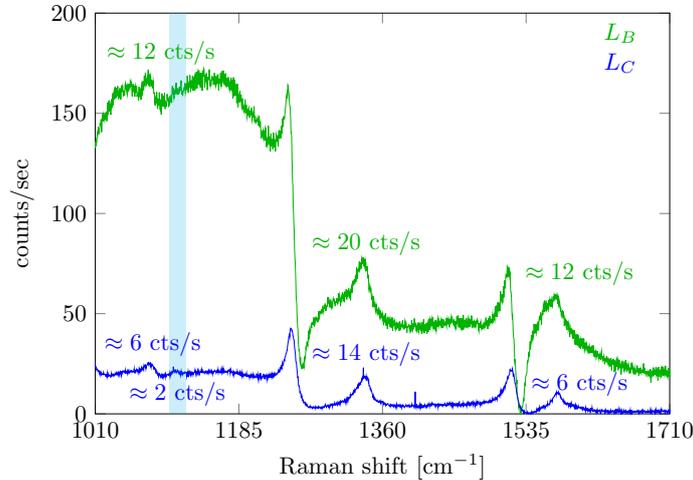


Figure 7.4: Zoom on the Raman spectra of the L_B and L_C structures (linear scale). Each spectrum is shifted by the minimum of the corresponding spectrum for improved visualization. The approximate signal counts of the Stokes peaks are also depicted. The 1110 cm^{-1} peak is marked by the cyan shaded area.

the dominant 1340 cm^{-1} peak starts to emerge when N increases. The smaller peaks at 1080 , 1110 and 1575 cm^{-1} only appear when the background is sufficiently low. A zoom on the 1340 cm^{-1} peak (cyan dashed line) is shown in Figure 7.5(b). For clarity the background is locally subtracted. The signal reaches a maximum value for $10 \leq N \leq 20$ and then decays again with increasing N . The waveguides functionalized with $N = 10$ or $N = 20$ antennas hence show a better performance in terms of absolute signal enhancement compared to the $N = 40$ waveguide. In contrast, the reduced background in the $N = 40$ case allows to resolve the complete spectrum. In that respect the $N = 40$ waveguide exhibits a better performance, albeit at a reduced limit of detection since the signal enhancement is decreased. For a fixed N we have shown in the previous section that one might need to select a non-optimal antenna geometry (L_C over L_B) that allows a complete resolution of the spectrum at the cost of a lower SERS enhancement. Here, a non-optimal antenna number ($N = 40$) for a fixed antenna geometry allows a better resolution of the total spectrum, but again at reduced SERS enhancement. Apart from signal optimization it is hence equally important to reduce the SiN background such that the smallest spectral features can be distinguished. In order to outline the interplay between signal enhancement and background reduction, we

have developed an analytical model that incorporates all relevant design parameters and elucidates on the absolute detection performance.

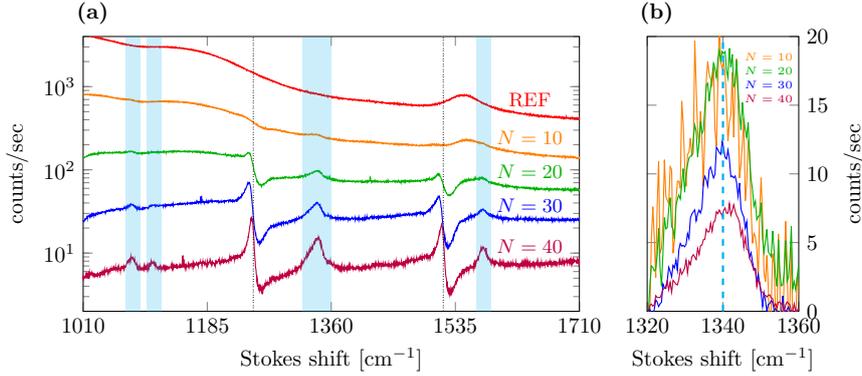


Figure 7.5: (a) Raman spectra of a reference waveguide (REF) and waveguides functionalized with $N = 10, 20, 30, 40$ antennas. (b) Zoom on the 1340 cm^{-1} signal peak. The background is locally subtracted to obtain the pure NTP signal. Taken from [28].

7.5 Analytical on-chip SERS model

For the derivation of our on-chip SERS model we start from the the mode-coupled power formula as discussed in an earlier chapter. The power $P_{wg}(\mathbf{r}_e, \lambda)$ coupled into the forward propagating waveguide mode $\mathbf{E}^m(\mathbf{r}, \lambda)$ as a result of a radiating dipole at position \mathbf{r}_e is given by

$$\frac{P_{wg}(\mathbf{r}_e, \lambda)}{P_0} = \frac{3}{8\pi} \frac{n_g(\lambda)}{n_m} \left(\frac{\lambda}{n_m} \right)^2 \frac{\epsilon_o \epsilon(\mathbf{r}_e) |\mathbf{e}_d \cdot \mathbf{E}^m(\mathbf{r}_e, \lambda)|^2}{\int \int_{A_{wg}} \epsilon_o \epsilon(\mathbf{r}) |\mathbf{E}^m(\mathbf{r}, \lambda)|^2 d\mathbf{r} |_{z=z_0}} \quad (7.1)$$

where $P_0 = \frac{\omega^4 |\mathbf{d}|^2}{12\pi\epsilon_0 c^3}$ is the power radiated in free space at a wavelength $\lambda = \frac{2\pi c}{\omega}$, $\epsilon(\mathbf{r})$ the relative permittivity at position \mathbf{r} , $n_g(\lambda)$ the group index of the waveguide mode, n_m the refractive index of the medium in which the dipole is placed and \mathbf{e}_d the unit vector along the dipole moment vector \mathbf{d} . [30] The integral in the denominator is calculated over a waveguide cross-section A_{wg} in the xy -plane and evaluated at the dipole position $z = z_0$ (coordinate axes are defined in Figure 7.6(a) and 7.6(b)).

7.5.1 SERS signal

Functionalizing the waveguide with an antenna however introduces a perturbation to the modal field $\mathbf{E}^m(\mathbf{r}, \lambda)$ of the waveguide and breaks the longitudinal invariance. Moreover, the above formula assumes a real dielectric permittivity, which does not hold for metals. In order to investigate whether formula (7.1) can still be used when the waveguide is functionalized with a metallic antenna, two sets of simulations were performed. In the first set (Figure 7.6(a)), the coupling of dipole radiation into the fundamental TE-mode is investigated. This allows an explicit evaluation of the left-hand side of equation (7.1). Dipole sources with fixed dipole moment vector \mathbf{d} were placed at different positions \mathbf{r}_e around the antenna (shown as black dots in Figure 7.6(a)). In the second set (Figure 7.6(b)), the fundamental TE-mode is launched into the waveguide functionalized with the same antenna in order to calculate the fields at positions \mathbf{r}_e where the dipole sources in the first simulation set were located. In this way it is possible to calculate the right-hand side of equation (7.1). The integral in the denominator is approximated in all calculations by evaluating it on a reference waveguide without antenna, so using the true modal fields and not the perturbed fields. This approximation allows a sufficient accurate evaluation of the right-hand side of equation (7.1) as is confirmed in Figure 7.6(c) where the simulation results are depicted. The red curve represents an incoherent superposition of the power coupled into the TE-mode as a result of the complete set of dipole emitters. The dashed blue curve is the predicted power that couples into the TE-mode, and is calculated as an incoherent superposition of the predicted coupled powers for each dipole. One can see that the predicted power coupled into the TE-mode (simulation set 2) matches very well with the explicit calculation using dipole emitters (simulation set 1). So equation (7.1) still gives an accurate prediction of the mode-coupled power when the waveguide is functionalized with a nanoplasmonic antenna.

Since the right-hand side of equation (7.1) faithfully represents the power coupled into the fundamental TE-mode, it is used to derive an analytical model predicting the total Stokes power coupled into the waveguide as a result of an array of N antennas with period Λ on top of the waveguide. Each array is centered on the waveguide with a distance L_1 to the front and back facet of the chip (see Figure 7.7(a)). For antennas with a symmetry axis along the y -axis (see Figures 7.6(a-b)), the NTP monolayer on each antenna will generate an equally large forward and backward propagating Stokes power $P_A(\lambda_P, \lambda_S) = \eta_A(\lambda_P, \lambda_S)P_{pump}$ for a given pump power

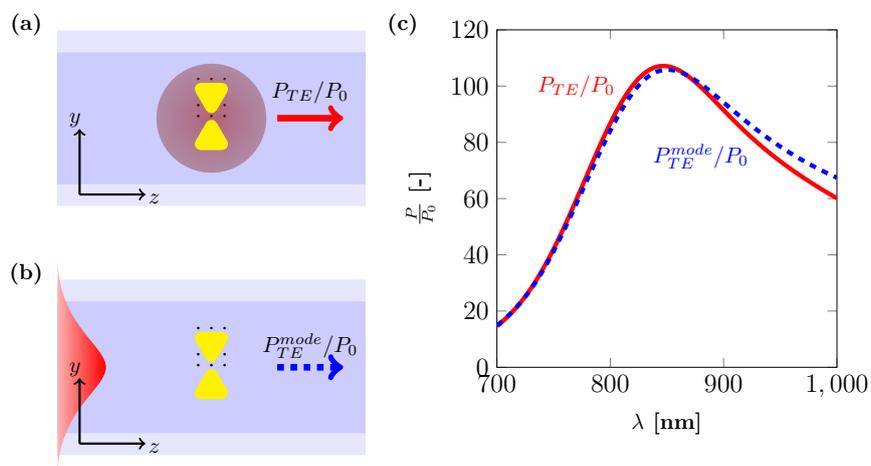


Figure 7.6: **(a)** Simulation set 1: collection of incoherent dipoles emitting into the fundamental TE-mode (red). **(b)** Simulation set 2: Calculation of the electric fields at the dipole positions of simulation set 1 using a fundamental TE-mode excitation. Based on these fields one can extract the predicted dipole emission (dashed blue). **(c)** Simulation results: incoherent superposition of the TE-coupled power due to all dipoles (red) and predicted TE-coupled power using the electric field values (dashed blue).

P_{pump} . The single antenna conversion efficiency $\eta_A(\lambda_P, \lambda_S)$ is an antenna dependent factor incorporating the integrated field enhancement profile near the metal surface and the molecular density and Raman cross-section. The total transmission loss induced by one antenna at wavelength λ is given by $1 - e_\lambda^{-1}$, whereby e_λ is the linear antenna extinction. Apart from the intrinsic waveguide losses α_{wg} , the pump and Stokes light will hence also be attenuated by e_P and e_S respectively. Two distinct cases can now be considered. In the forward scattering case, the forward propagating Stokes signals are collected at the output facet of the chip as shown in Figure 7.7(b), while in the backward scattering case the backscattered Stokes light from all antennas is collected at the input facet (Figure 7.7(c)). Each of these two cases will be considered in the following subsections.

7.5.1.1 Forward scattering

In this subsection we will outline the calculation of the total forward propagating Stokes power P_{tot}^F . First of all we write the dipole strength $|\mathbf{d}|^2$ as a function of the guided pump power P_{pump} in the waveguide

$$|\mathbf{d}|^2 = \alpha_m^2 |\mathbf{E}_d^m(\mathbf{r}_e, \lambda)|^2 \frac{n_g(\lambda) P_{pump}}{\int \int c \epsilon_o \epsilon(\mathbf{r}) |\mathbf{E}^m(\mathbf{r}, \lambda)|^2 d\mathbf{r} |_{z=z_{input}}} \quad (7.2)$$

where α_m is the molecular polarizability and $\mathbf{E}_d^m(\mathbf{r}_e, \lambda) = \mathbf{e}_d \cdot \mathbf{E}^m(\mathbf{r}_e, \lambda)$ the field strength along the dipole direction. The integral is evaluated at the input $z_{input} = 0$ of the waveguide. [13] The quantity

$$\eta(\mathbf{r}_e, \lambda_P, \lambda_S) = \frac{n_g(\lambda_P) n_g(\lambda_S) \lambda_S^2 |\mathbf{E}_d^m(\mathbf{r}_e, \lambda_P)|^2 |\mathbf{E}_d^m(\mathbf{r}_e, \lambda_S)|^2}{n_m \left(\int \int \epsilon(\mathbf{r}) |\mathbf{E}^m(\mathbf{r}, \lambda_P)|^2 d\mathbf{r} \right) \left(\int \int \epsilon(\mathbf{r}) |\mathbf{E}^m(\mathbf{r}, \lambda_S)|^2 d\mathbf{r} \right)} \quad (7.3)$$

solely depends on the modal fields and can hence be evaluated on a reference waveguide. Both integrals in the denominator are evaluated along the same cross-sectional area of the reference waveguide. The power P_{wg}^n coupled into the forward propagating TE-mode due to a collection of incoherently radiating dipoles around antenna n ($n = 1 \dots N$) is then approximately given by

$$\frac{P_{wg}^n}{P_{pump}} \approx \left(\frac{\rho\sigma(\lambda_S)}{2} e^{-\alpha(\lambda_P)(L_1+(n-1)\Lambda)} e_P^{1-n} \right) \times \int \int \int_{V_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e \quad (7.4)$$

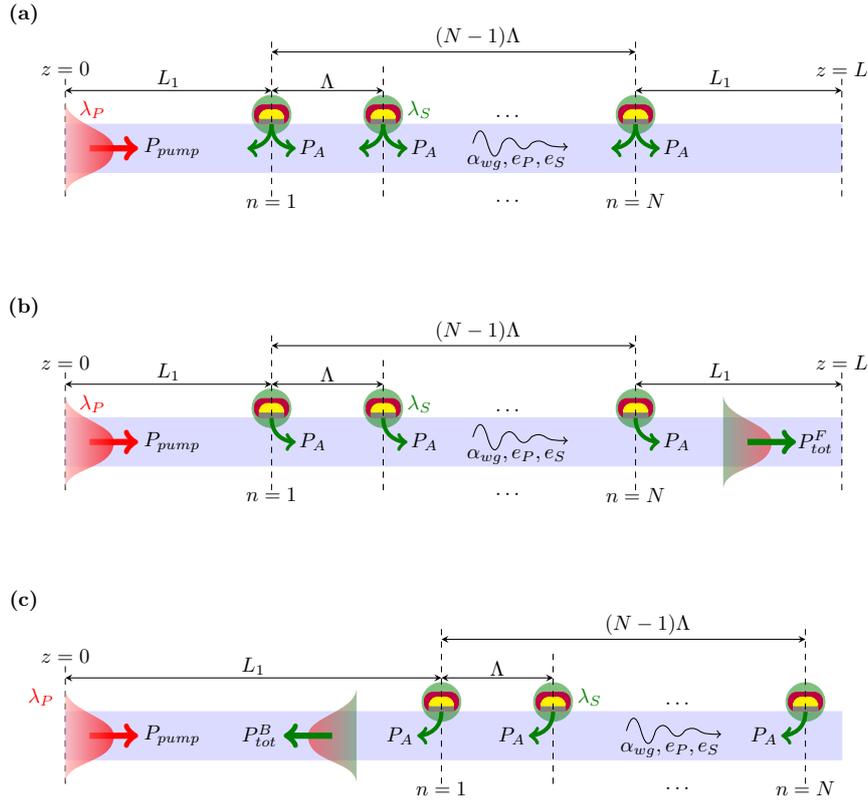


Figure 7.7: **(a)** Longitudinal cross-section of the chip. The pump beam, with power P_{pump} at wavelength λ_P , excites a Stokes signal λ_S of which the total power at the output facet is P_{tot} . The plasmonic array consists of N antennas with period Λ and is separated L_1 from both the input and output facet of the chip. Each of the antennas generates P_A guided Stokes power for a given pump power (in both forward and backward direction). Apart from the waveguide losses α_{wg} , the pump and Stokes light is attenuated by the pump e_P and Stokes e_S extinction respectively. **(b)** Forward scattering case. **(c)** Backward scattering case. Adapted from [28].

in which ρ is the molecular density, $\sigma(\lambda_S) = \pi^2 \alpha_m^2 / (\epsilon_0^2 \lambda_S^4)$ the Raman cross section [31], $\alpha(\lambda_P) = \alpha_P$ the waveguide loss at the pump wavelength, e_P the linear antenna extinction at the pump wavelength, V_m the volume in which the molecules are situated and $EF(\mathbf{r}_e, \lambda_{P/S})$ the field enhancement factor, at the dipole position \mathbf{r}_e , at the pump (P) and Stokes (S) wavelength respectively:

$$EF(\mathbf{r}_e, \lambda) = \frac{|\mathbf{E}_d^{ant}(\mathbf{r}_e, \lambda)|}{|\mathbf{E}_d^m(\mathbf{r}_e, \lambda)|} \quad (7.5)$$

whereby $\mathbf{E}_d^{ant}(\mathbf{r}_e, \lambda)$ is the local field around the antenna surface. For notational simplicity we rewrite formula (7.4) as

$$\frac{P_{wg}^n}{P_{pump}} \approx \eta_A(\lambda_P, \lambda_S) e^{-\alpha_P(L_1+(n-1)\Lambda)} e_P^{1-n}, \quad (7.6)$$

where the single antenna conversion efficiency

$$\eta_A(\lambda_P, \lambda_S) = \frac{\rho \sigma(\lambda_S)}{2} \int \int \int_{V_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e \quad (7.7)$$

is a dimensionless antenna dependent factor which incorporates the specific field enhancement profile near the antenna surface. Formula (7.4) takes into account the waveguide loss and the loss induced by the antennas in front of the n^{th} antenna as a result of which the actual excitation power decays. The signal (at Stokes wavelength λ_S) still has to pass $N - n$ antennas and propagate along a distance $L_1 + (N - n)\Lambda$, such that the power $P_{wg,out}^n$ reaching the output of the waveguide is given by

$$\begin{aligned} & \frac{P_{wg,out}^n}{P_{pump}} \\ & \approx \eta_A(\lambda_P, \lambda_S) \left(e^{-\alpha_P(L_1+(n-1)\Lambda)} e_P^{1-n} \right) \left(e^{-\alpha_S(L_1+(N-n)\Lambda)} e_S^{n-N} \right) \\ & = \eta_A(\lambda_P, \lambda_S) \left(e^{-\alpha_P(L_1-\Lambda)-\alpha_S(L_1+N\Lambda)} e_P e_S^{-N} \right) \left(\left(\frac{e_S}{e_P} \right) e^{(\alpha_S-\alpha_P)\Lambda} \right)^n. \end{aligned} \quad (7.8)$$

