FACULTY OF ENGINEERING AND ARCHITECTURE

On-Chip Raman Spectroscopy: Background Challenges and Their Mitigation

Kristof Reynkens

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Photonics Engineering

Supervisors
Prof. Roel Baets, PhD - Prof. Stéphane Clemmen, PhD

Department of Information Technology
Faculty of Engineering and Architecture, Ghent University

October 2021

GHENT UNIVERSITY

ISBN 978-94-6355-525-8
NUR 926
Wettelijk depot: D/2021/10.500/73

Members of the Examination Board

Chair

Prof. Filip De Turck, PhD, Ghent University

Other members entitled to vote

Prof. Nicolas Le Thomas, PhD, Ghent University
Jérôme Michon, PhD, Université Paris-Saclay, France
Prof. Andre Skirtach, PhD, Ghent University
Haolan Zhao, PhD, Ghent University

Supervisors

Prof. Roel Baets, PhD, Ghent University
Prof. Stéphane Clemmen, PhD, Ghent University

Acknowledgements

Firstly, I would like to start by expressing my gratitude towards my promotors Roel and Stéphane for their guidance and advice over the past few years. Although things looked bleak at some point during the PhD we still managed to turn things around despite the limited time frame and the presence of a worldwide pandemic. The many discussions and meetings we had over the years helped me to raise my game and gave me alot of insight and learnings into the world of photonics. Furthermore, feedback and corrections were readily available when needed. Secondly, I would like to thank the members of the jury for their valuable comments and questions that further improved the quality of this manuscript.

Special thanks to the supporting staff of the Photonic Research Group: Ilse, Ilse, Kristien, Bert, Jasper, Steven, Muhammad and Liesbet who selflessly provided help and advice over the course of my PhD and were always open for a friendly chat. I am also very grateful towards Haolan and Ali for their many hours of aid over the years during experimental work, simulations and calculations. Of course, I would also like to thank the other members of the Raman team: Nina, Zuyang, Yang and Xiaomin. The many (often lengthy) discussions we had in the Raman meetings provided me with valuable feedback and perspective when needed. Furthermore, I would like to thank Nicolas for the fruitful discussions we had during these meetings. Also a shout-out to Tom for the lab work we did together. Finally, I would like the acknowledge the help from the COCOON group members: Michiel, Jolien and Christophe in performing the ALD depositions required for this work.

The life of a PhD student can often be challenging and frustrating as a seemingly endless amount of problems seem to pop up around every corner. Luckily, this also creates a camaraderie with fellow colleagues as they have to deal with the same kind of difficulties. I now look back with great fondness to the wonderful atmosphere we had during lunches and after-work ping pong games which never failed to lighten the mood. First and foremost I would like to thank Jeroen for the many small coffee breaks and hilarious e-mail conversations we had over the years. I would also like to thank the fifth-floor lunch crew: Mimis (So,... begins story), Clemens (Hoe gaat het met jou?), Mahmoud (What‘s uuuuuuuup?), Irfan (NépolyYen the great), Emad, Mattias, Sarah, Lukas, Camiel, ... These sophisticated lunch conversations

really helped to shape me as a human being. I am also grateful towards my current and former office mates: Khannan (thanks for introducing me to the Expanse!), Xiaoning, Amin, Ivo, Xiangfeng, Anna and Iman. Furthermore, I also enjoyed the many games with the members of the PRG football team (the infrared devils): Javad (who learned me a thing or two of dressing in style), Gilles, Isaac, Stijn, Jing, Sören, ... I would also like to thank Margherita and Savvina for kindly hosting me during many dinner parties and events. Last but not least I would like to express my gratitude towards the ping pong crew: Alejandro: Maaaan, I hope you achieve your dream of opening a papaya farm one day (Please be careful to not touch his new ping pong bat though), Alessio (most friendly and patient guy I know), Irfan again (twice mentioned, what an honour!), Abdul (an even dirtier player than me), Stijn and Ewoud.

Behind the scenes there were also many family members and friends who supported me throughout the years. I would like to start by thanking my mother for all the unconditional love and support she provided ever since I was a small child. Furthermore, I would like to express my gratitude towards my grandmother, my two brothers: Tom (who has a PhD in statistics, I might have mentioned this once or twice) and Niels, their partners: Liese and Camille, my godson: Arthur, my mother-in-law: Christel and my father-in-law: Guido and his family: Julie, Louis and Ida. Also special thanks to the (partial) members of the Brussel crew: Andreas, Domien, Joris, Mathijs, Simon, Eloy and Sander who provided much needed joy during though (lock-down) times. I am also grateful towards all the people who directly or indirectly followed my progress and provided good wishes along the way. Finally, I would like to thank my sun, my moon and all my stars, my girlfriend Hannah who stood by my side for all these years and had to endure many canceled weekend activities, holidays and trips such that I could work for the PhD. I would like to take the opportunity here to express my gratitude towards her for all the sacrifices she made and for the joy she brings in my life every day.