Since N antennas contribute incoherently to the signal, the total amount of Stokes light at the output of the waveguide is given by

$$\begin{aligned} \frac{P_{wg,out}^{tot}}{P_{pump}} & \approx \eta_A(\lambda_P, \lambda_S) \left(e^{-\alpha_P(L_1-\Lambda)-\alpha_S(L_1+N\Lambda)} e_P e_S^{-N} \right) \times \\ & \sum_{n=1}^N \left(\left(\frac{e_S}{e_P} \right) e^{(\alpha_S-\alpha_P)\Lambda} \right)^n. \end{aligned} \quad (7.9)$$

For the considered waveguide platform it is reasonable to approximate $\alpha_P \approx \alpha_S = \alpha_{wg}$ such that the forward propagating Raman power is given by

$$\begin{aligned}
 \frac{P_{wg,out}^{tot}}{P_{pump}} &= \frac{P_{tot}^F}{P_{pump}} \\
 &\approx \eta_A(\lambda_P, \lambda_S) \left(e^{-\alpha_{wg}(2L_1+(N-1)\Lambda)} e_P e_S^{-N} \right) \sum_{n=1}^N \left(\frac{e_S}{e_P} \right)^n \\
 &\approx \eta_A(\lambda_P, \lambda_S) \left(e^{-2\alpha_{wg}L_1} e_S^{1-N} \right) \left(\frac{1 - \left(\frac{e_S}{e_P} \right)^N}{1 - \left(\frac{e_S}{e_P} \right)} \right) \\
 &= FOM_F(N, \lambda_P, \lambda_S) e^{-2\alpha_{wg}L_1}, \tag{7.10}
 \end{aligned}$$

where $L = 2L_1 + (N - 1)\Lambda \approx 2L_1$ has been used. This result applies to Stokes light co-propagating with the pump beam. The quantity

$$\boxed{FOM_F(N, \lambda_P, \lambda_S) = \eta_A(\lambda_P, \lambda_S) e_S^{1-N} \left(\frac{1 - \left(\frac{e_S}{e_P} \right)^N}{1 - \left(\frac{e_S}{e_P} \right)} \right)} \tag{7.11}$$

contains all necessary parameters to assess the SERS signal strength for a given waveguide geometry and is hence considered to be the relevant figure of merit (*FOM*) in comparing the performance of integrated antenna arrays. This equation is one of the key results of our on-chip SERS work and furtheron it is shown that the experimentally obtained SERS signals are in good correspondence with the above formula. It is important to note that the *FOM* also incorporates the modal properties of the waveguide through $\eta_A(\lambda_P, \lambda_S)$ and is hence also dependent on the exact waveguide geometry and excitation polarization. For a fixed antenna geometry and λ_P and λ_S , the optimum number of antennas N_{opt} that should be patterned on a waveguide to maximize the SERS signal is determined by

$$\left. \frac{dFOM_F(N, \lambda_P, \lambda_S)}{dN} \right|_{N_{opt}} = 0 \Rightarrow N_{opt} = \frac{\log\left(\frac{\log(e_S)}{\log(e_P)}\right)}{\log\left(\frac{e_S}{e_P}\right)}. \tag{7.12}$$

For the specific case of a monolayer, one has a molecular surface density ρ_s (number of molecules per m^2) rather than a volume density ρ

(number of molecules per m^3). In this limiting case, the single antenna conversion efficiency is theoretically defined by

$$\eta_A(\lambda_P, \lambda_S) = \frac{\rho_s \sigma(\lambda_S)}{2} \int \int_{A_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e \quad (7.13)$$

where the integration is now performed over the surface area A_m covered by the monolayer. Numerically one however always needs to model the presence of such a monolayer by introducing a finite thickness t_m . As a result, an effective monolayer volume $V_m \approx t_m \times A_m$ is defined. For any numerical evaluation the fields are hence integrated over this volume V_m such that

$$\int \int \int_{V_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e \approx t_m \times \int \int_{A_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e \quad (7.14)$$

The single antenna conversion efficiency is then given by

$$\begin{aligned} \eta_A(\lambda_P, \lambda_S) \\ \approx \frac{\rho_s \sigma(\lambda_S)}{2t_m} \int \int \int_{V_m} \eta(\mathbf{r}_e, \lambda_P, \lambda_S) EF(\mathbf{r}_e, \lambda_P)^2 EF(\mathbf{r}_e, \lambda_S)^2 d\mathbf{r}_e. \end{aligned} \quad (7.15)$$

From a numerical point of view, an effective volume density $\rho_{eff} = \rho_s/t_m$ is introduced, which is then multiplied with the integral over the effective monolayer volume V_m .

7.5.1.2 Backward scattering

The backward scattering case can be treated similarly to the forward scattering case. Using the same terminology and reasoning as in the previous subsection, we write the power P_{wg}^n that is coupled into the backward propagating TE-mode due to a collection of incoherently radiating dipoles around antenna n ($n = 1 \dots N$) as

$$\frac{P_{wg}^n}{P_{pump}} \approx \eta_A(\lambda_P, \lambda_S) e^{-\alpha_P(L_1+(n-1)\Lambda)} e_P^{1-n} \quad (7.16)$$

since the pump power is still propagating in the forward direction. Contrary to the forward scattering case, the Stokes light generated at antenna n now has to propagate back along the previous $n-1$ antennas and propagate along

a distance $L_1 + (n - 1)\Lambda$, such that the power $P_{wg,in}^n$ reaching the input of the waveguide is given by

$$\frac{P_{wg,in}^n}{P_{pump}} \approx \eta_A(\lambda_P, \lambda_S) \left(e^{-\alpha_P(L_1+(n-1)\Lambda)} e_P^{1-n} \right) \left(e^{-\alpha_S(L_1+(n-1)\Lambda)} e_S^{1-n} \right). \quad (7.17)$$

The incoherent superposition of the Stokes signals from N antennas results in a total amount of Stokes light at the input of the waveguide of

$$\begin{aligned} \frac{P_{wg,in}^{tot}}{P_{pump}} &= \frac{P_{tot}^B}{P_{pump}} \\ &\approx \eta_A(\lambda_P, \lambda_S) e^{-2L_1\alpha_{wg}} e_P e_S \sum_{n=1}^N \left(\frac{1}{e_P e_S} \right)^n \\ &\approx \eta_A(\lambda_P, \lambda_S) e^{-2L_1\alpha_{wg}} \left(\frac{1 - \left(\frac{1}{e_P e_S} \right)^N}{1 - \left(\frac{1}{e_P e_S} \right)} \right) \\ &= FOM_B(N, \lambda_P, \lambda_S) e^{-2\alpha_{wg}L_1}, \end{aligned} \quad (7.18)$$

where $\alpha_P \approx \alpha_S = \alpha_{wg}$ and $\alpha_{wg}\Lambda \approx 0$ have been used. Similar to the earlier defined figure of merit, $FOM_B(N, \lambda_P, \lambda_S)$ contains all relevant parameters to assess the backscattered Stokes signal:

$$FOM_B(N, \lambda_P, \lambda_S) = \eta_A(\lambda_P, \lambda_S) \left(\frac{1 - \left(\frac{1}{e_P e_S} \right)^N}{1 - \left(\frac{1}{e_P e_S} \right)} \right) \quad (7.19)$$

Contrary to forward scattering, the backscattered signal has no optimum antenna number, but saturates for $N \rightarrow \infty$ as shown in Figure 7.8. Moreover it is clear that the maximum backscattered signal always equals or exceeds the maximum forward scattered signal. This can be understood from the fact that the signal from the first antenna, which is the strongest, is fully available at the output, as opposed to the forward scattering case. The signal from subsequent antennas will then experience an ever increasing attenuation by the previous antennas, such that the backscattering signal eventually saturates for a sufficiently large number of antennas. It is hence beneficial to collect the backscattered light. However, in that case the previously described alignment procedure cannot be used and other optimized alignment schemes for backscattering collection need to be developed.

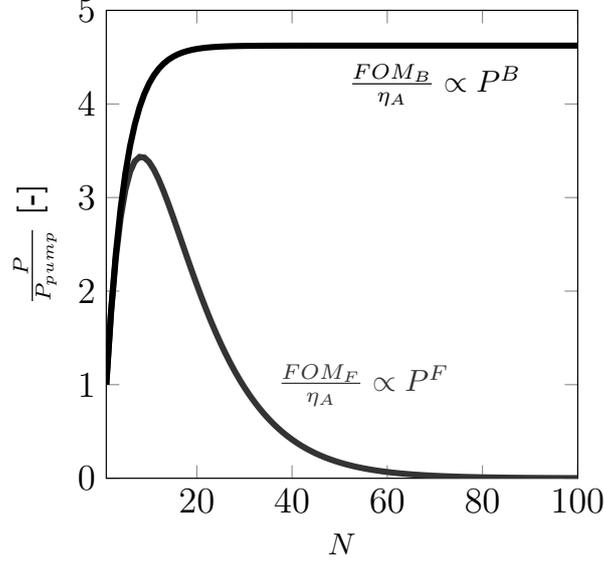


Figure 7.8: Comparison of the forward P^F and backward P^B scattering for $e_P = 1.16$ and $e_S = 1.10$.

7.5.2 Relation to the measurable signal

Since the experimental work described throughout this chapter involves forward Stokes scattering, we will stick to that case furtheron. In order to correlate the Stokes power at the output facet of the chip and the actual detected signal counts C_S , normalized with the integration time T , we need to take into account the characteristics of all optics in our setup. The Stokes power reaching the detector surface is given by

$$\begin{aligned}
 P_S &= \gamma^{out} \times T_{out} \times P_{tot} = \gamma^{out} \times T_{out} \times (FOM e^{-\alpha_{wg}L} P_{pump}) \\
 &= \gamma^{out} \times T_{out} \times (FOM e^{-\alpha_{wg}L} (\gamma^{in} P_{in})) \\
 &= FOM (\gamma^{in} \gamma^{out} P_{in} e^{-\alpha_{wg}L}) \times T_{out}
 \end{aligned} \tag{7.20}$$

where γ^{in} and γ^{out} are the coupling efficiencies in and out of the chip respectively, P_{in} is the power just before the input facet of the chip and T_{out} is the transmission through all optics between the output of the chip and the detector surface. The quantity $\gamma^{in} \gamma^{out} P_{in} e^{-\alpha_{wg}L}$ can be written as

$$\gamma^{in} \gamma^{out} P_{in} e^{-\alpha_{wg}L} = P_T / T_{PM} \tag{7.21}$$

in which P_T is the transmitted power as measured by the power meter (see the description of the measurement setup in a previous chapter) and T_{PM}

is the optical transmission between the output of the chip and the input of the powermeter. This transmission is related to T_{out} by $T_{out} \approx 100 \times T_{PM} \times T_{spec}$ in which T_{spec} is the transmission between the input slit of the spectrometer and the detector surface. The factor 100 stems from the fact that the fiber splitter only transmits 1% of the power to the power meter. So eventually P_S can be written as $P_S = FOM \times P_T \times 100T_{spec}$. By using the spectrometer sensitivity χ (defined as the number of electrons per count) and quantum efficiency $QE(\lambda_S)$ (number of electrons per number of incident photons), P_S can be related to the measured signal counts C_S per unit integration time T by

$$P_S = \left(\frac{C_S}{T}\right) \left(\frac{hc}{\lambda_S}\right) \left(\frac{\chi}{QE(\lambda_S)}\right). \quad (7.22)$$

Finally one gets

$$\frac{C_S}{T} = FOM \times 100T_{spec} \left(\frac{\lambda_S}{hc}\right) \left(\frac{QE(\lambda_S)}{\chi}\right) P_T = FOM \times C^*(\lambda_S). \quad (7.23)$$

The transmitted power $100P_T$ is about 1mW for the given measurement setup. If we consider e.g. the 1340 cm^{-1} line then $\frac{QE(\lambda_S)}{\chi} \approx 1$ counts/photon and $T_{spec} \approx 0.44$ (using data supplied by the manufacturer) such that $C^* \approx 1.94 \times 10^{15}$ counts/sec. $C^*(\lambda_S)$ can now be used as a conversion factor when the experimental signal counts are fitted to the analytical model. In this way one can rigorously quantify the parameters e_P , e_S and η_A and compare them with the theoretical predictions.

7.5.3 Background signal

As previously touched upon, the background signal generated by the SiN introduces a trade-off between signal enhancement and resolving the spectrum. Therefore it is also important to derive an analytical formula for this background signal in order to elucidate these effects. Similar reasonings as the ones described in the previous subsections can be applied to derive the total background signal generated along the waveguide. We will only derive the forward scattered background, however the same analysis can be done for the backward scattered background. The background mostly stems from the SiN core and is hence generated when the pump light propagates along the total length $L = 2L_1 + (N - 1)\Lambda$ of the waveguide. For a functionalized waveguide, the attenuation of the pump and Stokes light due to the antenna array also needs to be incorporated.

The dipoles which give rise to the background are mainly situated in the core, so for a given core cross-sectional area A_{core} , equation (7.3) has to be integrated over the complete waveguide core. This quantity is defined as

$$\eta_{core}(\lambda_P, \lambda_S) = \eta_c = \int \int_{A_{core}} \eta(\mathbf{r}_0, \lambda_P, \lambda_S) d\mathbf{r}_0. \quad (7.24)$$

For a given molecular core density ρ_c and core scattering cross-section σ_c , the total background P_{bg} propagating in the forward direction is then calculated through

$$\begin{aligned} & \frac{P_{bg}}{P_{pump}} \\ &= \frac{\rho_c \sigma_c \eta_c}{2} \left\{ \left(\int_0^{L_1} e^{-\alpha_P z} e^{-\alpha_S(L-z)} dz \right) e_S^{-N} \right. \\ &+ \left(\int_{L_1+(N-1)\Lambda}^{2L_1+(N-1)\Lambda} e^{-\alpha_P z} e^{-\alpha_S(L-z)} dz \right) e_P^{-N} \\ &+ \left. \sum_{n=0}^{N-2} \left(\int_{L_1+n\Lambda}^{L_1+(n+1)\Lambda} e^{-\alpha_P z} e^{-\alpha_S(L-z)} dz \right) e_P^{-(n+1)} e_S^{-(N-n-1)} \right\}. \quad (7.25) \end{aligned}$$

Using $\alpha_P \approx \alpha_S = \alpha_{wg}$ and $L \approx 2L_1$ ($2L_1 \gg (N-1)\Lambda$), the total forward propagating background is then given by

$$\begin{aligned} & \frac{P_{bg}}{P_{pump}} \\ &= \frac{\rho_c \sigma_c \eta_c}{2} e^{-\alpha_{wg} L} \left\{ \left(e_P^{-N} + e_S^{-N} \right) \frac{L}{2} + \Lambda \frac{e_S^{1-N}}{e_P} \left(\frac{1 - \left(\frac{e_S}{e_P} \right)^{N-1}}{1 - \left(\frac{e_S}{e_P} \right)} \right) \right\} \\ &\approx \frac{\rho_c \sigma_c \eta_c}{2} e^{-\alpha_{wg} L} \left\{ \left(e_P^{-N} + e_S^{-N} \right) \frac{L}{2} \right\} \\ &= \eta_B(\lambda_P, \lambda_S) e^{-\alpha_{wg} L} \left(e_P^{-N} + e_S^{-N} \right) \frac{L}{2} \\ &\approx \eta_B(\lambda_P, \lambda_S) e^{-2\alpha_{wg} L_1} \left(e_P^{-N} + e_S^{-N} \right) L_1. \quad (7.26) \end{aligned}$$

One can see that the background increases linearly with the total propagation length ($\propto L_1$) and decays exponentially with the number of antennas. In order to achieve the highest Signal-to-Noise Ratio (SNR) it is hence important to keep L_1 as short as possible. The previously defined conversion factor C^* can also be used to convert the background signal to background counts. In the next subsection we will combine the SERS and background signal and discuss the SNR of the forward scattering case.

7.5.4 Signal-to-Noise Ratio

Typically the overall noise of a CCD detector is given by

$$N = \sqrt{\delta_{dark}^2 + \delta_{readout}^2 + \delta_{shot}^2} \quad (7.27)$$

where δ_{dark}^2 is the variance of the dark noise, $\delta_{readout}^2$ the variance of the readout noise and δ_{shot}^2 the variance of the shot noise. [32, 33] The dark noise arises from thermally generated charges in the silicon sensor. Recent improvements in CCD design have greatly diminished dark noise and reduced their contribution to the total readout noise to less than 10 electrons per pixel at room temperature. For the ultimate sensitivity, cooling the CCD to low temperatures (-100°C) is still required. The readout noise is an inherent property of the sensor and usually sets the sensitivity limit for most cameras. The readout noise is a combination of noise sources, which originate from the process of amplifying and converting the photoelectrons into a voltage. Over the years readout noise has improved, but fundamentally the noise increases for faster camera readouts due to the increasing bandwidth that is required. Finally, the incoming photons have an inherent noise as well, known as shot noise. This shot noise is Poissonian distributed and determined by the square root of the total signal counts C_S ,

$$\delta_{shot} = \sqrt{C_S}. \quad (7.28)$$

[32, 33] For our particular measurements, the noise is dominated by shot noise such that

$$N \approx \sqrt{C_S}. \quad (7.29)$$

The total signal consists of the useful SERS signal S on one hand

$$S = FOM_F \times C^* \quad (7.30)$$

and the unwanted background signal B on the other hand

$$B = \eta_B (e_P^{-N} + e_S^{-N}) L_1 \times C^* \quad (7.31)$$

such that

$$N = \sqrt{C_S} = \sqrt{S + B} \quad (7.32)$$

where the conversion factor C^* is used to convert the signal power to signal counts. Defining $x = e_S/e_P$, we can write the SERS Signal-to-Noise Ratio

SNR as

$$SNR = \frac{S}{N} = \left(\sqrt{C^* N_{avg}} \right) \frac{\eta_A e_S^{1-N} \left(\frac{1-x^N}{1-x} \right)}{\sqrt{\eta_A e_S^{1-N} \left(\frac{1-x^N}{1-x} \right) + \eta_B L_1 e_S^{-N} (1+x^N)}} \quad (7.33)$$

In Figure 7.9(a) we plot the normalized P^F and SNR curves (normalized to 1). A zoom on the maxima of these curves is shown in Figure 7.9(b). In case the SiN background is negligible (i.e. $\eta_B L_1 = 0$), the SNR is maximal for the same N_{opt} as the signal. The shot noise is then purely determined by the SERS signal itself. When the SiN background becomes more dominant (larger $\eta_B L_1/\eta_A$ ratios), the SNR is maximal at an $N > N_{opt}$. Hence, the optimal SNR does not coincide with an optimal SERS enhancement in case the SiN background dominates on the SERS signal. This clearly confirms our earlier statement that a trade-off between spectral resolution (associated to SNR) and absolute SERS enhancement (associated to P^F) might occur due to the inherent SiN background. Moreover it is required

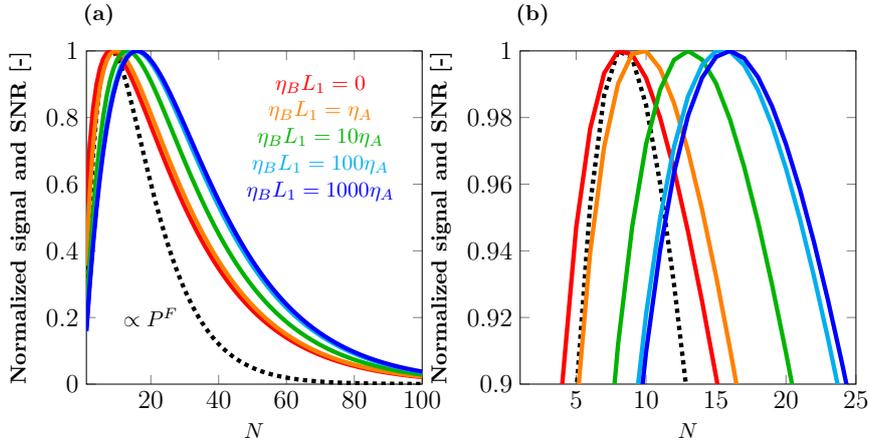


Figure 7.9: (a) Normalized P^F (black dotted) and SNR (colored) curves for $e_P = 1.16$, $e_S = 1.10$, $N_{avg} = 1$ and different $\eta_B L_1 / \eta_A$ ratios. (b) Zoom on the maxima of the curves.

that $SNR > 1$ for all Stokes peaks to allow detection. In Figure 7.10 we plot the absolute value of the SNR for $e_P = 1.16$, $e_S = 1.10$, $N_{avg} = 10$ and $\eta_A = 2 \times 10^{-15}$, $\eta_B L_1 = 200\eta_A$, which are realistic values for a typical experiment. One can see that a minimum number of antennas N_{min} is required in order to allow detection (i.e. to have $SNR > 1$). Once the

antenna number exceeds the maximum number of antennas N_{max} , the SNR again drops below 1. The red shaded areas depict zones where detection is impossible, while an antenna number within the blue shaded area does allow detection. In the most ideal case, one wants $N_{min} = 1$ such that detection of SERS signals from a single antenna is possible. As a result it would become possible to simultaneously probe large areas of analytes ($> \lambda^2$) and detecting all SERS events originating from different locations by monitoring just a single waveguide output, in contrast to microscopy based systems where one has to serially scan all hotspot locations. In the next chapter we will discuss a possible strategy to minimize the background associated shot noise in order to improve the SNR .

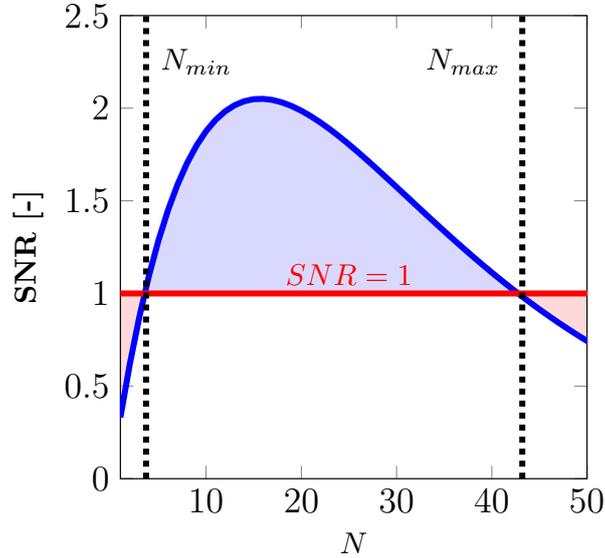


Figure 7.10: SNR for $e_P = 1.16$, $e_S = 1.10$, $N_{avg} = 10$ and $\eta_A = 2 \times 10^{-15}$, $\eta_B L_1 = 200\eta_A$.

7.6 Generalized on-chip SERS model

The model derived in the previous section assumes N identical antennas. However, due to fabrication errors there will always be differences among each of the antennas in the array. These differences have an impact on both the extinction e_P and e_S as well as on the single antenna conversion efficiency $\eta_A(\lambda_P, \lambda_S)$. In this section we develop a generalized on-

chip SERS model and describe a procedure to determine the potential deviations among different antennas (i.e. the disorder in the system) using a randomized fit to this generalized model. A similar procedure can be applied to the background associated shot noise as well.

We will only outline the forward scattering case for which the power generated by antenna n ($n = 1 \dots N$) is given by:

$$\frac{P_{wg}^n}{P_{pump}} \approx \eta_A(\lambda_P, \lambda_S) e^{-\alpha_P(L_1+(n-1)\Lambda)} e_P^{1-n}. \quad (7.34)$$

Since $L_1 \gg N\Lambda$ this can be simplified to

$$\frac{P_{wg}^n}{P_{pump}} \approx \eta_A(\lambda_P, \lambda_S) e^{-\alpha_P L_1} e_P^{1-n}. \quad (7.35)$$

The formula however assumes a constant pump extinction from the previous $n - 1$ antennas. Furthermore the factor $\eta_A(\lambda_P, \lambda_S)$ is also assumed constant for each antenna. If we allow that each of the antennas has a different extinction e_P^m and single antenna conversion efficiency $\eta_A^m(\lambda_P, \lambda_S)$ ($m = 1 \dots N$), then the power generated by antenna n is given by

$$\frac{P_{wg}^n}{P_{pump} e^{-\alpha_P L_1}} \approx \eta_A^n(\lambda_P, \lambda_S) \prod_{m=1}^{n-1} (e_P^m)^{-1}. \quad (7.36)$$

Applying a similar reasoning to the Stokes light that has to propagate along $N - n$ other antennas (and a length L_1), the Stokes power generated by antenna n that reaches the output is then given by

$$\frac{P_{wg,out}^n}{P_{pump} e^{-2\alpha_{wg} L_1}} \approx \eta_A^n(\lambda_P, \lambda_S) \left(\prod_{m=1}^{n-1} (e_P^m)^{-1} \right) \left(\prod_{m=n+1}^N (e_S^m)^{-1} \right). \quad (7.37)$$

The total generated Stokes power from all N (potentially different) antennas is then finally

$$\boxed{\frac{P_{tot}}{P_{pump} e^{-2\alpha_{wg} L_1}} \approx \sum_{n=1}^N \eta_A^n(\lambda_P, \lambda_S) \left(\prod_{m=1}^{n-1} (e_P^m)^{-1} \right) \left(\prod_{m=n+1}^N (e_S^m)^{-1} \right)}. \quad (7.38)$$

For the background one can analogously write

$$\boxed{\frac{P_{bg}}{P_{pump} e^{-2\alpha_{wg} L_1}} \approx \eta_B(\lambda_P, \lambda_S) L_1 \left(\prod_{m=1}^N (e_P^m)^{-1} + \prod_{m=1}^N (e_S^m)^{-1} \right)}. \quad (7.39)$$

In this case $\eta_B(\lambda_P, \lambda_S)L_1$ is a constant since it only depends on the waveguide parameters. The shot noise associated to the background is proportional to $\sqrt{P_{bg}}$. The number of counts/sec can be obtained by multiplying equations (7.38) and (7.39) with the earlier defined $C^*(\lambda_S)$. For a fixed λ_P and λ_S , $\eta_A(\lambda_P, \lambda_S)C^*(\lambda_S)$ will be denoted by η_A^* .

In order to determine the potential deviations among different antennas, we developed a randomized fitting procedure that can be applied to the experimental signal and shot noise data. This procedure will be outlined below. Initially, three random numbers are independently generated for each antenna ($n = 1 \dots N$) in the array: one for the pump extinction, one for the Stokes extinction and one for the single antenna conversion efficiency. The complete procedure is also depicted in Figure 7.11. This process is then repeated $N_{samples} = 1000$ times ($N_{samples}$ has to be sufficiently high to obtain a statistically relevant distribution). So for an array of N antennas the total number of randomly generated points is $3 \times N \times N_{samples}$. Using these random numbers, formulas (7.38) and (7.39) are subsequently evaluated in order to generate distributions of the signal and shot noise counts. Ultimately the mean value $\mu_S(N)/\mu_N(N)$ (blue and red dotted curves in the bottom graph) and standard deviation $\sigma_S(N)/\sigma_N(N)$ from the obtained signal (S) and shot noise (N) distributions are extracted for each number of antennas N . The 3σ -intervals for both distributions ($[\mu - 3\sigma, \mu + 3\sigma]$) define an area that marks the possible signal (blue area) and noise (red area) counts for a given uncertainty on the antenna parameters (according to formulas (7.38) and (7.39)). The standard deviations of the e_P , e_S and η_A distributions should be chosen such that all experimental signal and shot noise datapoints are covered by the generated signal and shot noise distributions respectively (provided all other constraints, outlined below, are also fulfilled); i.e. all experimental signal and shot noise data points should be contained in the $[\mu_S - 3\sigma_S, \mu_S + 3\sigma_S]$ interval and the $[\mu_N - 3\sigma_N, \mu_N + 3\sigma_N]$ interval respectively.

The mean value of each of the three normal distributions (e_P , e_S and η_A) is determined by a fit of the experimental data to the ideal model. One could obtain a perfect fit by fitting either the signal or shot noise data to their respective ideal analytical expression. These fitting parameters would however not generate a perfect fit to the other non-fitted equation, since both equations depend on e_P and e_S . Therefore a constrained fit, in which one minimizes the fitting error to both equations simultaneously, is performed such that both the signal and shot noise data are represented well using the fitted values for e_P and e_S . The fitted e_P , e_S and η_A^* values (mind that

η_A^* and not η_A is used because we fit the counts/sec) obtained from the (ideal) constrained fit are then chosen to be the mean values (μ_{e_P} , μ_{e_S} and $\mu_{\eta_A^*}$) for their respective normal distributions. The standard deviation σ_x is chosen such that the $3\sigma_x$ intervals represent realistic deviations from the mean value (x denotes one of the three parameters). For the extinction this means e.g. that $\mu_{e_{P/S}} - 3\sigma_{e_{P/S}} \geq 1$ (because 1 is the lower boundary for the linear extinction). Since η_A^* is always positive, $\mu_{\eta_A^*} - 3\sigma_{\eta_A^*} > 0$ should also be satisfied. In the end, this procedure eventually allows to determine the deviations among different antennas (i.e. σ_{e_P} , σ_{e_S} and $\sigma_{\eta_A^*}/\sigma_{\eta_A}$) based on a randomized fit to the signal and shot noise data. This procedure will be applied in the next section to a concrete experimental example.

7.7 Comparison between theory and experiment

For the particular bowtie antenna studied in Figure 7.5 ($\alpha \approx 60^\circ$, $L \approx 106 \pm 8$ nm and $\Delta \approx 48 \pm 13$ nm), the extinction spectrum $e(\lambda)$ and single antenna conversion efficiency $\eta_A(\lambda_P, \lambda_S)$ are numerically evaluated using Lumerical FDTD Solutions. These theoretical predictions are subsequently compared with the experimental data.

7.7.1 Simulation setup

We used a refractive index of $n_{rib} = 1.9$ for the SiN rib (with width $w_{rib} = 700$ nm and height $h_{rib} = 220$ nm), $n_{bclad} = 1.45$ for the SiO₂ bottomcladding and $n_{tclad} = 1$ for the top cladding (air). The bowtie geometry parameters were set to $\alpha = 60^\circ$, $L = 106$ nm and $\Delta = 48$ nm. A thin native oxide layer ($t_{nox} = 2$ nm) between the SiN and the Ti has also been incorporated. [18] The metal stack thicknesses are fixed to $t_{Ti} = 2$ nm and $t_{Au} = 30$ nm and a built-in refractive index model for Au (Johnson and Christy [34]) and Ti (CRC [35]) is used. An additional surface layer with thickness $t_{NTP} = 1$ nm and index $n_{NTP} = 3$ is used to model the NTP monolayer. The antenna region (including the Ti adhesion layer and the NTP monolayer) is meshed with a uniform mesh of 0.5 nm in the plane of the antenna (yz -plane) and 2 nm in the x -direction. A mesh refinement to 1 nm (0.5 nm) is applied in regions where the thickness in the x -direction is ≤ 2 nm (≤ 1 nm). The estimated surface area of an NTP molecule is 0.18 nm², so the surface density is then $\rho_s = 5.56 \times 10^{18}$ molecules/m². [29] The Raman cross-section is $\sigma \approx 0.358 \times 10^{-30}$ cm²/sr, which was obtained by applying the λ_S^{-4} scaling to the original data of the 1340 cm⁻¹ line. [36]

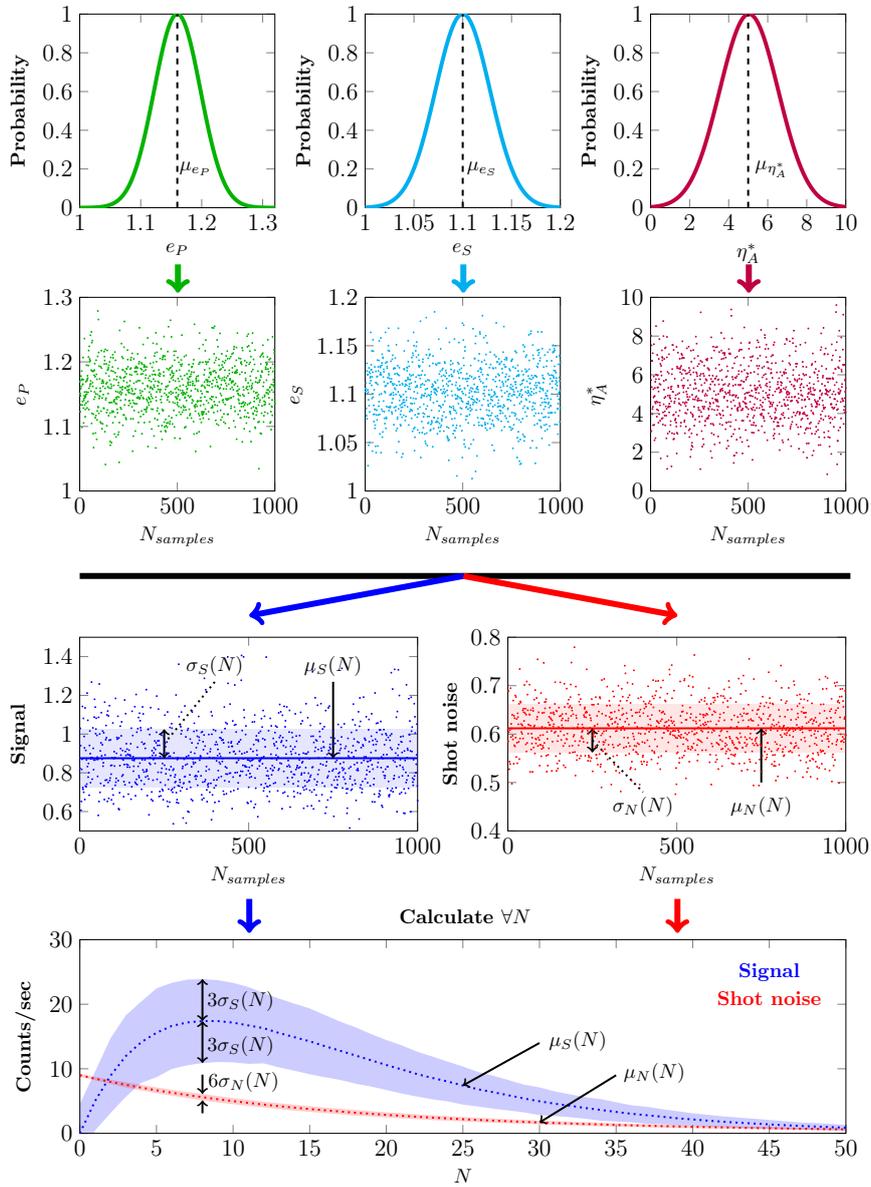


Figure 7.11: Randomized fitting procedure in order to determine the deviations among different antennas through the signal (blue area) and shot noise (red area) distributions, which are in itself generated by the e_P , e_S and η_A^* distributions.