Brussels, 2021
Kristof Reynkens

Contents

List of Acronyms

A

ALD	Atomic Layer Deposition
AL_2O_3	Aluminum Oxide
APD	Avalanche Photodetector
AWG	Arrayed Waveguide Grating

B

BD	Balanced Detector
BOX	Buried Oxide Layer

C

CARS	Coherent Anti-Stokes Raman Spectroscopy

CCD	Charge-Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CSRS	Coherent Stokes Raman Scattering
CW	Continuous Wave

D

dB	Decibel
DC	Directional Coupler
DI	Deionized
DMAB	4-dimercaptoazobenzene
DMSO	Dimethyl Sulfoxide
DUV	Deep-UV

E

EDFA	Erbium-Doped Fiber Amplifier

F

FBG	Fiber Bragg Grating
FDTD	Finite-Difference Time-Domain
FP	Fabry-Perot
FSR	Free Spectral Range
FWHM	Full Width at Half Maximum

I

IMLL	Integrated Mode-Locked Laser
IPA	Isopropyl Alcohol

L

LSPR	Localized Surface Plasmon Resonance

M

MLL	Mode-Locked Laser

MMI	Multi-Mode Interferometer
MZI	Mach-Zehnder Interferometer

N

NIR	Near Infrared
NL	Nonlinear
NR-DFWM	Nonresonant Degenerate Four-Wave Mixing
NTP	4-Nitrothiophenol
NWERS	Nanophotonic Waveguide-Enhanced Raman Spectroscopy

O

OSA	Optical Spectrum Analyzer

P

PECVD	Plasma Enhanced Chemical Vapour Deposition
PICs	Photonic Integrated Circuits

PSPP	Propagating Surface Plasmon Polariton

R

RF	Rejection Filter
RPM	Rounds Per Minute

S

SBR	Signal-to-Background Ratio
SEM	Scanning Electron Microscopy
SERS	Surface-Enhanced Raman Spectroscopy
SE-CARS	Surface-Enhanced Coherent anti-Stokes Raman Spectroscopy
SE-SRS	Surface-Enhanced Stimulated Raman Spectroscopy
Si_3N_4	Silicon Nitride
SiO_2	Silicon Oxide
SNR	Signal-to-Noise Ratio

SOI	Silicon On Insulator
SPP	Surface Plasmon-Polariton
SPR	Surface Plasmon Resonance
SRG	Stimulated Raman Gain
SRL	Stimulated Raman Loss
SRS	Stimulated Raman Spectroscopy

T

TA	Transient Absorption
Ta_2O_5	Tantala
TE	Transverse Electric
TiO_2	Titania
TM	Transverse Magnetic
TPA	Two-Photon Absorption

U

UV	Ultraviolet

X

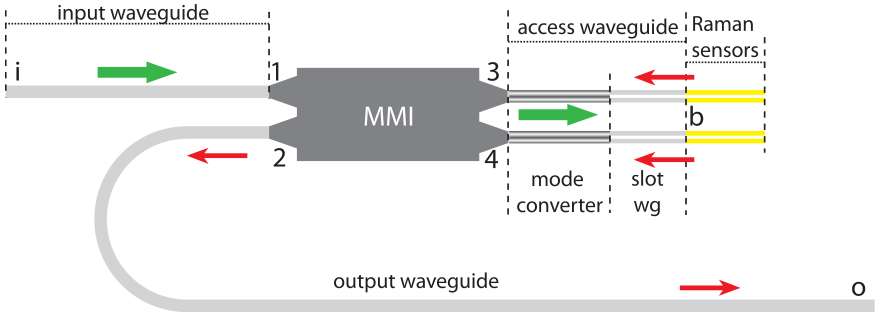
XPM	Cross-Phase Modulation

Nederlandse Samenvatting

Ramanspectroscopie is erg in trek omdat het gedetailleerde informatie geeft over de chemische structuur, fase en kristalliniteit van onderzochte moleculen. Het laat toe om de trillingen van de moleculen op een niet-invasieve manier te onderzoeken en de onderzochte monsters hebben geen voorbereidingswerk nodig. Het is daarnaast ook compatibel met nabije-infrarood golflengtes waardoor het populair is in het chemisch, biomedisch en farmaceutisch onderzoeksveld. Ramanspectroscopie is ondanks zijn vele kwaliteiten nog altijd voornamelijk beperkt tot een labo-omgeving. De oorzaak hiervoor ligt bij de intrinsieke zwakte van het Ramanverstrooiingsproces waardoor de Raman-apparaten (zoals hoog-vermogen lasers en diep-gekoelde detectoren) duur en omvangrijk zijn. Hierdoor is er recent een grote interesse ontstaan voor silicium fotonica waarbij het doel is om alle elementen van het Raman spectroscopisch systeem (lasers, detectoren, Raman sensoren, filters en spectrometers) op een optische chip van enkele millimeters groot te integreren. Omdat deze chips via CMOS-compatibele technologie gefabriceerd worden, zijn ze daarnaast ook compact, goedkoop en massaal produceerbaar. Een geïntegreerde Raman sensor werd voor het eerst in onze groep gedemonstreerd (Dhakal et al.| Optics Letters, 2014) in de vorm van nanofotonische golfgeleiderspectroscopie (NWERS) waarbij er een druppel isopropanol gedetecteerd werd met een single-mode Si_3N_4 strip golfgeleider. NWERS zorgt voor een versterking van het Raman signaal t.o.v. een Raman microscoop omdat het gebruik maakt van het versterkte optische veld nabij het oppervlak van de golfgeleider en de lange interactielengtes mogelijk gemaakt doordat het licht gevangen is in de golfgeleider. Ondanks dat NWERS vele voordelen heeft, zijn er momenteel ook nog twee voorname nadelen aan verbonden: 1) de versterking van het Raman signaal is niet voldoende om niet langer gebruik te moeten maken van diep-gekoelde detectoren (deze zijn moeilijk te integreren op een chip) 2) in de kern van de diëlektrische golfgeleiders wordt er een ongewenste fotonische achtergrond gegenereerd die het ratio van het signaal-tot-achtergrond (SBR) verslechtert.