Single antenna extinction spectra $E(\lambda)$ (in dB) are calculated through $E(\lambda) = T_{ref}(\lambda) - T_{ant}(\lambda)$ in which $T_{ref}(\lambda)$ is the power transmission (in dB) through the reference waveguide and $T_{ant}(\lambda)$ the power transmission (in dB) of a waveguide functionalized with one antenna. Linear extinction spectra $e(\lambda) \triangleq e_\lambda$ are then given by $e(\lambda) = 10^{E(\lambda)/10}$. A field and index profile monitor are used to extract the local field $|\mathbf{E}(\mathbf{r}, \lambda)|$ and index $n(\mathbf{r})$ around the antenna. The single antenna conversion efficiency

$$\begin{aligned} \eta_A(\lambda_P, \lambda_S) &= \frac{\rho_s \sigma}{2t_{NTP}} \frac{\int \int \int_{V_m} n_g(\lambda_P) n_g(\lambda_S) \lambda_S^2 |\mathbf{E}(\mathbf{r}, \lambda_P)|^2 |\mathbf{E}(\mathbf{r}, \lambda_S)|^2 d\mathbf{r}}{\left(\int \int n(\mathbf{r})^2 |\mathbf{E}^m(\mathbf{r}, \lambda_P)|^2 d\mathbf{r} \right) \left(\int \int n(\mathbf{r})^2 |\mathbf{E}^m(\mathbf{r}, \lambda_S)|^2 d\mathbf{r} \right)} \end{aligned} \quad (7.40)$$

is calculated by integrating the local fields over the effective monolayer volume V_m in which the index satisfies $n(\mathbf{r})|_{\mathbf{r} \in V_m} = n_{NTP}$. The denominator is calculated using the modal fields $\mathbf{E}^m(\mathbf{r}, \lambda)$ of a non-functionalized reference waveguide and the local field enhancement is given by the ratio of the local and modal electric fields: $EF(\mathbf{r}, \lambda) = \frac{|\mathbf{E}(\mathbf{r}, \lambda)|}{|\mathbf{E}^m(\mathbf{r}, \lambda)|}$. At a certain position, the Raman enhancement factor EF_R is calculated as $EF_R = EF(\lambda_P)^2 EF(\lambda_S)^2$.

The simulated extinctions are $e_P = 1.16$ ($E_P = 0.63$ dB) and $e_S = 1.10$ ($E_S = 0.43$ dB) if we consider a pump wavelength $\lambda_P = 785$ nm and the 1340 cm^{-1} Stokes peak. The predicted N_{opt} for this 1340 cm^{-1} peak is 8 to 9 antennas, while $\eta_A \approx 2.19 \times 10^{-15}$. For each 1W of pump power the antenna is therefore expected to generate 2.19 fW of guided Stokes power. Figure 7.12 shows the Raman enhancement EF_R for the 1340 cm^{-1} line, evaluated halfway the gold thickness. For example, in the center of the gap at 5 nm from the tip of the antenna (marked by the black dot in Figure 7.3(a)) an $EF_R \approx 1.14 \times 10^4$ is expected. Due to the large gap, the Raman enhancement is relatively low. However, fabrication induced variations on a larger gap will have less pronounced effects on EF_R as compared to variations on a very narrow gap. Since our primary goal is to investigate the properties of this platform and establish whether our analytical model can faithfully reproduce the experimental data, it is more convenient to work with a larger gap and avoid possible large variations in EF_R . This allows a more reliable comparison between theory and experiment.

Furthermore, it is interesting to assess the efficiency with which the total emitted Stokes radiation is coupled to the fundamental TE-mode. The β -factor is a measure for this efficiency and is determined by the ratio

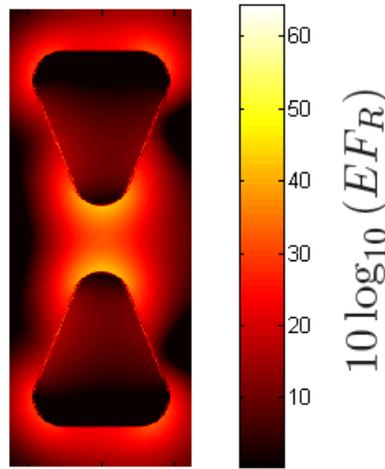


Figure 7.12: Raman enhancement EF_R for the 1340 cm^{-1} line, evaluated halfway the gold thickness for a bowtie with dimensions $\alpha = 60^\circ$, $L = 106\text{ nm}$ and $\Delta = 48\text{ nm}$.

of the Stokes power coupled into the fundamental TE-mode P_{TE} and the total radiated Stokes power P_{rad} , $\beta = P_{TE}/P_{rad}$. For the calculations we considered a fixed dipole emitter near the antenna surface. P_{rad} can then be calculated using a 3D power monitor of which the boundaries are in the far field of the dipole source. P_{TE} can be simultaneously determined using a mode expansion monitor which extracts the power carried by the TE-mode. For the 1340 cm^{-1} line, $\beta \approx 17\%$ for a dipole in the center of the gap and at 15 nm above the top surface of the waveguide. The total modal power is equally distributed in forward and backward propagating TE-light, so about 8.5% of the generated power couples into the forward propagating mode. As far as the pump efficiency is concerned, it is important to note that the pump power which is not used by the first antenna is still fully available for the next antennas in the array (apart from a small waveguide loss reduction). In that respect it is possible to transfer almost all the pump light to the antennas, which makes the overall energy transfer quite efficient. It should be noted however that part of the exchanged power is lost due to non-radiative heating and hence does not contribute to the optical excitation. By sweeping the geometrical parameters one could in principle find a configuration which optimizes the β -factor (at the Stokes wavelength) or the useful energy transfer (at the pump wavelength). However, since the overall Raman process is a

combination of two wavelength contributions, it is likely that a separate optimization of the structure for either the pump or Stokes light will not yield the optimal SERS signal. Hence it remains essential to optimize the suggested figure of merit FOM as this eventually determines the maximal SERS signal.

7.7.2 Fit of the experimental data

In Figure 7.13 we plot the experimentally obtained signal and shot noise counts of the 1340 cm^{-1} peak as a function of the number of antennas for a fixed bowtie geometry (the datapoints are extracted from Figure 7.5(b)). In the previous section we thoroughly outlined a procedure to estimate the spread on the experimental parameters E_P , E_S and η_A . This procedure will now be applied to the given experimental data.

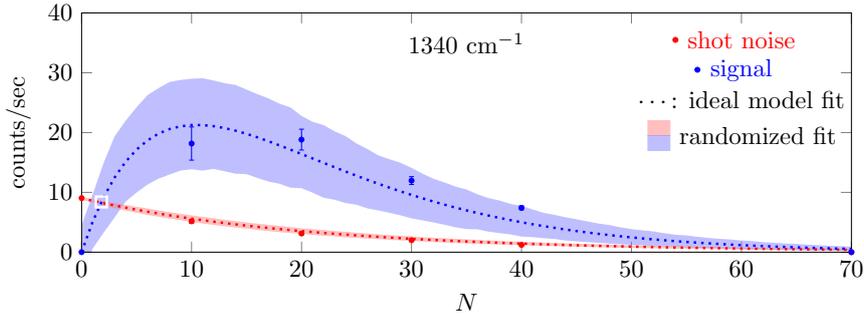


Figure 7.13: Signal (blue circles) and shot noise (red circles) at the 1340 cm^{-1} peak for the ($\alpha \approx 60^\circ$, $L \approx 106 \pm 8\text{ nm}$ and $\Delta \approx 48 \pm 13\text{ nm}$) bowtie. The dotted lines represent a constrained fit to the ideal model while the shaded areas represent a randomized fit to estimate the uncertainty on the experimental parameters. The white square denotes the minimum number of antennas N_{min} required to generate a detectable signal.

The dotted lines represent a constrained fit to the ideal model, from which the mean values of the E_P , E_S and η_A distributions were extracted (mind that η_A only differs from η_A^* by a constant factor C^*). From the randomized fit we then obtained $E_P \approx 0.49 \pm 0.15\text{ dB}$, $E_S \approx 0.35 \pm 0.11\text{ dB}$ and $\eta_A = (2.60 \pm 0.77) \times 10^{-15}$. Theoretically $E_P = 0.63\text{ dB}$, $E_S = 0.43\text{ dB}$ and $\eta_A = 2.19 \times 10^{-15}$ were predicted for this particular bowtie structure. The theoretically predicted parameters are all within the error bars of the experimentally fitted data, which clearly establishes the validity of our model and its ability to provide quantitative predictions

of the absolute Raman power coupled into a single mode waveguide. Given this excellent correspondence, we expect the fabricated structures to have a Raman enhancement factor EF_R on the order of 1.14×10^4 near the two antenna gap tips. Decreasing the gap size should boost the maximum EF_R by another two or three orders of magnitude. From the fitted conversion efficiency η_A one moreover derives that a single antenna produces (2.60 ± 0.77) fW of guided Stokes power for each 1W of guided pump power. Compared to the free space Raman scattering P_0 of a single NTP molecule in a bulk air environment, $P_A = \eta_A P_{pump}$ is $\approx 3.94 \times 10^6 P_0$. This includes the excitation and emission enhancement of all molecules in the monolayer as well as the coupling efficiency to the guided mode. Since only half of the Stokes light is carried by the forward propagating mode, the total power coupled into the fundamental TE-mode is therefore $\approx 8 \times 10^6 P_0$.

Our observations also reveal that a minimum number of antennas N_{min} is required to generate a detectable signal (marked by the white square in Figure 7.13). If $N < N_{min}$ then the shot noise still dominates on the signal. It has to be noted however that the relevant signal is generated in a very small region $(N - 1)\Lambda$ compared to the overall length $2L_1 + (N - 1)\Lambda \approx 2L_1$, while the shot noise is mainly attributed to this non-useful length $2L_1 \approx 10000 \mu m$. Practically the antenna section can be made as short as $\approx 10 \mu m$ for a single antenna. Consequently, designs in which dedicated on-chip splitters are used to separate the pump and Stokes light should in principle allow to reduce the background by a factor of 1000. These devices are therefore expected to have a $\sqrt{1000} \approx 32$ times higher SNR , pushing towards the $N_{min} = 1$ limit such that signals generated by a single antenna can still be detected. These experimental observations are all in line with the theoretically expected trends (see e.g. Figure 7.10). When the background is reduced by a dedicated chip design, one hence maintains the possibility of optimizing the antenna array in terms of absolute SERS enhancement. Since N_{opt} is completely determined by λ_P and λ_S , the highest absolute SERS signal strength will then be obtained for an array geometry for which $FOM(N_{opt}(\lambda_P, \lambda_S), \lambda_P, \lambda_S) \equiv FOM(\lambda_P, \lambda_S)$ is optimized for a given λ_P and λ_S . By numerical simulation one can show that for a fixed angle $\alpha \approx 60^\circ$ and gap $\Delta \approx 51$ nm, the optimal length $L \approx 110$ nm (for the given pump and NTP Stokes wavelength). The same procedure can be repeated for other gap sizes or other geometries. Hence, using our model it is possible to find an antenna array which maximizes the absolute on-chip SERS enhancement for a given pump and Stokes wavelength. The developed single mode SERS probe moreover allows

quantitative control on the Raman enhancement and subsequent coupling of the enhanced Stokes light with the underlying waveguide. In combination with other on-chip spectral functionalities, such as arrayed waveguide gratings [37], the presented SERS probe is forecasted to allow multiplexed detection of extremely weak Raman signals on a highly dense integrated platform. We also envisage that integrated nanoantennas, similar to the ones reported here, could be used as transducer between quantum dot emitters and the fundamental waveguide mode, potentially enabling applications in on-chip quantum communication and quantum computation. [38, 39]

7.7.3 Comparison of FOM_F

Now that we have introduced all relevant parameters, it is interesting to compare the single antenna conversion efficiency η_A and FOM_F for the three different antennas (single rod (SR), double rod (DR) and bowtie (BT)) introduced in the previous chapter. There it was found that the maximum Raman enhancement EF_R^{max} (for $\lambda_P = 785$ nm and $\lambda_S = 877$ nm) of a single rod antenna is about two orders of magnitude lower than EF_R^{max} of a bowtie antenna. Here we will consider the same antennas, but now coated with a 1 nm monolayer of NTP (the same layer as used in the previous simulations described above). As opposed to EF_R^{max} , this will quantify an average Raman enhancement because now an ensemble of points (and not just one point) is considered. For each of the antennas we calculated the extinction E , single antenna conversion efficiency η_A and FOM_F as shown in Figure 7.14. Please note that all extinction spectra are slightly redshifted when the antennas are coated with NTP. If we consider the maximum η_A for each of the antennas (at the selected λ_S marked by the cyan shaded line), then it is striking that despite the two orders of magnitude difference in maximum Raman enhancement, the maximum η_A of a SR only differs by a factor 9 from the maximum η_A of a BT antenna and by a factor 6 from the maximum DR η_A . The difference in maximum FOM_F is even smaller (only a factor of 6 difference between the SR and BT and a factor of 4 between the SR and DR). This discrepancy is ascribed to the fact that in case of a monolayer (or a collection of molecules/particles) near the antenna surface, one looks at an average effect of the Raman enhancement. Since all nanoplasmonic antennas redistribute the local fields (regardless of whether they are gapped or not), it is reasonable that the average Raman enhancement effect will not have as dramatic differences among several antenna types as compared to the maximum Raman enhancement. Since the fabrication of gapped nanostructures is however (much) harder than

the fabrication of isolated antennas, one could argue that the reduction in FOM_F comes at the benefit of easier (and potentially more robust) fabrication.

Moreover we analyzed the effect of the gap on η_A and FOM_F for a bowtie antenna with fixed apex angle $\alpha = 60^\circ$, length $L = 110$ nm and thickness $t_{Au} = 30$ nm. All antennas are again coated with an NTP monolayer of 1 nm. From Figure 7.15 one can see that for the selected Stokes wavelength $\lambda_S = 877$ nm, the η_A of a 10 nm gap antenna only differs by a factor of 3.5 from the η_A of a 50 nm gap antenna. Similarly, the FOM_F of a 10 nm antenna is only 3.7 times higher than the FOM_F of a 50 nm antenna. The maximum Raman enhancement for the $\lambda_S = 877$ nm line of a 10 nm antenna is $\approx 10^6$ while it only amounts $\approx 2 \times 10^4$ for a 50 nm gap antenna. Despite this two orders of magnitude difference, one does not gain much in averaged SERS enhancement (related to FOM_F) when going to smaller gaps. On the other hand, it becomes increasingly difficult to make small gaps with high reproducibility. In that respect one might want to resort to larger gaps at the cost of a slightly lower FOM_F . The above results are quite striking in the sense that not much can be gained with gapped nanostructures when one considers monolayer coated antennas which are integrated on a waveguide.

Finally, we also studied the impact of the adhesion layer on η_A and FOM_F for a bowtie antenna with fixed apex angle $\alpha = 60^\circ$, length $L = 106$ nm and gap $\Delta = 48$ nm. Three different situations were considered: no adhesion layer between the Au and the waveguide, a 2 nm thick Ti adhesion layer and a 2 nm thick MPTS ((3-Mercaptopropyl)trimethoxysilane, [40]) adhesion layer. The MPTS layer is modeled using a refractive index of 1.444. [40] For each of these three cases, E , η_A and FOM_F are depicted in Figure 7.16. For the selected λ_S , the η_A of an antenna with MPTS adhesion layer is 3.45 higher as compared to the 2 nm Ti adhesion layer case (FOM_F is 3.76 times higher). Compared to no adhesion layer at all, the MPTS case has a 1.3 times higher η_A and 1.65 times higher FOM_F . By changing the 2 nm Ti adhesion layer by a 2 nm MPTS adhesion layer one can hence increase the maximum SERS signal by a factor 3.76.

7.8 ROC versus SEROC

In this section we will compare the relative strength of spontaneous Raman On-Chip (ROC) and Surface Enhanced Raman On-Chip (SEROC). To this end, we consider the bulk Raman sensing experiment with IPA as described

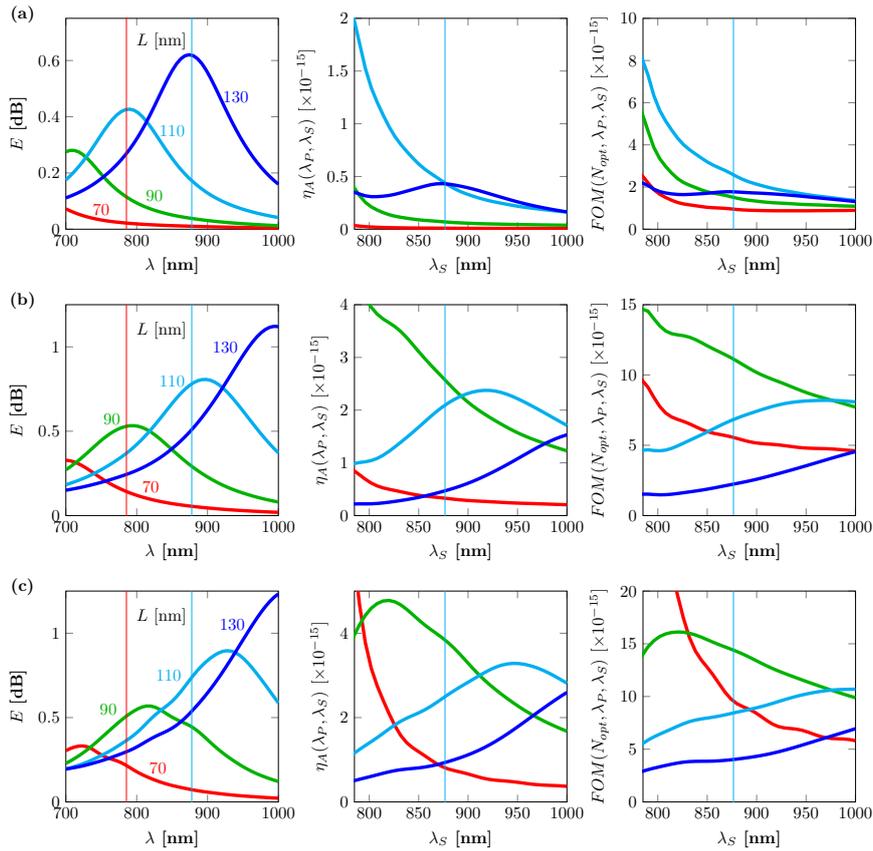


Figure 7.14: Comparison of the extinction E (left), single antenna conversion efficiency η_A (middle) and FOM_F (right) for single rod (a), double rod (b) and bowtie (c) antennas. For each of the antennas, four different lengths L (70, 90, 110, 130 nm) were considered, while the other parameters were kept fixed. The Au thickness is $t_{Au} = 30$ nm, the single and double rod width $W = 55$ nm, the bowtie apex angle $\alpha = 60^\circ$ and the double rod and bowtie gap $\Delta = 10$ nm. Each antenna is also covered with a 1 nm NTP monolayer. The red and cyan line again mark the pump $\lambda_P = 785$ nm and selected Stokes $\lambda_S = 877$ nm wavelength.