Deze fotonische achtergrond kan drastisch gereduceerd worden door een nanoplasmonische slot golfgeleider i.p.v. een diëlektrische golfgeleider te gebruiken zoals Raza et al. (APL Photonics, 2018) aangetoond hebben. Deze slots laten toe om de interactielengte tot een fractie te verkleinen omdat ze gebruik maken van plasmonische versterking waardoor ze nog steeds Raman signalen van dezelfde sterkte kunnen genereren als in NWERS. Omdat de sterkte van de fotonische

achtergrond proportioneel is met deze interactielengte leidt dit tot een sterke reductie van de achtergrond. Als we deze oppervlakte-versterkte Ramanspectroscopie (SERS) sensor gaan integreren met een circuit dat in staat is om het verstrooide Raman licht te analyseren, is het natuurlijk ook belangrijk dat er geen verdere achtergrond gegenereerd wordt in dit circuit. Helaas hebben fotonische componenten zoals filters en spectrometers een lengte van meerdere millimeters waardoor deze een sterke achtergrond genereren. Daarom werd er in vorige metingen op de nanoplasmonische slot golfgeleider enkel gebruik gemaakt van een minimale toegangsgolfgeleider als het enige element tussen de SERS sensor en het chip-facet waar het laser licht op gefocust wordt. Helaas laat zo een circuit-layout niet toe om andere spectroscopische componenten te integreren omdat de achtergrond sterk afhankelijk is van de lengte van deze toegangsgolfgeleider. Daarom heb ik in deze thesis een alternatief circuit onderzocht op basis van een multi-mode interferometer (MMI). Ik heb aangetoond dat dit circuit erin slaagt om het Raman verstrooide licht van de SERS sensoren efficiënt te verzamelen en het laat daarnaast ook om de integratie van spectroscopische elementen toe op zijn outputgolfgeleider zonder dat deze een grote achtergrond bijdrage genereren.



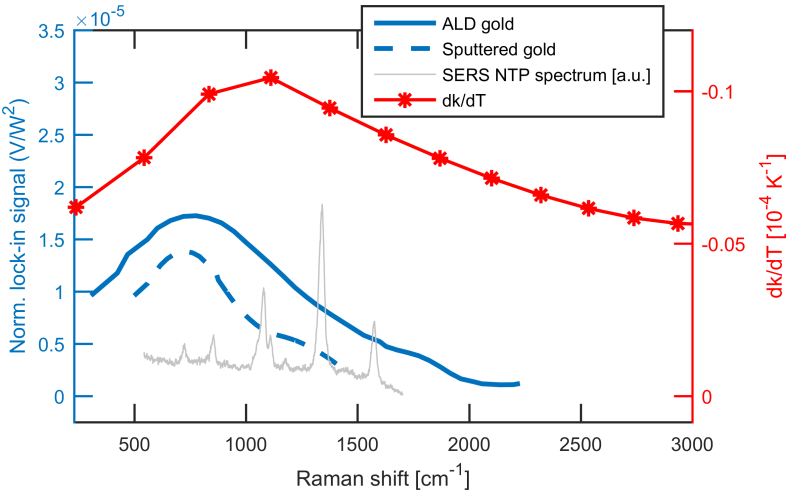
Figuur 1: MMI-nanoplasmonische slot configuratie waar het voorwaarts-propagerende excitatielicht (groene pijl) van het achterwaarts-propagerende Raman signaal gescheiden wordt (rode pijl). Grijze elementen zijn van Si_3N_4 gemaakt terwijl de gele elementen voornamelijk uit goud bestaan. Herprint van Reynkens et al. (Optics Express, 2020).

Onze architectuur met een 2x2 MMI en achterwaartse collectie van het Raman verstrooide light wordt getoond in Fig.1. De MMI zorgt ervoor dat het excitatielicht van de input naar de sensor kan propageren en dat het Raman-verstrooide licht daarna van de sensor naar de outputgolfgeleider kan propageren. Het excitatielicht en de bijhorende achtergrond blijven voorwaarts propageren en worden in de sensor geabsorbeerd. De afwezigheid van het excitatielicht in de outputgolfgeleider laat toe om lange ($\sim$ mm) spectroscopische elementen te integreren zonder dat deze een extra achtergrond bijdrage leveren. Door het Raman licht van de twee Raman sensoren (bedekt met 4-Nitrothiophenol (NTP)) te collecteren in de outputgolfgeleider hebben we in deze thesis aangetoond dat de MMI-plasmonische slot configuratie effectief werkt. Verder hebben we bepaald dat de achtergrond in deze configuratie voornamelijk in de toegangsgolfgeleider (& Raman sensor) en de outputgolfgeleider

gegenereerd worden en dat de MMI en de inputgolfgeleider geen grote bijdrage hebben. De achtergrond die in de outputgolfgeleider gegenereerd wordt heeft zijn oorsprong in de reflecties die de excitatielaser op zijn pad tegenkomt naar de SERS sensors. Deze zorgen ervoor dat er een beetje excitatielicht in de outputgolfgeleider teruggekoppeld wordt (simulaties tonen aan dat de reflectie aan het plasmonische slot facet de grootste is). Het feit dat de MMI geen grote achtergrond bijdrage heeft, is te verwachten omdat deze redelijk kort is (112 μm) en multi-mode zodat het meeste van de achtergrond gegenereerd in de MMI naar buiten uitgestraald wordt en niet naar de outputpoort koppelt. De sterkte van de achtergrond gegenereerd in de output- en toegangsgolfgeleider is van dezelfde grootorde ondanks dat de outputgolfgeleider veel langer is (11mm vs 43.5 μm). Er is dus maar een minimale kost in achtergrond tussen deze configuratie en degene die in een gewone meting van de SERS sensor gebruikt wordt (met alleen maar de toegangsgolfgeleider). Finaal hebben we ook een theoretische berekening gemaakt die aantoont dat onze MMI-configuratie een superieur signaal-tot-achtergrond ratio heeft (voor de relevante outputgolfgeleider lengtes) t.o.v. een simpele transmissie-configuratie en een directionele koppelings-configuratie die alle twee ook in staat zijn om dezelfde circuit functionaliteit uit te voeren. De MMI-plasmonische slot configuratie is dus een goede kandidaat voor de verdere integratie met andere spectroscopische componenten.