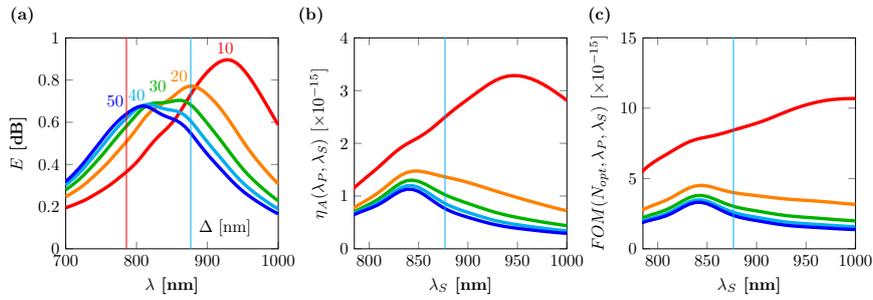


Figure 7.15: Comparison of the extinction E (a), single antenna conversion efficiency η_A (b) and FOM_F (c) for a bowtie antenna with fixed apex angle $\alpha = 60^\circ$, length $L = 110$ nm and thickness $t_{Au} = 30$ nm, but varying gap Δ . The red and cyan line again mark the pump $\lambda_P = 785$ nm and selected Stokes $\lambda_S = 877$ nm wavelength.

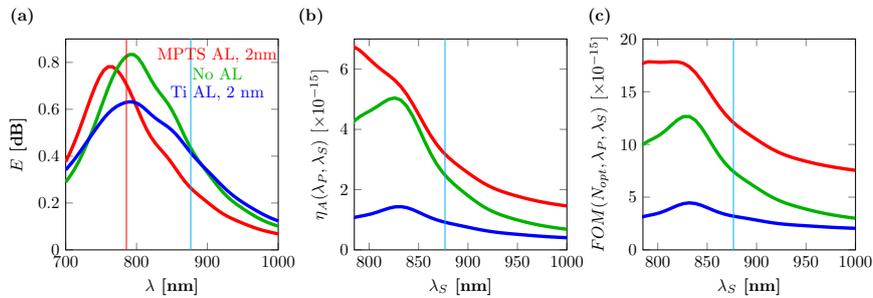


Figure 7.16: Comparison of the extinction E (a), single antenna conversion efficiency η_A (b) and FOM_F (c) for a bowtie antenna with fixed apex angle $\alpha = 60^\circ$, length $L = 106$ nm and gap $\Delta = 48$ nm, but varying adhesion layer (AL). The red and cyan line again mark the pump $\lambda_P = 785$ nm and selected Stokes $\lambda_S = 877$ nm wavelength. Only the gold area is coated with an NTP monolayer.

in an earlier chapter. There it was shown that the ratio of the collected spontaneous Stokes power to the transmitted power equals

$$\frac{P_{coll}}{P_t} = \zeta(L) \approx \frac{\rho\sigma\eta_0 L}{2}, \quad (7.41)$$

in case the complete top cladding of a waveguide with length L is covered with IPA (density $\rho = 7.9 \times 10^{27} \text{ m}^{-3}$). The conversion efficiency $\rho\sigma\eta_0 = 4 \times 10^{-10} \text{ cm}^{-1}$ for the 819 cm^{-1} IPA peak, which has a cross-section $\sigma = 7.9 \times 10^{-31} \text{ cm}^2 \text{ sr}^{-1}$. [13] In order to compare the strength of this spontaneous Raman signal with a SERS signal, we have ‘coated’ a bowtie antenna with a uniform IPA layer of thickness t_{IPA} and calculated the single antenna conversion efficiency η_A for the 819 cm^{-1} IPA peak. The calculation procedure is exactly the same as the one for an NTP monolayer. In Figure 7.17(a) η_A is plotted as a function of t_{IPA} . As could be expected, the signal increases when the probing volume ($\propto t_{IPA}$) increases. The local peak at $t_{IPA} = 10 \text{ nm}$ could be explained by a better overlap between the analyte and the hotspot volume.

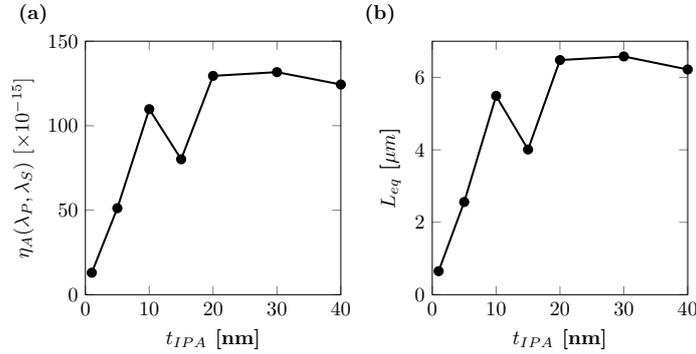


Figure 7.17: Single antenna conversion efficiency η_A (a) and equivalent length L_{eq} (b) for a bowtie antenna with fixed apex angle $\alpha = 60^\circ$, length $L = 106 \text{ nm}$ and gap $\Delta = 48 \text{ nm}$, covered by a layer of IPA with varying thickness t_{IPA} .

It is now instructive to calculate which equivalent length L_{eq} would be required to generate the same signal as one antenna. This length can easily be found by

$$L_{eq} = \frac{2\eta_A}{\rho\sigma\eta_0} \quad (7.42)$$

and is plotted in Figure 7.17(b), taking into account the conversion efficiency $\rho\sigma\eta_0 = 4 \times 10^{-10} \text{ cm}^{-1}$. One can see that an antenna covered

with an IPA layer of 1 – 20 nm generates a signal equivalent to the signal that would be generated by a $\approx 1 - 7 \mu\text{m}$ SiN waveguide which is totally covered by IPA. This means that as soon as a bulk volume of IPA covers a length of $\approx 10 - 70 \mu\text{m}$, the SERS signal generated by one antenna will already be one order of magnitude lower compared to the spontaneously generated Raman signal. As a result, the integration of nanoplasmonic antennas on top of waveguides has no added value for bulk Raman sensing (so for a sufficiently large probing volume) since the antenna effect is too localized to compete with the much larger spontaneous Raman scattering. As a rough rule of thumb, one could state that one antenna corresponds with a $1 - 10 \mu\text{m}$ waveguide section when it concerns the generation of Raman signals. Nevertheless, this on-chip spontaneous Raman platform still suffers from the same drawback of conventional Raman spectroscopy in the sense that spontaneous Raman scattering is usually extremely weak. It hence becomes increasingly difficult to detect very small amounts of analyte or single nanoparticles (or molecules) through spontaneous Raman spectroscopy. At this point, SEROC can boost the signal provided that the analyte or nanoparticle finds its way to the hotspot. In such a case the problem becomes diffusion limited, but one could envisage schemes where small microfluidic channels are patterned on top of the antenna region in order to improve the interaction probability between the particles and the antenna. Analytes can also adsorb to the metal surface, in which case one could detect very small amounts of these analytes or chemical changes due to interaction with other molecules.

7.9 Comparison with other SERS substrates

Apart from a comparison between ROC and SEROC, it is also instructive to include a comparative analysis between SEROC and other SERS substrates. A commonly used figure of merit in comparing the relative performance of SERS substrates is the substrate enhancement factor SEF :

$$SEF = \frac{I_{SERS}/N_{abs}}{I_{bulk}/N_{bulk}} \quad (7.43)$$

where I_{SERS} and I_{bulk} is the Raman intensity obtained in a SERS and spontaneous Raman experiment respectively (taken under identical conditions, i.e. same excitation wavelength and power). N_{bulk} is the average number of molecules in the scattering volume V_s for a spontaneous Raman measurement. Typically V_s is taken to be the confocal volume when

the spectrum is collected through a confocal Raman microscope. N_{abs} is the number of adsorbed molecules on the SERS substrate. [41] In [11] a new type of SERS substrate, based on wafer-level arrays of Au-capped nanopillars, is compared with a commercially available Klarite substrate. Below we will provide a comparative analysis of our on-chip SERS probe with these two SERS substrates.

We will first derive an analytical *SEF* formula for our on-chip SERS probes. In case of forward scattering, the maximum generated SERS intensity is proportional to $FOM_F(N_{opt})e^{-2\alpha_{wg}L_1}$, while a spontaneous Raman signal collected through a confocal Raman microscope is proportional to $2\rho\sigma(\lambda_P/n)$ (see Chapter 3). The molecular density ρ can be written in terms of the number of bulk molecules and the scattering volume V_s as $\rho = N_{bulk}/V_s$ such that

$$SEF = \frac{nFOM_F(N_{opt})e^{-2\alpha_{wg}L_1}}{2\frac{N_{bulk}}{V_s}\sigma\lambda_P} \frac{N_{bulk}}{N_{abs}} \quad (7.44)$$

$$= \frac{nV_sFOM_F(N_{opt})e^{-2\alpha_{wg}L_1}}{2\sigma N_{abs}\lambda_P} \quad (7.45)$$

In the experiments discussed before, the following numbers represent realistic values for the parameters in the *SEF* formula: $n = 1.33$, $FOM_F(N_{opt} \approx 9) \approx 10^{-14}$, $e^{-2\alpha_{wg}L_1} \approx 0.5$, $\sigma \approx 0.358 \times 10^{-30} \text{ cm}^2$, $\lambda_P = 785 \text{ nm}$. The scattering volume V_s is estimated to be approximately $10 \mu\text{m}^3$. The *SEF* for our on-chip SERS probe is then

$$SEF \approx \frac{10^9}{N_{abs}}. \quad (7.46)$$

It is important to note that this *SEF* is generated by a very small number of antennas (i.e. $N_{opt} = 9$). The *SEF* of the substrates that we will discuss next for comparison are all calculated using a large number of active SERS hotspots (i.e. $\gg 9$). A more fair comparison therefore reports the SEF_1 of single SERS hotspots as well. To this end, we will divide the overall *SEF* by the number of estimated SERS hotspots. In our case, a single antenna (hotspot) has $FOM_F \approx 2 \times 10^{-15}$, such that the single antenna (hotspot) SEF_1 is

$$SEF_1 \approx \frac{10^9}{5N_{abs}}. \quad (7.47)$$

The estimated surface area of a typical bowtie antenna is $32200 \times 10^{-18} \text{ m}^2$ and the surface density of NTP molecules is $\rho_s = 5.56 \times 10^{18} \text{ molecules/m}^2$

such that $N_{abs} \approx 179000 \approx 2 \times 10^5$. For full monolayer coverage, our on-chip SERS probe hence has $SEF \approx 5000$ and $SEF_1 = 1000$. Because the tightly confined hotspots are mainly generated around the tips it is reasonable to assume that the dominant part of the SERS signal is generated by a much lower number of molecules N_{mol}^{hot} . We estimate a combined hot spot surface area of approximately $1200 \times 10^{-18} \text{ m}^2$ in which case $N_{abs} = N_{mol}^{hot} \approx 7000$ and $SEF^{hot} \approx 149250$ and $SEF_1^{hot} \approx 30000$.

A similar analysis to determine N_{mol}^{hot} , SEF^{hot} and SEF_1^{hot} for the Au-cap nanopillar and Klarite substrates will be described below. In [11] the available surface area of one Au-cap is estimated to be $2.2 \times 10^4 \text{ nm}^2$, such that about 12×10^4 molecules can bind to one Au-cap (assuming full monolayer coverage). Based on the field plots in [11] we roughly estimate that between 25 to 50 % of the molecules are within the hotspot. Taking 25%, $N_{mol}^{hot} = 3 \times 10^4$ for the nanopillar. For the Klarite substrate, the hot spot area is estimated to be $1.25 \mu\text{m}^2$ (again assuming that 25% of the full monolayer is within a hotspot). This amounts to $N_{mol}^{hot} = 6 \times 10^6$ hotspot molecules for the Klarite substrate. SEF^{hot} is now obtained by multiplying SEF by a factor 4, because only 25% of the full monolayer is assumed to contribute significantly to the overall SERS signal.

The nanopillar density is estimated to be $1.23 \times 10^{14} \text{ pillars/m}^2$, while the laser focal spot area is about $19.6 \times 10^{-12} \text{ m}^2$ such that the SERS signal is generated by ≈ 2400 nanopillars (hotspots). In the same focal spot, there are about 5 active Klarite hotspots. SEF_1^{hot} is obtained by dividing SEF^{hot} by the number of active hotspots in the focal spot of the laser (so by 2400 for the nanopillar substrate and by 5 for the Klarite substrate).

Finally, it is important to consider that the number of molecules in each of the hotspot areas is quite different. For our SEROC substrate there are only 7000 hotspot molecules, while there are 6×10^6 hotspot molecules for the Klarite substrate. Therefore, one should also consider the relative strength of each of these substrates for the same amount of hotspot molecules. To this end, we define $SEF_1^{hot,eq} = SEF_1^{hot}/N_{mol}^{hot}$, meaning that we only consider 1 molecule in the hotspot area.

The results are summarized in the table below:

Substrate	SEF	N_{mol}^{hot}	SEF^{hot}	SEF_1^{hot}	$SEF_1^{hot,eq}$
Klarite	8.4×10^4	6×10^6	3.4×10^5	6.8×10^4	0.01
Nanopillar	2.3×10^5	3×10^4	1×10^6	4×10^2	0.1
SEROC	5×10^3	7×10^3	1.5×10^5	3×10^4	4

One can see that the pure SEF of our SEROC substrates is much lower as compared to the nanopillars or Klarite substrates. However, in case of

our bowtie antennas the dominant hotspots are mainly situated near the antenna tips such, making the SEF^{hot} a more fair number in assessing the enhancement capabilities of SEROC. Comparing SEF^{hot} shows that all substrates have similar enhancement properties, with the nanopillar still being the best. Optimized SEROC designs (i.e. lowering the waveguide loss and improving FOM) should boost the SEROC SEF^{hot} by at least one order of magnitude, making it competitive with the existing substrates. By comparing SEF_1^{hot} one can see that a large part of the strength of the Au-cap nanopillars can be attributed to the high density of SERS hotspots, because SEF_1^{hot} is considerably lower for the nanopillars as compared to Klarite or SEROC substrates. The true strength of SEROC becomes however visible when comparing $SEF_1^{hot,eq}$, which could be considered as a comparison of the relative strength of the hotspot areas. SEROC substrates are expected to have at least a one order of magnitude higher hotspot strength as compared to the nanopillars and at least two orders of magnitude higher hotspot strength as compared to Klarite substrates. So in case only a very small amount of target adsorbate molecules is present, SEROC will show a (much) higher sensitivity as compared to the other substrates (for the same excitation power and wavelength). In other words, if a Klarite substrate hits its detection limit for a given concentration of adsorbate molecules, a SEROC substrate will be able to detect a 100 times lower concentration of adsorbate molecules. Currently, the losses due to in- and outcoupling limit the SEROC sensitivity but further integration of on-chip laser sources and spectrometers should (partially) mitigate this issue.

7.10 Conclusion

In this chapter we outlined the main result of this PhD thesis, i.e. the generation of SERS signals from integrated bowtie nanoantennas, excited and collected by a single mode silicon nitride waveguide. We have investigated the influence of the SERS signal on the plasmon resonance and the number of antennas on the waveguide. Moreover we established an analytical model that allows quantitative predictions of the Stokes power coupled into the fundamental TE-mode and exhibits an excellent correspondence with the experimentally observed trends. Furthermore we discussed the trade-off between signal optimization and noise reduction and studied the on-chip SERS performance of three common antenna types. Finally, the spontaneous on-chip Raman platform was compared with the on-chip SERS platform. In the next chapter we will have a closer look at

the background reduction problem in order to improve the *SNR* of these on-chip SERS probes.

References

- [1] Anker, J. N.; Hall, W. P.; Lyandres, O.; Shah, N. C.; Zhao, J.; Van Duyne, R.P. Biosensing with plasmonic nanosensors. *Nat. Mater.* **2008**, *7*, 442–453.
- [2] Willets, K. A.; Van Duyne, R.P. Localized surface plasmon resonance spectroscopy and sensing. *Annu. Rev. Phys. Chem.* **2007**, *58*, 267–297.
- [3] Halas, N. J.; Lal, S.; Chang, W-S.; Link, S.; Nordlander, P. Plasmons in strongly coupled metallic nanostructures. *Chem. Rev.* **2011**, *111*, 3913–3961.
- [4] Giannini, V.; Fernández-Domínguez, A. I.; Heck, S. C.; Maier, S. A. Plasmonic nanoantennas: fundamentals and their use in controlling the radiative properties of nanoemitters. *Chem. Rev.* **2011**, *111*, 3888–3912.
- [5] Chu, Y.; Banaee, M. G.; Crozier, K. B. Double-resonance plasmon substrates for surface-enhanced Raman scattering with enhancement at excitation and Stokes frequencies. *ACS Nano* **2010**, *4*, 2804–2810.
- [6] McFarland, A. D.; Young, M. A.; Dieringer, J. A.; Van Duyne, R. P. Wavelength-scanned surface-enhanced Raman excitation spectroscopy. *J. Phys. Chem. B* **2005**, *109*, 11279–11285.
- [7] Ye, J.; Wen, F.; Sobhani, H.; Britt Lassiter, J.; Van Dorpe, P.; Nordlander, P.; Halas, N.J. Plasmonic Nanoclusters: Near Field Properties of the Fano Resonance Interrogated with SERS. *Nano Lett.* **2012**, *12*, 1660–1667.
- [8] Kasera, S.; Biedermann, F.; Baumberg, J.J.; Scherman, O.A.; Mahajan, S. Quantitative SERS Using the Sequestration of Small Molecules Inside Precise Plasmonic Nanoconstructs. *Nano Lett.* **2012**, *12*, 5924–5928.
- [9] Gallinet, B.; Siegfried, T.; Sigg, H.; Nordlander, P.; Martin, O.J.F. Plasmonic Radiance: Probing Structure at the Angström Scale with Visible Light. *Nano Lett.* **2013**, *13*, 497–503.