In het eerste deel van de thesis hebben we gefocust op het mitigeren van de foton achtergrond. In het tweede deel focussen we op het versterken van het Raman signaal zodat we niet langer diepgekoelde detectie nodig hebben. We weten dat door het gebruik van de plasmonische slot, de achtergrond door de diëlektrische kern sterkt onderdrukt wordt. Helaas hebben de plasmonische slots net als NWERS het probleem dat hun gegenereerde Raman signalen niet sterk genoeg zijn om het gebruik van diep-gekoelde detectoren te vermijden. Daarom hebben Zhao et al. (Optics Letters, 2018) gestimuleerde Ramanspectroscopie (SRS) onderzocht op een Si_3N_4 strip golfgeleider (met CW lasers en een lock-in detectie schema). Het gebruik van SRS op de chip zorgt voor een toename van het gegeneerde signaal met vijf grootteordes t.o.v. NWERS, hierdoor kan de detectie op kamertemperatuur gebeuren. Op zijn beurt heeft SRS dan weer last van de achtergrond die gegenereerd wordt door het gebruik van diëlektrische strip golfgeleiders ($\sim$ mm). Daarom is de volgende stap voor de Raman sensor, de combinatie van plasmonische versterking en SRS zodat de interactielengte tot een fractie kan herleid worden en ondertussen de detectie op kamertemperatuur kan uitgevoerd worden. Het tweede doel van deze thesis is dus om oppervlakte-versterkte gestimuleerde Ramanspectroscopie (SE-SRS) en oppervlakte-versterkte coherente anti-Stokes Ramanspectroscopie (SE-CARS) te onderzoeken (CARS is een veelvoorkomend alternatief voor SRS). In dit werk hebben we eerst een theoretische vergelijking gemaakt tussen de verschillende Raman technieken waarbij we voor elke techniek dezelfde onderzochte stof (NTP) en hetzelfde gemiddelde Pump (en Stokes) vermogen genomen hebben. Onze berekeningen hebben aangetoond dat SE-CARS en SE-SRS beide zorgen voor een toename van het Raman signaal met meer dan drie grootteordes t.o.v.

SERS op de plasmonische slot. Deze verbetering komt ten koste van een meer complex excitatie/collectie schema (i.e. gepulseerde lasers voor SE-CARS en lock-in detectie voor SE-SRS) maar maakt een detectie op kamertemperatuur mogelijk. Op hetzelfde moment zorgen SE-CARS en SE-SRS er ook voor dat er een verbetering is in signaal-tot-achtergrond ratio van minstens drie grootteordes t.o.v. SRS op de diëlektrische strip golfgeleider. In dit werk verkiezen we SE-SRS boven SE-CARS omdat het veel gevoeliger is voor de detectie van relevant biologische stoffen. Deze stoffen kunnen gemakkelijk een concentratie hebben die drie grootteordes lager ligt dan de monolaag van NTP die we in onze berekeningen gebruiken. Onder deze voorwaarden zal SE-CARS een signaal generen dat drie grootteordes onder dat van SE-SRS ligt omdat het een kwadratische afhankelijkheid heeft met de concentratie t.o.v. lineair voor SE-SRS. Daarom is SE-SRS gevoeliger voor de detectie van verdunde stoffen en stoffen met een lage concentratie. Voor een geïntegreerde Raman sensor in de context van een labo-op-chip is de detectie van relevante biologische stoffen zeer belangrijk waardoor SE-SRS op dit gebied een duidelijk voordeel heeft t.o.v. SE-CARS.



Figuur 2: Ongewenst SRS signaal gemeten op de nanoplasmonische slots met ALD goud (vol blauw & linker as) en gesputterd goud (gestippeld blauw). Variatie van de extinctie coëfficiënt dk/dT van goud met de golflengte (rood & rechter as). Verwachte NTP spectrum voor de SE-SRS meting (grijs & linker as [a.u.]). Herprint van Reynkens et al. (Optics Letters, 2021).

We hebben in deze thesis aangetoond dat de experimentele demonstratie van SE-SRS op de nanoplasmonische slot uitdagend is. De intrinsieke opwarming van de nanoplasmonische slot (die veroorzaakt wordt door de lineaire absorptie) zorgt voor een thermo-optisch effect dat het Stokes signaal zodanig beïnvloedt dat het Raman signaal van de onderzochte stof (NTP) volledig overschaduwd wordt. Dit thermo-optisch effect is niet alleen twee grootteordes sterker dan onze geschatte