- [10] Siegfried, T.; Ekinici, Y.; Martin, O.J.F.; Sigg, H. Gap Plasmons and Near-Field Enhancement in Closely Packed Sub- 10 nm Gap Resonators. *Nano Lett.* **2013**, *13*, 5449–5453.
- [11] Li, J.; Chen, C.; Jans, H.; Xiumei, X.; Verellen, N.; Vos, I.; Okumura, Y.; Moshchalkov, V.V.; Lagae, L.; Van Dorpe, P. 300 mm Wafer-level, ultra-dense arrays of Au-capped nanopillars with sub-10 nm gaps as reliable SERS substrates. *Nanoscale* **2014**, *6*, 12391–12396.
- [12] Seok, T. J.; Jamshidi, A.; Eggleston, M.; Wu, M.C. Mass-producible and efficient optical antennas with CMOS-fabricated nanometer-scale gap. *Opt. Express* **2013**, *21*, 16561–16569.
- [13] Dhakal, A.; Subramanian, A.Z.; Wuytens, P.; Peyskens, F.; Le Thomas, N.; Baets, R. Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides. *Opt. Lett.* **2014**, *39*, 4025–4028.
- [14] Peyskens, F.; Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Bright and dark plasmon resonances of nanoplasmonic antennas evanescently coupled with a silicon nitride waveguide. *Opt. Express* **2015**, *23*, 3088–3101.
- [15] Peyskens, F.; Subramanian, A.Z.; Dhakal, A.; Le Thomas, N.; Baets, R. In *Enhancement of Raman Scattering Efficiency by a Metallic Nano-antenna on Top of a High Index Contrast Waveguide*, CLEO 2013, San Jose, California, United States, June 09-14, 2013.
- [16] Arnaud, L.; Bruyant, A.; Renault, M.; Hadjar, Y.; Salas-Montiel, R.; Apuzzo, A.; Léron del, G.; Morand, A.; Benech, P.; Le Coarer, E.; Blaize, S. Waveguide-coupled nanowire as an optical antenna. *J. Opt. Soc. Am. A* **2013**, *30*, 2347–2355.
- [17] Arango, F. B.; Kwadrin, A.; Koenderink, A. F. Plasmonic Antennas Hybridized with Dielectric Waveguides. *ACS Nano* **2012**, *6*, 10156–10167.
- [18] Février, M.; Gogol, P.; Aassime, A.; Mégy, R.; Delacour, C.; Chelnokov, A.; Apuzzo, A.; Blaize, S.; Lourtioz, J-M.; Dagens, B. Giant Coupling Effect between Metal Nanoparticle Chain and Optical Waveguide. *Nano Lett.* **2012**, *12*, 1032–1037.

- [19] Février, M.; Gogol, P.; Lourtioz, J-M.; Dagens, B. Metallic nanoparticle chains on dielectric waveguides: coupled and uncoupled situations compared. *Opt. Express* **2013**, *21*, 24504–24513.
- [20] Février, M.; Gogol, P.; Barbillon, G.; Aassime, A.; Mégy, R.; Bartenlian, B.; Lourtioz, J-M.; Dagens, B. Integration of short gold nanoparticles chain on SOI waveguide toward compact integrated biosensors. *Opt. Express* **2012**, *20*, 17402–17410.
- [21] Chamanzar, M.; Xia, Z.; Yegnanarayanan, S.; Adibi, A. Hybrid integrated plasmonic-photonic waveguides for on-chip localized surface plasmon resonance (LSPR) sensing and spectroscopy. *Opt. Express* **2013**, *21*, 32086–32098.
- [22] Lin, S.; Zhu, W.; Jin, Y.; Crozier, K.B. Surface-Enhanced Raman Scattering with Ag Nanoparticles Optically Trapped by a Photonic Crystal Cavity. *Nano Lett.* **2013**, *13*, 559-563.
- [23] Kong, L.; Lee, C.; Earhart, C.M.; Cordovez, B.; Chan, J.W. A nanotweezer system for evanescent wave excited surface enhanced Raman spectroscopy (SERS) of single nanoparticles. *Opt. Express* **2015**, *23*, 6793-6802.
- [24] Measor, P.; Seballos, L.; Yin, D.; Zhang, J.Z.; Lunt, E.J.; Hawkins, A.R.; Schmidt, H. On-chip surface-enhanced Raman scattering detection using integrated liquid-core waveguides. *Appl. Phys. Lett.* **2007**, *90*, 211107-1-211107-3.
- [25] Fan, M.; Andrade, G.F.S.; Brolo, A.G. A review on the fabrication of substrates for surface enhanced Raman spectroscopy and their applications in analytical chemistry. *Anal. Chim. Acta* **2011**, *693*, 7–25.
- [26] Subramanian, A.Z.; Neutens, P.; Dhakal, A.; Jansen, R.; Claes, T.; Rottenberg, X.; Peyskens, F.; Selvaraja, S.; Helin, P.; Du Bois, B.; Leyssens, K.; Severi, S.; Deshpande, P.; Baets, R.; Van Dorpe, P. Low-Loss Singlemode PECVD Silicon Nitride Photonic Wire Waveguides for 532900 nm Wavelength Window Fabricated Within a CMOS Pilot Line. *IEEE Photon. J.* **2013**, *5*(6), 2202809.
- [27] Romero-García, S.; Merget, F.; Zhong, F.; Finkelstein, H.; Witzens, J. Silicon nitride CMOS-compatible platform for integrated photonics

- applications at visible wavelengths. *Opt. Express* **2013**, *21*(12), 14036–14046.
- [28] Peyskens, F.; Dhakal, A.; Van Dorpe, P.; Le Thomas, N.; Baets, R. Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform. *ACS Photonics* **2016**, *3*(1), 102–108.
- [29] Mahmoud, M.A. Surface-Enhanced Raman Spectroscopy of Double-Shell Hollow Nanoparticles: Electromagnetic and Chemical Enhancements. *Langmuir* **2013**, *29*, 6253–6261.
- [30] Chul Jun, Y.; Briggs, R.M.; Atwater, H.A.; Brongersma, M.L. Broadband enhancement of light emission in silicon slot waveguides. *Opt. Express* **2009**, *17*, 7479–7490.
- [31] Long, D.A. *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*; John Wiley & Sons Ltd: Chichester, England, 2002; pp 96-97.
- [32] CCD Signal to Noise Ratio. <http://www.andor.com/learning-academy/ccd-signal-to-noise-ratio-calculating-the-snr-of-a-ccd> (accessed Jan 02, 2016).
- [33] CCD Sensitivity and Noise. <http://www.andor.com/learning-academy/ccd-sensitivity-and-noise-what-is-noise-and-what-does-it-matter> (accessed Jan 02, 2016).
- [34] Johnson, P.B.; Christy, R.W. Optical Constants of the Noble Metals. *Phys. Rev. B* **1972**, *6*, 4370–4379.
- [35] Haynes, W. M., Eds. *CRC Handbook of Chemistry and Physics 95th Edition (Internet Version 2015)*; CRC/Taylor and Francis, 2015.
- [36] Thomas, M.; Mühlig, S.; Deckert-Gaudig, T.; Rockstuhl, C.; Deckert, V.; Marquetand, P. Distinguishing chemical and electromagnetic enhancement in surface-enhanced Raman spectra: The case of para-nitrothiophenol. *J. Raman Spectrosc.* **2013**, *44*, 1497-1505.
- [37] Martens, D.; Subramanian, A.; Pathak, S.; Vanslebrouck, M.; Bienstman, P.; Bogaerts, W.; Baets, R. Compact silicon nitride

- arrayed waveguide gratings for very near-infrared wavelengths. *IEEE Photon. Technol. Lett.* **2015**, *27*, 137-140.
- [38] Pelton, M. Modified spontaneous emission in nanophotonic structures. *Nature Photon.* **2015**, *9*, 427-435.
- [39] Tame, M.S.; McEnery, K.R.; Özdemir, S.K.; Lee, J.; Maier, S.A.; Kim, M.S. Quantum plasmonics. *Nature Phys.* **2013**, *9*, 329-340.
- [40] (3-Mercaptopropyl)trimethoxysilane. <http://www.sigmaaldrich.com/catalog/product/aldrich/175617?lang=en®ion=BE> (accessed Jan 13, 2016).
- [41] Le Ru, E.C.; Etchegoin, P.G. Quantifying SERS enhancements. *MRS Bulletin* **2013**, *38*, 631-640.

8

Background Mitigation

8.1 Introduction

In the previous chapter we developed an analytical model for on-chip SERS and showed that the inherent SiN background sets an ultimate limit on the detection performance. Dedicated on-chip designs, where the pump and Stokes beam can be separated, are expected to allow a substantial reduction of the shot noise and hence improve on the SERS Signal-to-Noise ratio. Here we will discuss such a design, based on a Y-splitter, and report on preliminary experimental results. It should be noted however that this research is still ongoing, hence no hard experimental conclusions can be drawn yet.

8.2 Y-splitter design

In Figure 8.1 we depict the Y-splitter design used to separate the pump and Stokes light. The upper arm is used to inject a pump beam with power P_{pump} at wavelength λ_P (red) into the chip. The straight arm of the Y-splitter contains the antenna array (see enlarged inset). Each antenna generates a certain amount of Stokes power at wavelength λ_S that will propagate both forwards and backwards (green lobes around each antenna). At the same time the pump beam will generate some background light

while propagating along the straight antenna arm. Both the SERS signal and the background light will then couple into the upper and lower arm of the Y-splitter. Moreover, the forward propagating pump and Stokes light can reflect on the gold array or on the end facet. On one hand this will give an additional contribution to the SERS light propagating backwards in the lower arm. On the other hand however, the pump reflection will also result in additional background. The ultimate goal is to minimize any pump and background reflection and optimize the backward propagating SERS light such that the lower antenna arm mainly contains the relevant SERS signal.

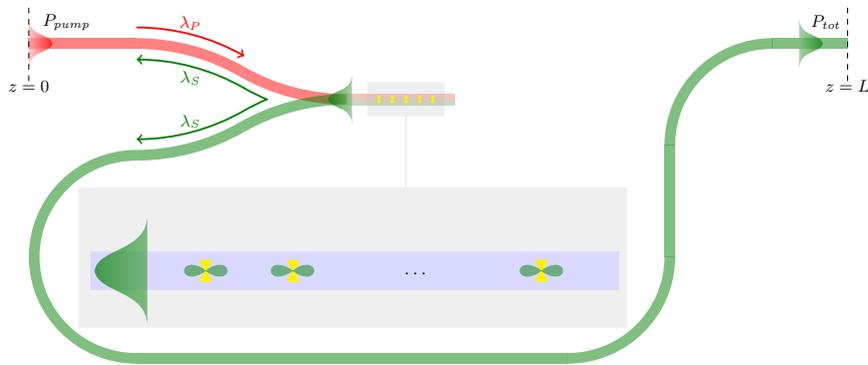


Figure 8.1: Y-splitter design to separate the pump and Stokes light.

As far as the SERS light is concerned, we have three main contributions:

- S1: The backward scattered Stokes light generated from the antenna array.
- S2: The forward scattered Stokes light from the antenna array, which is reflected at the end facet of the straight antenna arm.
- S3: SERS light generated by the reflected pump at the end facet of the straight antenna arm.

The background on the other hand is mainly generated by the following contributions:

- B1: The backward scattered background light originating from the straight antenna arm.

- B2: The forward scattered background light from the upper and straight antenna arm which is reflected at the end facet of the straight antenna arm.
- B3: The forward scattered background light from the upper and straight antenna arm which is reflected by the Au array.
- B4: Background light generated by the reflected pump at the end facet of the straight antenna arm.
- B5: Background light generated by pump light which is reflected by the Au array.

Reflections originating from the Y-splitter junction are neglected in this analysis. Based on these contributions we will derive an analytical formula for the expected SERS and background signal of this Y-splitter design.

The geometrical parameters of the aforementioned Y-splitter are depicted in Figure 8.2. The transmission of the Y-splitter from the upper arm to the straight arm is T_1 and from the straight to the lower arm T_2 . The transmission of the bends with radius R is T_B . At the end facet of the straight antenna arm, light can reflect with a reflection coefficient given by $R(\lambda)$. A single antenna is assumed to have a symmetrical radiation profile, emitting as much in the forward and backward direction. This radiated power, into the guided TE mode, is calculated through the single antenna conversion efficiency $\eta_A(\lambda_P, \lambda_S) = \eta_A$ and is given by

$$P_A = \eta_A P_{pump}. \quad (8.1)$$

In our analysis we assume that the waveguide loss at the pump and Stokes wavelength is more or less the same, i.e. $\alpha_P \approx \alpha_S = \alpha_{wg}$ and that the single antenna extinction at the pump and Stokes wavelength is e_P and e_S respectively. The straight antenna arm has a total length $L_2 + (N-1)\Lambda + L_3$ where Λ is the period between each of the antennas. This length is assumed to be much smaller than the length of the input and output waveguide ports $L_1 + 2\pi R + L_4 + L_V + L_5$. For simplicity we will hence take the limit $\alpha_{wg}L_2 = \alpha_{wg}L_3 = \alpha_{wg}\Lambda \rightarrow 0$ since the inherent waveguide losses over such a small length can in general be neglected. Apart from the reflection at the end facet, the Au array will also exhibit a certain reflection coefficient $R_A(N, \lambda)$. It should be noted that this reflection will depend on the number of antennas N in the array and the spacing inbetween them. Finally we define the quantity

$$\xi = T_1 T_2 T_B^4 e^{-\alpha_{wg}(L_1 + 2L_2 + 2L_3 + 2\pi R + L_4 + L_V + L_5)}, \quad (8.2)$$

which contains several loss factors due to the SiN circuit, but not due to the gold. When R is sufficiently large, the bend loss can be neglected such that $T_B \approx 1$.

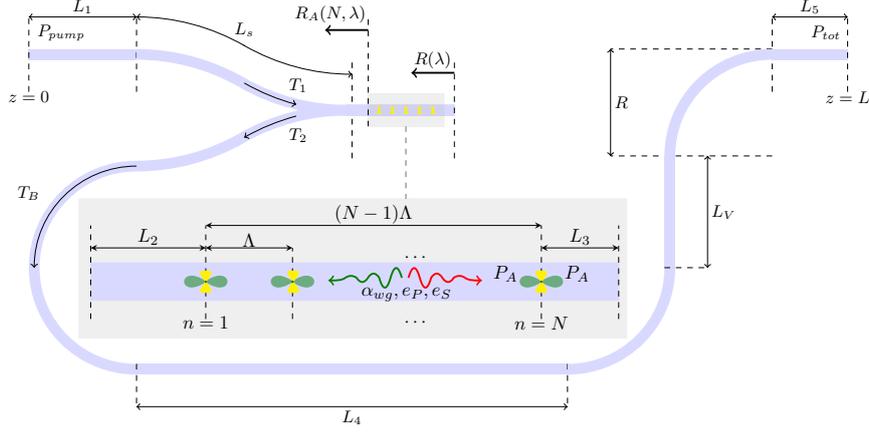


Figure 8.2: Geometrical parameters of the Y-splitter design.

8.2.1 Signal

Now that we have introduced all relevant parameters, we will calculate the separate contributions to the SERS signal. The pump power available at the first antenna is

$$P_{pump}^1 = P_{pump} e^{-\alpha_{wg}(L_1+L_s+L_2)} T_1. \quad (8.3)$$

The power generated by antenna n ($n = 1 \dots N$) is then

$$P_n = P_{pump}^1 \eta_A e^{-\alpha_{wg}(n-1)\Lambda} e_P^{1-n} \quad (8.4)$$

and the eventual backscattered Stokes power (S1 contribution) which reaches the output is

$$\begin{aligned} & \frac{P_n^{S1}}{P_{pump}^1} \\ &= \eta_A e^{-\alpha_{wg}(n-1)\Lambda} e_P^{1-n} e^{-\alpha_{wg}((n-1)\Lambda+L_2+L_s+2\pi R+L_4+L_V+L_5)} e_S^{1-n} T_2 T_B^4 \end{aligned}$$

which eventually equals

$$\begin{aligned} \frac{P_n^{S1}}{P_{pump}} &\approx \eta_A (e_P e_S)^{1-n} (T_1 T_2 T_B^4) e^{-\alpha_{wg}(L_1+2L_2+2L_s+2\pi R+L_4+L_V+L_5)} \\ &= \eta_A \xi (e_P e_S)^{1-n}. \end{aligned} \quad (8.5)$$

For the S2 contribution we consider the forward propagating power which is reflected with strength $R(\lambda_S)$. The power P_n^R , generated by antenna n , just after reflection at the end facet is then

$$\frac{P_n^R}{P_{pump}^1} = \eta_A e^{-\alpha_{wg}(n-1)\Lambda} e_P^{1-n} e^{-\alpha_{wg}((N-n)\Lambda+L_3)} e_S^{n-N} R(\lambda_S). \quad (8.6)$$

Taking into account the dielectric waveguides losses and the additional attenuation by propagating the Stokes light across N antennas, the power P_n^{S2} originating from reflected forward scattering that reaches the output is

$$\begin{aligned} \frac{P_n^{S2}}{P_{pump}} &= \eta_A \xi e^{-2\alpha_{wg}L_3} e_P^{1-n} e_S^{n-2N} R(\lambda_S) \\ &\approx \eta_A \xi e_P^{1-n} e_S^{n-2N} R(\lambda_S). \end{aligned} \quad (8.7)$$

Finally, the reflected pump light will again generate Stokes light in the backward direction (S3 contribution). The pump power at antenna N after reflection of the pump beam is

$$P_{pump}^N = P_{pump}^1 e^{-\alpha_{wg}(N-1)\Lambda} e_P^{-N} e^{-\alpha_{wg}L_3} R(\lambda_P) e^{-\alpha_{wg}L_3}. \quad (8.8)$$