Raman response, maar het heeft ook zijn eigen niet-triviale spectrum dat overeenkomt met dat van de thermo-extinctie coëfficiënt van goud in de literatuur (zie Fig. 2). Om de oorsprong van dit ongewenste achtergrondsignaal te begrijpen, moeten we eerst kijken naar het feit dat in SE-SRS de Raman gain zich manifesteert als een kleine toename van de Stokes intensiteit. Door de pump laser te moduleren, zal de Raman imprint ook gemoduleerd worden zodat het opgepikt wordt door een lock-in versterker. Helaas zorgt het gebruik van een gemoduleerde pump er ook voor dat er een gemoduleerde opwarming in het goud gebeurt door zijn absorptie. De periodische opwarming zorgt voor een verandering van de absorptie en brekingsindex van het goud wat leidt tot een extra parasitaire modulatie van de Stokes laser. De verandering in de brekingsindex en absorptie worden beschreven door de thermobrekingscoëfficiënt (dn/dT) en thermo-extinctie coëfficiënt (dk/dT) en onze berekeningen tonen aan dat dk/dT het sterkste parasitaire signaal veroorzaakt. In Fig. 2 zien we dat dk/dT een sterke spectrale overeenkomst toont met onze spectra. Het is niet triviaal om dit thermo-optisch effect waar te nemen (met één laser) omdat de geïnduceerde niet-lineaire verandering in transmissie (door een toename van het vermogen) te zwak is om in een simpel transmissie experiment geobserveerd te worden. We hebben de thermische oorsprong van ons parasitair signaal verder bepaald door aan te tonen dat het duidelijk zwakker wordt bij een toename van de modulatie frequentie, wat wijst op een traag proces zoals het thermisch effect. Verder hebben onze berekeningen aangetoond dat er enkel een milde temperatuurmodulatie nodig is in de plasmonische slot (2.5 K) om zo een sterke parasitaire achtergrond te genereren. Thermische simulaties op een 3D model van de plasmonische slot hebben daarnaast meer inzicht gegeven in de opwarmingsdynamiek en ondersteunde de experimentele data. Deze simulaties volgen daarnaast ook het bi-exponentiele frequentie gedrag van het parasitair signaal wat zijn oorsprong heeft in de snellere hitte verspreiding in het goud t.o.v. de omliggende materialen. Deze elementen tonen aan dat ons parasitair signaal van thermische oorsprong is. We hebben finaal ook nog verschillende strategieën besproken om deze thermische achtergrond in de toekomst te onderdrukken.

English Summary

Raman spectroscopy is a powerful sensing tool that provides detailed information on the chemical structure, phase and crystallinity of the probed molecules. In Raman spectroscopy, the vibrational modes of molecules are probed in a non-invasive manner without the need for additional sample preparation. Furthermore, its compatibility with NIR wavelengths allows for the analysis of a wide range of molecules in the chemical, biomedical and pharmaceutical field. Despite these qualities, Raman spectroscopy has been mainly limited to laboratory environments. This can be understood from the inherent weakness of the Raman scattering process itself which requires the use of expensive and bulky instrumentation such as high power lasers and deeply-cooled CCD detectors. In order to overcome this limitation, silicon photonics has recently emerged which holds the promise of integrating the components of a conventional Raman spectroscopic system (i.e. lasers, detectors, Raman sensors, spectrometers and filters) onto a millimeter-sized chip. Furthermore, these chips are fabricated using CMOS-compatible technology and are therefore compact, cheap and mass-producible. Our group (Dhakal et al. | Optics Letters, 2014), for the first time demonstrated an integrated Raman sensor by performing nanophotonic waveguide enhanced Raman spectroscopy (NWERS) on a single mode Si_3N_4 strip waveguide covered with isopropanol. In NWERS, the Raman signal is boosted (compared to a conventional Raman microscope) by taking advantage of the field enhancement near the dielectric waveguide surface and the long interaction lengths enabled by confining the optical mode to the waveguide core. Despite its advantages, NWERS still has to deal with two main challenges: 1) the provided signal boost is still not strong enough to avoid the use of deeply-cooled CCD detectors (which are difficult to integrate on a photonic chip) 2) an undesired photon background is generated in the dielectric waveguide core which lowers the signal-to-background ratio (SBR) of any acquired Raman spectra.

Concerning the background, Raza et al. (APL Photonics, 2018) showed that this photon background can be drastically reduced by using nanoplasmonic slot waveguides over dielectric waveguides. The plasmonic field enhancement of those slots allows for the interaction length to be reduced to a fraction of that required in NWERS while still proving the same signal enhancement. Of course, when integrating such a plasmonic sensor with a circuit capable of analyzing the Raman scattered light, it will be vital that no further background is generated in the circuit surrounding the sensor. Unfortunately, for photonic components such as spectral

notch filters and arrayed waveguide gratings, the photon background contribution becomes significant as their lengths can easily exceeds millimeters. That is why previous characterizations on the plasmonic slot used a minimalistic access waveguide as the only interface between the facet of the chip (where an external excitation beam is focused) and the Raman plasmonic sensor itself. However, such a circuit layout does not allow for the integration of spectroscopic elements after the sensor since the strength of the background is critically depended on the length of that access waveguide. This is why I explored an alternative circuit layout based on a multi-mode interferometer (MMI). I showed that this circuit is capable of capturing the Raman scattered light of the surface-enhanced Raman spectroscopy (SERS) sensor while simultaneously allowing for the integration of spectroscopic elements on its output waveguide without inducing neither a large background penalty nor a large signal penalty.

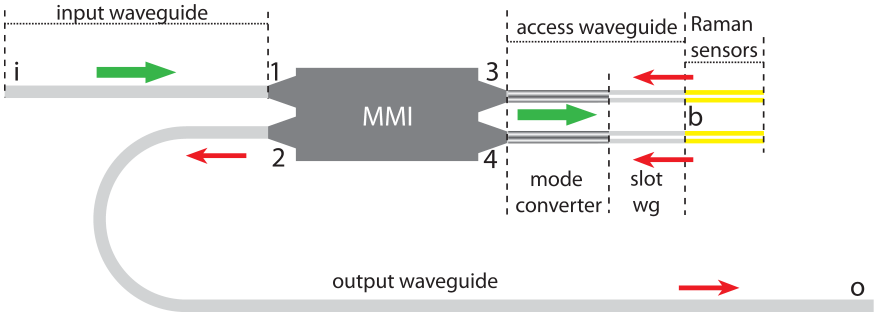


Figure 3: Schematic of the configuration relying on two SERS sensors and a MMI separating the forward-propagating excitation beam (green arrow) from the backward-propagating Raman scattered light (red arrows). Grey elements are made of Si_3N_4, while yellow elements are primarily made of gold. Image reprinted from Reynkens et al. (Optics Express, 2020).

In Fig. 3, our designed architecture is shown which relies on a 2x2 MMI and backward Raman collection. We can see that the MMI lets the excitation beam propagate from the input facet to the sensor and then lets the Raman back-scattered light propagate from the sensor to the output waveguide. The excitation beam and associated background from the entrance waveguide keep on forward propagating in the plasmonic sensor and are finally absorbed. The absence of a strong excitation beam in the output waveguide allows for the integration of mm-long spectroscopic elements without a strong background penalty. In this work, we experimentally demonstrated that the MMI-plasmonic slot configuration works effectively by collecting the Raman response of the two 4-Nitrothiophenol (NTP)-coated SERS sensors at the output waveguide. By varying both the access waveguide and output waveguide lengths we determined that the background in our circuit mainly originates from the back-scattered photon background in that short access waveguide (& the Raman sensor) and from the output waveguide. Other background sources due to the MMI and the input waveguide do not play a major role. The background generation in the output waveguide is due to forward-Raman scattering process

of the remaining excitation beam which is reflected at the (plasmonic slot) interfaces and coupled to the output waveguide. The absence of a strong background contribution of the MMI is expected since it is short (112 μm) (therefore inducing little background) and highly multimode such that most of the background photons generated in the MMI will be radiated out and not collected at the output port. The background contributions of the access waveguide and output waveguide are of the same order despite the output waveguide being significantly longer (11 mm vs 43.5 μm). The penalty in terms of extra photon background generation compared to a common SERS measurement on the plasmonic slot (with only the access waveguide) is thus minimal. Finally, we made a theoretically estimate that showed that for typical output waveguide lengths (mm to cm), our MMI-based configuration performs significantly better than an alternative configuration relying on forward-Raman scattering and a directional coupler-configuration which are both able to perform the same functionality.

Having mitigated the presence of the photon background in the first part of this thesis, the second part focuses on increasing the Raman signal in order to avoid the use of a deeply-cooled CCD. While we know that the plasmonic slot greatly suppresses the photon background contribution of the dielectric waveguide core, its enhancement is not strong enough to avoid the use of expensive deeply-cooled detectors. For this reason, on-chip stimulated Raman spectroscopy (SRS) has been explored by Zhao et al. (Optics Letters, 2018) using CW beams, a lock-in detection scheme and a Si_3N_4 strip waveguide. The stimulation with a Stokes beam results in a signal enhancement of more than five orders of magnitude compared to NWERS which allows for room-temperature detection. However, it suffers on its turn from a significant background contribution by making use of mm-long dielectric strip waveguides. Compared to the latter demonstration, using a plasmonic slot over a dielectric strip is expected to provide major advantages since it would greatly reduce the background contribution while at the same time allowing for room-temperature detection. Therefore our second goal in this PhD is to explore on-chip surface-enhanced stimulated Raman spectroscopy (SE-SRS) and surface-enhanced coherent anti-Stokes spectroscopy (SE-CARS) on the nanoplasmonic slot (where CARS is a common alternative coherent Raman technique to SRS). In this work we made a calculation comparing the different on-chip Raman techniques assuming the same analyte (a monolayer of NTP) and the same average pump (and Stokes) power. This calculation showed that both SE-CARS & SE-SRS provide a signal enhancement of more than three orders of magnitude compared to SERS on the nanoplasmonic waveguide. This improvement comes at the cost of a more complex excitation/detection scheme (i.e. pulsed lasers for SE-CARS and lock-in detection for SE-SRS) but allows for room temperature detection. Compared to on-chip SRS, SE-CARS and SE-SRS provide an improvement in the signal-to-background ratio of (more than) three orders of magnitude (thereby drastically reducing the background contribution of the waveguide core). SE-SRS is preferred over SE-CARS in this work since it is much more sensitive to biologically relevant analytes. For these analytes, the concentration to be detected could easily drop by three

or more orders of magnitude compared to the monolayer of NTP assumed in our calculations. In that regime, the generated signal levels of SE-CARS drops an additional three orders of magnitude compared to SE-SRS because of its quadratic dependence on the concentration (compared to linear for SE-SRS). This thus makes SE-SRS more sensitive to diluted analytes and better equipped to detect low analyte concentrations. As the detection of biologically relevant concentrations will be very important for an integrated Raman sensor in a lab-on-chip context, waveguide-based SE-SRS provides a clear advantage over SE-CARS in this aspect.