The power P_n^{S3} which eventually reaches the output can be calculated similar to the S1 and S2 calculations and is given by

$$\frac{P_n^{S3}}{P_{pump}} = \eta_A \xi e_P^{n-2N} e_S^{1-n} R(\lambda_P). \quad (8.9)$$

Since N antennas contribute to the signal, the total SERS power P_S at the output is then

$$\begin{aligned} \frac{P_S}{P_{pump}} &= \eta_A \xi \sum_{n=1}^N (P_n^{S1} + P_n^{S2} + P_n^{S3}) \\ &= \eta_A \xi \left(\left(\frac{1 - \left(\frac{1}{e_P e_S}\right)^N}{1 - \left(\frac{1}{e_P e_S}\right)} \right) + R(\lambda_S) e_S^{1-2N} \left(\frac{1 - \left(\frac{e_S}{e_P}\right)^N}{1 - \left(\frac{e_S}{e_P}\right)} \right) \right. \\ &\quad \left. + R(\lambda_P) e_P^{1-2N} \left(\frac{1 - \left(\frac{e_P}{e_S}\right)^N}{1 - \left(\frac{e_P}{e_S}\right)} \right) \right). \end{aligned} \quad (8.10)$$

In order to get a better insight into the relative strength of the different contributions, it is important to have an approximate formula for the reflection coefficients. The reflection $R(\lambda)$ at the end facet can be considered as a reflection on the interface between two media with respective indices $n_{eff}(\lambda)$, i.e. the effective index of the waveguide mode, and n_{surr} , i.e. the refractive index of the surroundings (e.g. air). $R(\lambda)$ can then be approximated [1] as

$$R(\lambda) \approx \left(\frac{n_{eff}(\lambda) - n_{surr}}{n_{eff}(\lambda) + n_{surr}} \right)^2. \quad (8.11)$$

The effective index of the waveguide mode around $\lambda_P = 785$ nm is $n_{eff}(\lambda_P) \approx 1.584$ and $n_{eff}(\lambda_S) \approx 1.542$ for $\lambda_S = 877$ nm. Assuming the waveguide is surrounded by air, the reflection coefficients are then $R(\lambda_P) \approx 5.1\%$ and $R(\lambda_S) \approx 4.6\%$ respectively. Moreover, the forward scattered SERS signals are always equal to or smaller than the backward scattered signals (see previous chapter). In combination with the relatively small reflection coefficients, the S1 contribution (backward scattered signal) will usually be much larger as compared to the S2 and S3 contributions (forward scattered signals). Therefore it is reasonable to approximate the SERS output power as

$$\boxed{\frac{P_S}{P_{pump}} \approx \eta_A \xi \left(\frac{1 - \left(\frac{1}{e_P e_S}\right)^N}{1 - \left(\frac{1}{e_P e_S}\right)} \right)}. \quad (8.12)$$

It is important to note that coupling effects between individual antennas are not incorporated. Due to coupling effects, it is not possible anymore to define a simple single antenna conversion efficiency η_A , such that the above formula no longer holds. In Figure 8.3 we plot the typical output signal of a Y-splitter for realistic experimental parameters. After an initial increase, the signal quickly saturates when the number of antennas N increases.

8.2.2 Background

As far as the background is concerned, we start with the background generated by the straight antenna arm which is backscattered (B1 contribution). We use the same notations as in the previous chapter in order to express the

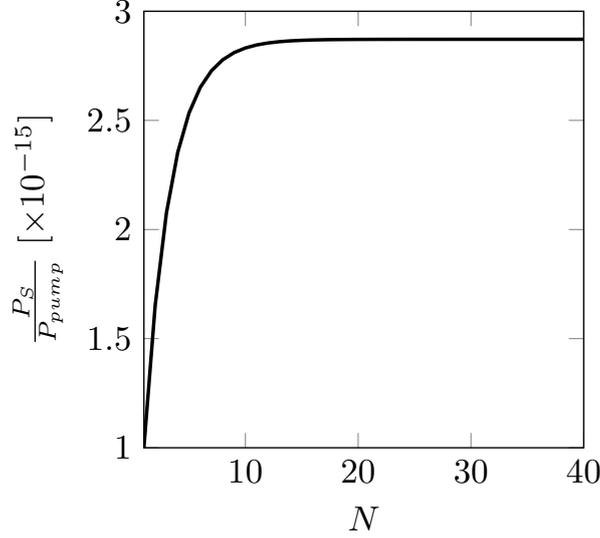


Figure 8.3: Typical output signal $\frac{P_S}{P_{pump}}$ of a Y-splitter. The following values were assumed: $e_P = 1.15$, $e_S = 1.34$, $\eta_A \xi = 10^{-15}$.

scattering strength η_B in a cross-sectional SiN area. The B1 contribution consists of 3 parts, given by

$$\begin{aligned}
\frac{P_{bg}^{B1}}{P_{pump}} = & \eta_B e^{-\alpha_{wg}(L_1+L_s)} T_1 \left\{ \left(\int_0^{L_2} e^{-\alpha_{wg}z} e^{-\alpha_{wg}z} dz \right) \right. \\
& + \left(\int_{L_2+(N-1)\Lambda}^{L_2+(N-1)\Lambda+L_3} e^{-\alpha_{wg}z} e^{-\alpha_{wg}z} dz \right) e_P^{-N} e_S^{-N} \\
& + \left. \sum_{n=0}^{N-2} \left(\int_{L_2+n\Lambda}^{L_2+(n+1)\Lambda} e^{-\alpha_{wg}z} e^{-\alpha_{wg}z} dz \right) e_P^{-(n+1)} e_S^{-(n+1)} \right\} \times \\
& \left\{ T_2 T_B^4 e^{-\alpha_{wg}(L_s+2\pi R+L_4+L_V+L_5)} \right\}. \quad (8.13)
\end{aligned}$$

The first part represents the background generated in the L_2 section, which is not attenuated by the antennas. The second part is the backscattered background from the L_3 section which is attenuated by the antennas two times (initially the pump is attenuated and in backpropagation the Stokes light is attenuated as well). The third part contains the background which is generated inbetween the antennas. In order to write the formulas in a more condensed form, the following approximation is often used throughout this

section:

$$1 - e^{-\epsilon L} \approx \epsilon L \quad (8.14)$$

when $\epsilon L \ll 1$ (which can be due to a small ϵ or a small L or a combination of both). Making use of the assumptions introduced before, one can write the B1 contribution eventually as

$$\begin{aligned} \frac{P_{bg}^{B1}}{P_{pump}} &= \eta_B \xi \left(L_2 + L_3 (e_P^{-N} e_S^{-N}) + \frac{\Lambda}{e_P e_S} \left(\frac{1 - \left(\frac{1}{e_P e_S}\right)^{N-1}}{1 - \left(\frac{1}{e_P e_S}\right)} \right) \right) \\ &\approx \eta_B \xi (L_2 + L_3 (e_P^{-N} e_S^{-N})) \end{aligned} \quad (8.15)$$

considering that the contributions proportional to Λ can in general be neglected since $\Lambda \ll L_{2,3}$. By applying analogous reasonings, it is possible to derive the other contributions as well. For the B2 contribution we take into account forward propagating background light generated along the path $0 \rightarrow L_1 + L_s + L_2$, the path $L_1 + L_s + L_2 \rightarrow L_1 + L_s + L_2 + (N-1)\Lambda$ and the path $L_1 + L_s + L_2 + (N-1)\Lambda \rightarrow L_1 + L_s + L_2 + (N-1)\Lambda + L_3$, which is then eventually backscattered in the bottom arm of the Y-splitter. The B2 contribution is then

$$\begin{aligned} \frac{P_{bg}^{B2}}{P_{pump}} &= \eta_B \xi R(\lambda_S) (e_S^{-2N} (L_1 + L_s + L_2) + L_3 (e_P^{-N} e_S^{-N})) \\ &\quad + \Lambda \frac{e_S^{1-2N}}{e_P} \left(\frac{1 - \left(\frac{e_S}{e_P}\right)^{N-1}}{1 - \left(\frac{e_S}{e_P}\right)} \right) \\ &\approx \eta_B \xi R(\lambda_S) e_S^{-2N} (L_1 + L_s + L_2). \end{aligned} \quad (8.16)$$

Analogously, the B3 contribution is approximately given by

$$\frac{P_{bg}^{B3}}{P_{pump}} \approx \eta_B \xi R_A(N, \lambda_S) e_S^{-2N} (L_1 + L_s + L_2). \quad (8.17)$$

The B4 contribution consists of background light generated by the pump beam which is reflected at the end facet of the straight antenna arm. In order to calculate this contribution, we neglect the background generated along the small $L_2 + (N-1)\Lambda + L_3$ length and only consider the background generated along the $2\pi R + L_4 + L_V + L_5$ length. One can show that this results in

$$\frac{P_{bg}^{B4}}{P_{pump}} = \eta_B \xi e_P^{-2N} R(\lambda_P) (2\pi R + L_4 + L_V + L_5). \quad (8.18)$$

Finally, the B5 contribution is generated by pump light reflected from the Au array. Again we neglect the background generated along $L_2 + (N-1)\Lambda + L_3$ and only consider the background generated in the $2\pi R + L_4 + L_V + L_5$ section, such that

$$\frac{P_{bg}^{B5}}{P_{pump}} = \eta_B \xi R_A(N, \lambda_P) (2\pi R + L_4 + L_V + L_5). \quad (8.19)$$

The overall background for this Y-splitter design is then

$$\begin{aligned} \frac{P_{bg}^B}{P_{pump}} \approx & \eta_B \xi \left(L_2 + \frac{L_3}{e_P^N e_S^N} (1 + R(\lambda_S)) \right. \\ & + (L_1 + L_s + L_2) \left(R_A(N, \lambda_S) + \frac{R(\lambda_S)}{e_S^{2N}} \right) \\ & \left. + (2\pi R + L_4 + L_V + L_5) \left(R_A(N, \lambda_P) + \frac{R(\lambda_P)}{e_P^{2N}} \right) \right). \end{aligned} \quad (8.20)$$

In a typical fabricated configuration, the following lengths are common: $L_2 = L_3 = 50 \mu\text{m}$, $\Lambda = 1 \mu\text{m}$, $L_1 = 80L_2$, $L_4 = 30L_2$, $L_V \ll L_2$, $L_s \approx L_2/2$, $L_5 = 80L_2$, $R = L_2$. For simplicity we set $L_o = L_1 + L_s + L_2 \approx 2\pi R + L_4 + L_V + L_5$, such that $\nu = L_o/L_2 \gg 1$. The background is then approximately given by

$$\begin{aligned} \frac{P_{bg}^B}{P_{pump}} \approx & \eta_B \xi L_2 \left(1 + \frac{1 + R(\lambda_S)}{e_P^N e_S^N} \right. \\ & \left. + \nu \left(R_A(N, \lambda_S) + \frac{R(\lambda_S)}{e_S^{2N}} + R_A(N, \lambda_P) + \frac{R(\lambda_P)}{e_P^{2N}} \right) \right). \end{aligned} \quad (8.21)$$

In order to estimate the relative contributions of each of these terms, we have simulated the reflection coefficient $R_A(N, \lambda)$ for a fixed ($\alpha = 60^\circ$, $L = 120 \text{ nm}$, $\Delta = 40$) bowtie geometry and varying N . In our initial background reduction experiments (see furtheron), bowties of similar dimension were used. Reflection spectra are plotted in Figure 8.4(a). The single antenna reflection peaks around 900 nm. By increasing the antenna number, a splitting with peaks at 785 nm and 990 nm is observed. The strenght of these reflection peaks initially increases with N but will eventually saturate for sufficiently large N . Moreover, the reflection at positions away from the peaks typically decreases with increasing N (compared to the reflection of a single antenna).

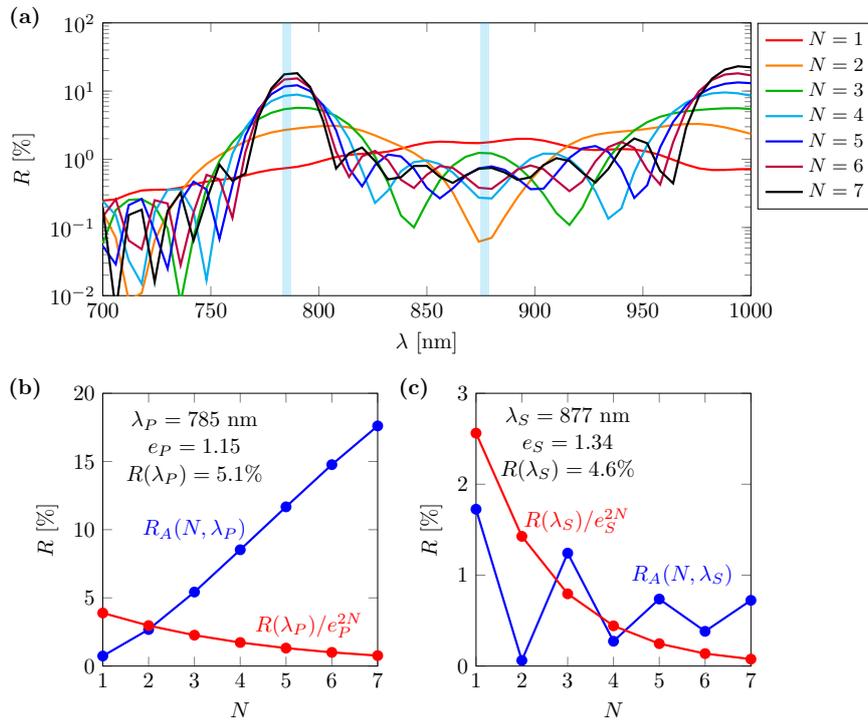


Figure 8.4: (a) Reflection spectra of a ($\alpha = 60^\circ$, $L = 120$ nm, $\Delta = 40$) bowtie antenna array with fixed period $\Lambda = 1$ μm . The Ti adhesion layer and the natural oxide layer were not taken into account. (b-c) Reflection of the antenna array $R_A(N, \lambda)$ and facet reflection $R(\lambda)/e^{2N}$ for (b) $\lambda_P = 785$ nm and (c) $\lambda_S = 877$ nm.

In Figure 8.4(b) and 8.4(c), the reflection at a given pump (λ_P) and Stokes (λ_S) wavelength is shown respectively (both wavelengths are marked by the cyan shaded areas in Fig. 8.4(a)). The linear antenna extinction e_λ is evaluated numerically and the facet reflection coefficient $R(\lambda)$ was calculated before for the two given wavelengths. For small antenna numbers N , the array $R_A(N, \lambda)$ and facet $R(\lambda)/e^{2N}$ reflection are of similar order. While the facet reflection always decays with increasing N , the array reflection will increase (Fig. 8.4(b)) or gradually decrease (Fig. 8.4(c)) depending on the exact wavelength. In case the wavelength coincides with a reflection peak, the array reflection will dominate on the facet reflection for large N (Fig. 8.4(b)). As a result, the background will increase with the number of antennas if the array reflection spectrum contains peaks at the given pump and Stokes wavelength. By properly designing the gold array, one should aim for minimizing $R_A(N, \lambda)$ while maintaining the same SERS signal strength. Figure 8.5 depicts the reflection spectra of a gold array with fixed N , but varying period Λ . Adjusting the period has a drastic effect on the array reflection of a specific wavelength and one can see that for the considered wavelengths (highlighted by the cyan shaded areas) the reflection can decrease by two to three orders. This effect could be ascribed to coupling effects between the different antennas in the array. As a result, the background reduction will be mainly limited by the facet reflection in such a case.

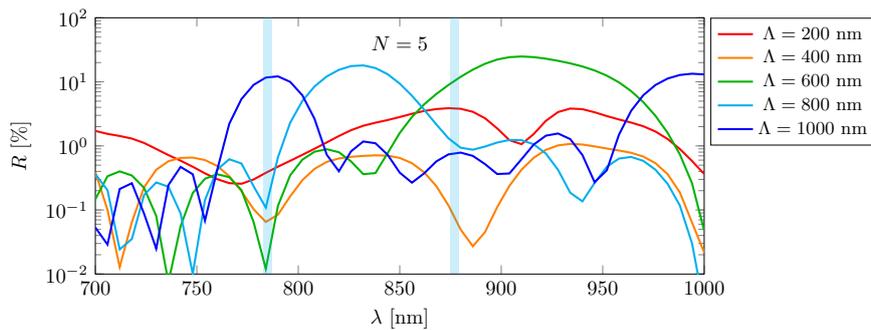


Figure 8.5: Reflection spectra of a ($\alpha = 60^\circ$, $L = 120$ nm, $\Delta = 40$) bowtie antenna array with fixed $N = 5$, but varying period Λ . The Ti adhesion layer and the natural oxide layer were not taken into account.

In conclusion, since $\nu \gg 1$ (typically $O(100)$) the background will mainly result from reflections of the gold array and the end facet. The facet reflections could be mitigated by introducing a properly designed grating which e.g. couples as much pump light as possible out of the waveguide. For single antennas one can of course not change the ‘array’ reflection, but for a larger number of antennas it is possible to mitigate the reflections as well. Depending on the targeted Stokes wavelength, it is hence feasible to achieve a considerable background reduction. The typical background signal of a Y-splitter is plotted in Figure 8.6 for realistic experimental parameters. In case the reflections are dominated by the gold array reflection (Figure 8.6(a)), the background initially increases with the number of antennas. For sufficiently large N the array reflection will saturate, and so will the background. When the reflections are dominated by the facet reflection (Figure 8.6(b)) the background decreases for increasing N . Regardless of which reflection mechanism dominates, the order of magnitude of the background is mainly determined by the length of the input and output sections, i.e. νL_2 . The larger ν , the larger the background. Moreover the lowest background signal that can be achieved is determined by the value $\eta_B L_2$.