In this work, we showed that the experimental demonstration of SE-SRS on the nanoplasmonic slot is challenging. The heating intrinsic to linear absorption in the gold nanostructure induces a thermo-optic effect that affects the Stokes beam overshadowing the Raman signature of the analyte (NTP in the present case). Not only is this thermo-optic effect two order of magnitude stronger than our estimated Raman response, but it also presents its own non-trivial spectrum because of the spectral variation of the thermo-extinction coefficient of gold (see Fig.4). To better

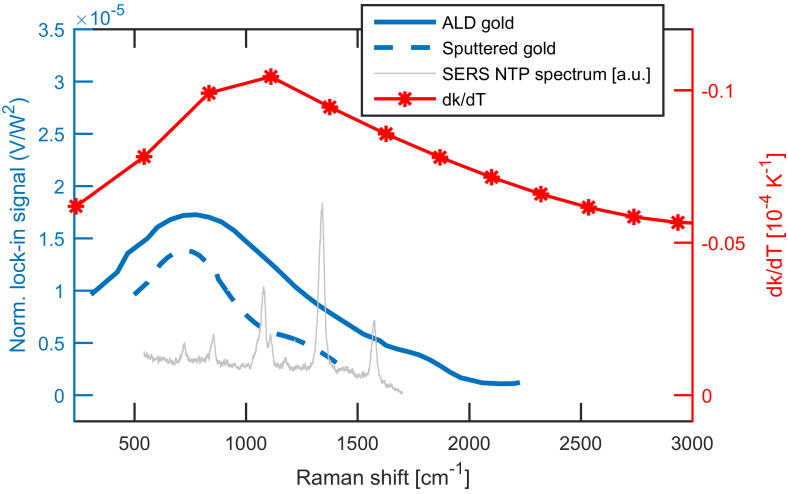


Figure 4: Recorded spurious SRS spectra for our nanoplasmonic slots made with ALD gold (bold blue & left-axis) and sputtered gold (dotted blue). Wavelength dependence of the extinction coefficient dk/dT of gold (red & right-axis). Expected NTP spectrum for our SE-SRS measurement (grey & left-axis [a.u.]). Image reprinted from Reynkens et al. (Optics Letters, 2021).

understand the origin of this spurious signal we need to consider that the Raman gain is imprinted as a small intensity increase on-top of the already present Stokes beam. By modulating the pump beam, this gain is imprinted as a shallow modulation on the Stokes beam that is picked up by a lock-in amplifier. Unfortunately, this modulated pump beam also leads to a modulated heating of the gold via its absorption. This periodic heating affects the absorption and refractive index of the gold that in turn leads to a parasitic modulation of the Stokes beam. The change of the refractive

index and absorption with temperature are quantified via the thermorefractive coefficient (dn/dT) and thermo-extinction coefficient (dk/dT) (where we calculated that the parasitic effect due to the dk/dT will be dominant). The wavelength dependence of dk/dT is presented in Fig. 4 and shows a clear similarity to our own measured spectra. Note that this thermo-optic effect is not trivial to observe (with a single laser) as the induced non-linear change in transmission (with increasing power) is too weak to be observed in a simple transmission experiment. We further confirmed the thermal origin of this spurious SRS signal by demonstrating its decline with the modulation frequency which points at a slow process like the thermal effect. Supportive calculations show that only a mild (2.5 K) thermal modulation in the plasmonic slot is sufficient for such a strong spurious background to be present. Thermal simulations on a 3D model of the plasmonic slot gave more insight into the dynamic of this heating and supported the experimental data. Furthermore, these thermal simulations recreated the bi-exponential frequency behaviour of the spurious SRS signal we deduced in our measurements which can be attributed to the faster heat flow in the gold compared to the surrounding materials. All this evidence points towards the fact that our captured spurious SRS signal has a thermal origin. Finally, we proposed several mitigation strategies to suppress this thermal background.

1 Introduction

1.1 Background and rationale

Raman spectroscopy is a non-invasive sensing technique used to detect and analyze a wide range of molecules in chemistry, biology, physics and life sciences [1–9]. In Raman spectroscopy, the vibrational frequencies of molecules are probed resulting in a distinct spectral fingerprint which allows to distinguish a diverse set of molecules in a chemically complex environment [10]. Furthermore, the compatibility to perform Raman spectroscopy in the NIR region allows for the detection of biologically-relevant molecules in aqueous solutions. Raman spectroscopy has however been mainly limited to in-lab applications because of the inherent weakness of the Raman scattering process and the resulting need for expensive and bulky optical instrumentation. Recently, there has been a push to miniaturize the Raman spectroscopic system to cheap handheld devices that can be used on-site [11]. In this context, integrated photonics has emerged which holds the potential to integrate the full Raman spectroscopic system on a single photonic chip (i.e. lasers, detectors, Raman sensors, spectrometers and filters). These photonics chips are fabricated using CMOS-compatible techniques (i.e. already existing standard process technology in a state-of-the-art CMOS fab) and are therefore cheap, mass-producible, robust and reliable [12]. Concerning the Raman sensor itself, in our group, Dhakal et al. [13] were the first ones to demonstrate nanophotonic waveguide enhanced

Raman spectroscopy (NWERS) using a single mode Si_3N_4 strip waveguide. This on-chip Raman sensor makes use of the evanescent field around the dielectric waveguide to probe the analyte over long optical path lengths (beyond the diffraction limit). Furthermore, the high index contrast between the waveguide core and cladding leads to a large enhancement of the evanescent field near the waveguide surface. Therefore waveguide-based Raman spectroscopy is accompanied with a large signal enhancement compared to a conventional (confocal) Raman microscope [14]. However, two main challenges still remain for NWERS. Firstly, an expensive deeply-cooled detector is still required as the provided signal enhancement is not strong enough for room-temperature detection. Chip-integration of such detectors will be challenging and therefore hinder the demonstration of a fully-integrated Raman spectroscopic system. Secondly, an unwanted photon background [15] is generated in the dielectric waveguide core that leads to a reduction of the signal-to-background ratio and thus lowers the detection sensitivity.