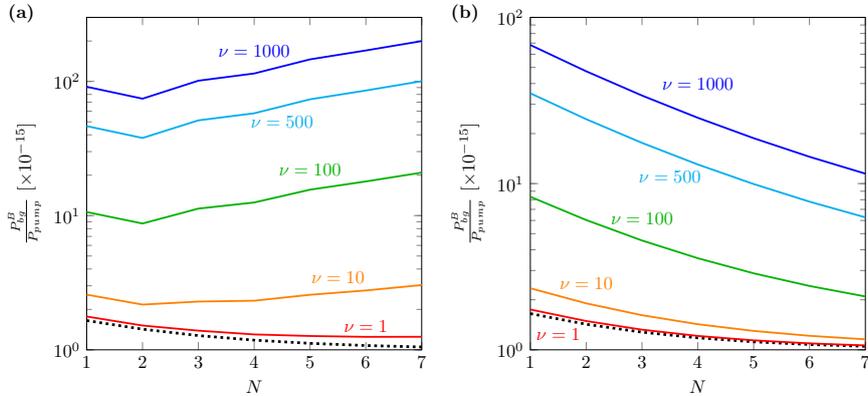


Figure 8.6: Typical background signal $\frac{P_{bg}^B}{P_{pump}^B}$ of a Y-splitter as a function of ν and N . The following fixed values were assumed: $e_P = 1.15$, $e_S = 1.34$, $\eta_B L_2 \xi = 10^{-15}$, $R(\lambda_P) = 5.1\%$, $R(\lambda_S) = 4.6\%$. The black dotted line represents the (ideal) limit when no reflections are present at all. **(a)** $R_A(N, \lambda_P)$ and $R_A(N, \lambda_S)$ values taken from the simulated reflection spectra of Figure 8.4(b) and 8.4(c). **(b)** Fixed $R_A(N, \lambda_P) = R_A(N, \lambda_S) = 0.1\%$.

8.2.3 Signal-to-Noise Ratio

As explained in the previous subsection, the background depends on a variety of factors and can be largely minimized by reducing the reflections from the gold array and the end facet. The typical length of the waveguides discussed in the previous chapter was about $L_{tot} = 10000 \mu m$, with a background signal proportional to $\eta_B L_{tot}$. For the Y-splitter, the background is solely determined by the value $\eta_B L_2$ in the ideal limit of small ν or zero reflections, so a maximum background reduction of L_{tot}/L_2 is expected. Typically $L_2 = 50 \mu m$, which would boil down to a 200-fold background reduction or a shot noise reduction by $\sqrt{200} = 14$. The strength of the backscattered signal obtained with the Y-splitter will at least equal the strength of the forward scattered signal from the previous chapter (for the same guided power) and will usually be slightly larger for increasing N (see also previous chapter). As a result, an overall improvement in Signal-to-Noise Ratio by at least one order of magnitude is expected. In the next section we will discuss some preliminary experimental results. These experiments were still ongoing during the writing of this PhD.

8.3 Preliminary experimental results

In Figure 8.7(a) we depict the Raman spectra obtained on Y-splitters with a different number of antennas N . The Raman spectrum of a straight reference waveguide (REF), obtained through forward scattering, is also shown. The dips in the reference waveguide spectrum are attributed to fabrication errors during e-beam lithography and remaining dirt on the waveguide. The total length of this reference waveguide approximately equals the length $L_1 + 2L_2 + 2L_s + 2\pi R + L_4 + L_V + L_5$ so it is reasonable to compare the background of the reference waveguide with the background of the Y-splitter in order to assess the average background mitigation.

One can see that the background of all Y-splitter designs is lower than the background of the reference waveguide and that the background increases for increasing N . For sufficiently large N there is a saturation of this trend as exemplified in Figure 8.7(b) where the background strength at 1340 cm^{-1} is plotted as a function of N . In the range $N = 1 \rightarrow 20$, the background exhibits the largest changes (more than one order of magnitude) while for $N = 40 \rightarrow 100$ the changes are much smaller. Since the other parts

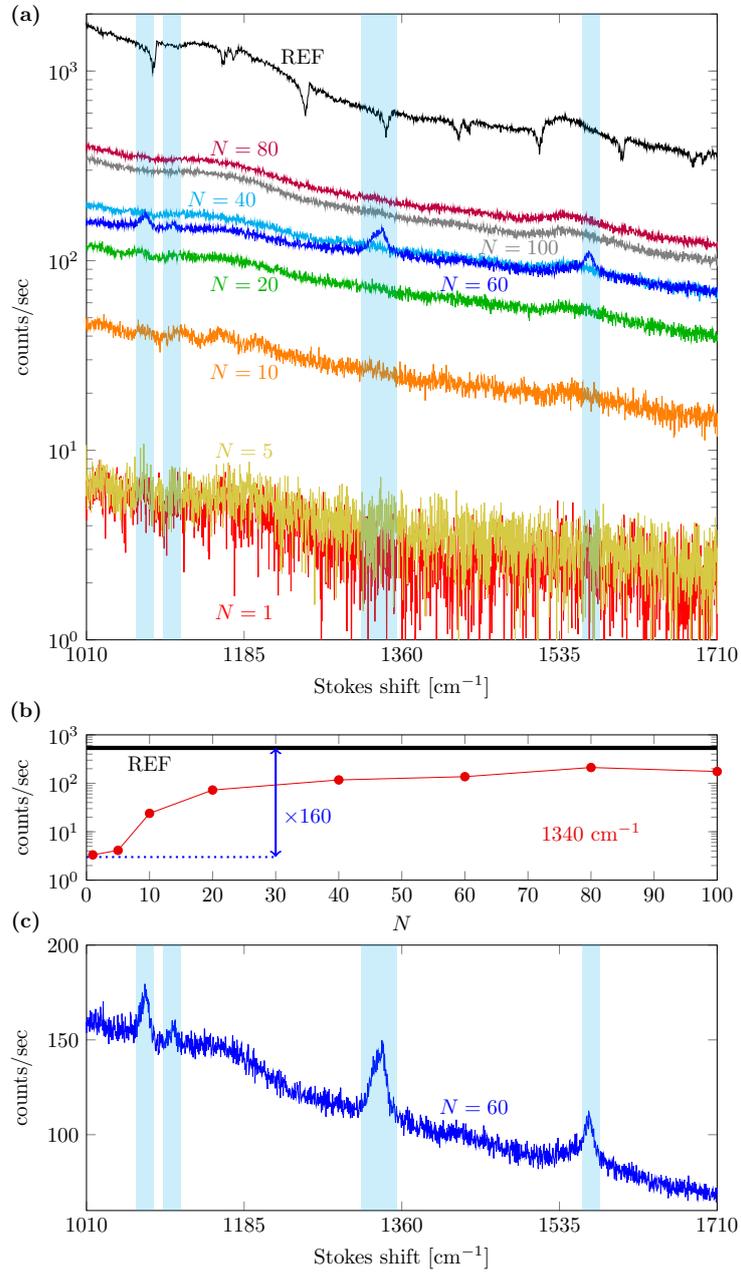


Figure 8.7: (a) Raman spectra of the previously described Y-splitter as a function of N . The Raman spectrum of a straight reference waveguide REF is added as well. (b) Background strength at 1340 cm^{-1} as a function of N . (c) NTP SERS signal obtained from the $N = 60$ Y-splitter.

of the geometry remain the same, such a trend can only be explained by reflections of the gold array (which will depend on N). Considering the model we developed in the previous section, we conclude that either the pump and/or Stokes wavelength matches with one of the maxima in the reflection spectrum because this indeed gives rise to an increasing background with increasing N . Compared to the reference waveguide, we achieved a background reduction by two orders of magnitude.

Peculiarly, an NTP signal is only observed for the $N = 60$ case (which is shown in Figure 8.7(c)). In principle, for backscattering one would expect that the signal saturates for larger N (see previous chapter), hence if the $N = 60$ signal is observed, then signals on all $N > 60$ waveguides should be observed as well. Potential explanations are an imperfect coating process or remaining dirt layers on the gold which inhibit proper binding of NTP. Since the antenna array was not optimized to minimize reflections from the gold array, the overall SNR is still low (≈ 5 for the 1340 cm^{-1} peak without averaging). As shown in the previous section, it is possible to reduce the array reflection by two to three orders of magnitude, which would result in a shot noise reduction of at least one order. For such a large antenna number $N = 60$, the facet reflection will moreover be negligible as well. So, by properly tuning the antenna array we expect to increase the SNR from 5 to 50. Currently we are improving the Y-splitter design in order to optimize the SNR , which ultimately should enable the detection of extremely weak Raman signals from a single antenna (i.e. $N_{min} = 1$ as discussed in the previous chapter). This would allow quantitative and multiplexed detection of several biological and chemical substances, such as viruses or DNA, on a fully integrated platform.

8.4 Conclusion

In this chapter we included some ongoing work involving Y-splitter chip designs to mitigate the large Raman background. We developed an analytical model identifying the different contributions to either the SERS and background signal. It was found that the SERS signal is mainly dominated by the backscattered light, while the background is predominantly due to reflections on the gold array or the end facet of the Y-splitter. Moreover, we have shown that (theoretically) it is possible to reduce the array reflection by properly tuning the array period. Initial experiments show that the background can be reduced up to two orders of magnitude, but it is expected that optimized Y-splitter designs can push

this further in order to allow detection of signals originating from a single antenna.

References

- [1] Baets, R.; Roelkens, G. Course notes *Photonics*, 2009–2010.

9

Conclusion and Outlook

9.1 Conclusion

In this thesis we investigated the possibility of integrating nanoplasmonic antennas on top of high-index contrast waveguides for on-chip Surface Enhanced Raman Spectroscopy (SERS). This single mode SERS probe can be combined with other photonic, fluidic or biological functionalities in order to develop a highly dense integrated platform for the multiplexed detection of extremely weak Raman signals.

At first instance, several fabrication strategies based on e-beam lithography were considered for the development of such a hybrid nanophotonic-plasmonic platform. Patterning nanoplasmonic antennas generally requires an e-beam resist which is much thinner than the thickness of the waveguides on which they have to be integrated. For an improved definition of these structures it is therefore desirable to fabricate the antennas first, and only afterwards etch a waveguide around them. This is realized by a two-step e-beam lithography process starting from a blank silicon nitride (SiN) slab wafer. In the first step we defined the antennas. Different processes were considered based on the type of e-beam resist. A negative e-beam resist (FOX-12 or ma-N 2400.6) initially requires the deposition of a metal stack on the complete wafer, after which the structures are milled by an ion miller. In case of a positive e-beam resist (PMMA), the resist is spun

on a SiN slab wafer and the metal is only deposited after exposure. A lift-off then eventually removes the residual metal. Structures defined using a negative resist generally have a more reproducible and smooth shape, while structures defined through lift-off are generally more grainy. On the downside, FOx-12 is very hard to remove, making the eventual structure less accessible to analytes. On the contrary, ma-N 2400.6 does not suffer from this issue and was hence found to be the best resist for patterning integrated nanoplasmonic antennas with high resolution and reproducibility. Since the negative resist process requires an ion milling of the remaining metal stack, we also investigated the impact of this milling on the waveguide losses of the eventual device. It was found that the ion milling has no detrimental impact on the loss. In the second e-beam step, the SiN waveguides were defined using ma-N 2403 and etched with a commercial ICP etcher. The best waveguide losses obtained with this two-step e-beam process were about 3 dB/cm around $\lambda = 785$ nm.

Subsequently we performed a comparative analysis of several types of integrated nanoplasmonic antenna and established the possibility of integrating narrow gap antennas on a SiN waveguide. In our study we compared single rod, double rod and bowtie antennas and investigated the plasmon resonance tuning as a function of the geometrical parameters of the antenna. The theoretically predicted trends were confirmed by our experiments and experimental interaction efficiencies of 19% around 800 nm were observed for bowtie antennas. We also numerically verified that the interaction efficiency can be optimized by tuning the waveguide geometry. Moreover, numerical simulations of the field enhancement were performed and it was shown that of the three investigated antenna types, the bowtie antenna has the highest maximum Raman enhancement. Apart from the broad dipolar resonance, we also predicted that bowtie antennas with a large apex angle exhibit a narrower (higher Q) dark resonance as a result of the propagation induced phase delay. This was also confirmed by experiment. In a natural way, evanescent waveguide excitation can hence be used to excite dark plasmon resonances.

After establishing plasmon resonance tuning of integrated nanoplasmonic antennas in the 700 – 1000 nm region, we turned our attention to the investigation of on-chip SERS, which was the key challenge of this PhD. Since it is important to isolate the pure SERS signal and avoid any other contribution from waveguide generated spontaneous Raman scattering, we have used 4-nitrothiophenol (NTP). By immersion in an NTP:ethanol solution it is assumed that a self-assembled monolayer of NTP forms on

top of the Au through a Au-S bond. Subsequent rinsing with pure ethanol will remove the unbound NTP such that only the Au surface is coated with NTP. Measurements on a non-functionalized reference waveguide (without any gold) indeed confirm that no NTP remains on the SiN. We also showed that the SERS spectra could be attributed to a plasmon resonance effect, meaning that a stable and reproducible enhancement factor can be associated with it. The effect of the number of antennas on the SERS spectrum was also investigated and it was found that an optimum antenna number exists such that the SERS signal is maximized. Moreover, an analytical on-chip SERS model was developed in order to establish the relevant design parameters and figure of merit. Since not all antennas on a chip are the same, we have also extended our ideal on-chip SERS model in order to incorporate potential differences among the antennas. Based on a dedicated fitting procedure to this generalized model, we were able to assess the experimental spread on the design parameters. Subsequently we compared the experimental data with the theoretical predictions. An excellent correspondence was found between the theoretically predicted and experimentally observed absolute Raman power coupled back into the waveguide mode. As a result, the suggested platform enables a complete quantitative control on the Raman enhancement and subsequent coupling of the enhanced Stokes light with the underlying waveguide. In combination with other on-chip spectral functionalities, such as arrayed waveguide gratings, the presented SERS probe is forecasted to allow multiplexed and quantitative detection of extremely weak Raman signals on a highly dense integrated platform.

During our on-chip SERS experiments we also identified a trade-off between signal optimization and background reduction. The shot noise associated to the inherent SiN background ultimately limits the smallest features that can be detected. The attenuation introduced by the antennas can reduce this shot noise, albeit sometimes at the cost of a reduction in the SERS enhancement. Strategies to simultaneously optimize the signal and mitigate the noise are therefore very interesting. Dedicated on-chip splitters are suggested to realize this goal. A Y-splitter where an array of antennas is patterned on the short straight arm while the two bend parts are used for excitation and collection respectively, could result in a more efficient separation of the SERS signal and the SiN background. This work is still ongoing, but preliminary results already show that the background can be reduced by a factor of 100 for such Y-splitter designs.

9.2 Outlook

Finally we will give an outlook on possible future research directions. In no way this list is exhaustive, but it rather reflects some interesting problems and ideas that originated during the course of this PhD.

9.2.1 Background reduction

As already described before, the issue of background associated shot noise sets a limit on the smallest features that can be detected. If one would like to use this platform for single molecule detection or the detection of ultralow analyte or particle concentrations, then it is of paramount importance to mitigate the background. We already embarked upon this background mitigation by the use of dedicated Y-splitters, but it would be interesting to explore other intricate designs and combine them with on-chip filtering such that the SERS signal can be delivered to an on-chip spectrometer for readout. Our single mode SERS probe provides the most efficient way of delivering Raman light to an on-chip spectrometer, making it ideal for complete integration of a SERS-based detection platform.

9.2.2 Non-Au SERS

In all our experiments we used Au (with a Ti adhesion layer) as the plasmonic metal. While having a good chemical stability, making it ideal for biological and chemical sensing applications, Au is not CMOS-compatible. Therefore it would be interesting to explore other metals or combinations of metals and dielectrics which are CMOS-compatible. Moreover it would be necessary to define the nanoplasmonic antennas by a higher throughput scheme because e-beam lithography is relatively slow. A possible route is the use of nanoimprint technology in combination with CMOS-compatible dielectric waveguides. This would greatly enhance the mass-reproducibility such that these on-chip SERS probes can be made cheap and disposable.

9.2.3 Nanoparticle trapping

The local field enhancement near a metallic antenna also creates a strong gradient in the electric field, as a result of which the antenna can exert an optical force proportional to the gradient of the intensity. [1] This so-called optical tweezing effect is also observed in the focus of a laser beam,

but in order to trap smaller particles it becomes increasingly difficult with diffraction-limited systems. Due to the huge field confinement, plasmonic antennas could be used for the stable trapping of a nanoparticle. While trapped, one could envisage reading out the Raman spectrum of that trapped particle. Potential applications could be the on-chip detection of viruses, contaminants or DNA.

9.2.4 Quantum Photonics

Another fascinating opportunity for the developed integrated antennas is their use in on-chip quantum photonic circuits. While photons are ideal for carrying quantum information, they are difficult to store for a long time however. In order to implement a quantum memory for quantum information carried by photons, it is required to map the photon quantum state to another medium. The spin of an electron is an excellent candidate for the storage of quantum information. [2] However, a commonly encountered problem is the limited coupling efficiency between the quantum system and the light, hindering the scalability and the rate of information transfer. Moreover it is essential to obtain the optimal parameters which maximize the SNR of the optical spin state readout. [3] Due to the resonant behaviour of a nanoplasmonic antenna, one can significantly change the spontaneous decay rate of a (quantum) emitter in the vicinity of such an antenna. [4] The antennas can then be engineered to exhibit a broad resonance that covers the emission spectrum of the quantum emitter, which eventually should allow a significant enhancement of the photon collection efficiency into a specific optical mode. This Purcell enhancement can also be used to optimize the optical coupling efficiency and the SNR of the spin state readout. [3]

References

- [1] Juan, M.L.; Gordon, R.; Pang, Y.; Eftekhari, F.; Quidant, R. Self-induced back-action optical trapping of dielectric nanoparticles *Nature Phys.* **2009**, *5*, 915–919.
- [2] Buckley, S.; Rivoire, K.; Vučković, J. Engineered quantum dot single-photon sources. *Rep. Prog. Phys.* **2012**, *75(12)*, 126503.
- [3] Wolf, S.A.; Rosenberg, I.; Rapaport, R.; Bar-Gill, N. Purcell-enhanced optical spin readout of Nitrogen-Vacancy centers in

diamond. Preprint on arXiv: <http://arxiv.org/abs/1505.01006>.

- [4] Pelton, M. Modified spontaneous emission in nanophotonic structures. *Nature Photon.* **2015**, 9, 427–435.