In order to reduce this photon background contribution, the nanoplasmonic slot waveguide has been proposed [16]. This plasmonic slot makes use of plasmonic field enhancement and surface-enhanced Raman spectroscopy (SERS) to drastically reduce the background generation while maintaining the same level of (broadband) Raman enhancement as NWERS. Besides the Raman sensor itself, it is also important that the analysing circuit surrounding the sensor does not have a major background contribution (since this would lead to a further deterioration of the signal sensitivity). However, spectroscopic elements such as spectral filters [17] or arrayed waveguide gratings (AWG) [18] typically have dimensions in the order of several millimeters which would results in a significant background contribution from their dielectric waveguides. In previous characterisations on the plasmonic slot, the access waveguide connecting the plasmonic slot to the chip facet was simply minimized in order to avoid an additional background penalty. However, increasing that access waveguide to the lengths required for the integration with other spectrosopic elements is not that trivial since the signal-to-background ratio (SBR) will quickly deteriorate [16]. Therefore, such a circuit layout does not allow for the integration of the Raman sensor with other spectroscopic elements. *In this PhD project, we will explore an alternative circuit layout based on a multi-mode interferometer (MMI). We will investigate if this circuit is capable of efficiently capturing the Raman scattered light of the SERS sensor while simultaneously allowing for the integration of spectroscopic elements on its output waveguide without inducing a large background penalty.*

Concerning the Raman sensor itself, while the nanoplasmonic slot greatly reduces the Si_3N_4 background contribution, it still requires deeply-cooled detectors in order to capture the Raman scattered light. Therefore, alternatives Raman techniques such as coherent Raman spectroscopy (CRS) can be employed to further

boost the Raman signal by many orders of magnitude (in CRS the molecular vibrations are coherently driven by the addition of an extra driving laser). Note that CRS is commonly implemented in two modalities: stimulated Raman spectroscopy (SRS) & coherent anti-stokes Raman spectroscopy (CARS). In [19], an on-chip SRS measurements has been demonstrated on a Si_3N_4 strip waveguide using CW beams and a lock-in detection scheme. In this measurement, a signal enhancement of more than five orders of magnitude has been shown compared to NWERS allowing for room-temperature detection. However, dielectric waveguides suffer from the photon background that superposes itself onto the Raman spectrum. *Therefore, in this PhD project we will explore the obvious next step for the Raman sensor which is to combine plasmonic enhancement with coherent Raman spectroscopy on the waveguide platform. In this way, we can combine the advantages of both techniques (i.e. a low photon background contribution and room-temperature detection). We will first theoretically determine if surface-enhanced stimulated Raman spectroscopy (SE-SRS) or SE-CARS is the most suited candidate for sensing before proceeding to the experimental demonstration of SE-SRS or SE-CARS on the nanoplasmonic slot waveguide.*

1.2 Thesis outline

In **Chapter 2**, I provide the theoretical basis of Raman scattering based on classical theory. First, I introduce the concept of spontaneous Raman scattering followed by a discussion on the Raman cross section and selection rules. Next, I give the theoretical background on coherent Raman scattering. Afterwards, I detail two coherent Raman techniques: CARS and SRS and give a small discussion on the differences between both. Finally, I give an introduction on plasmonic physics and quickly touch upon the enhancement mechanism in surface-enhanced (coherent) Raman scattering.

In **Chapter 3**, I firstly introduce the concept of waveguide based Raman spectroscopy and discuss its advantages over free-space Raman spectroscopy. Afterwards, I will review the general formulas to calculate the generated Raman signal for waveguide-enhanced (stimulated) Raman spectroscopy (which we will use in Chapter 4). Finally, I give an overview of the experimental work that has already been done on NWERS together with its remaining challenges (i.e. a limited signal generation and a strong background contribution due to its guiding material).

In **Chapter 4**, I will go into detail on the alternative on-chip Raman techniques to NWERS that can (partially) overcome its challenges. I start by giving a summary on the experimental work that has been performed on on-chip SERS & SRS together with their remaining challenges. Next, I will give an overview of the existing

literature on SE-SRS using free-space excitation (and collection). Then, I will make my own theoretical calculation of the expected signal strength and photoemission background for on-chip SE-SRS and compared those results to on-chip SERS & SRS. Afterwards, I discuss the literature on SE-CARS and I make a comparing calculation of on-chip SE-CARS to on-chip SE-SRS. Finally, I motivate why I prefer to explore SE-SRS over SE-CARS in this work and provide an overview table of my theoretical comparison.

In **Chapter 5**, I discuss the design and experimental demonstration of a analysing circuit surrounding the plasmonic waveguide sensor based on a MMI. Firstly, I discuss the MMI-design and its fabrication and continue with details on the followed measurement procedure. Then, I discuss the experimental results obtained and determine the origin of the extra background generation compared to a common SERS measurement on the plasmonic slot. Finally, I compare our MMI-configuration to alternatives architectures such as a simple transmission-configuration and a direction coupler-configuration and discuss possible improvements to the device .

In **Chapter 6**, I experimentally explore SE-SRS on nanoplasmonic waveguides. Firstly, I discuss the sample fabrication followed by the description of the measurement setup and the experimental verification of it is correct functioning by performing SRS measurements on a optical fiber and a dielectric slot waveguide. Then, I present the result of our SE-SRS measurements on the nanoplasmonic slot waveguide and demonstrate that a combined photothermal and thermo-optic effect in the gold material induces a strong spurious signal that constraints the detection limit for the analyte. Next, these experimental results are supported by theoretical estimates on the strength of the photothermal effect and further confirmed by thermal simulations for which a 3D model of the nanoplasmonic slot is build. Finally, I propose several methods to reduce or counteract this background.

In **Chapter 7**, I summarize the most important conclusions of this PhD project together with some possible future improvements.

In **Appendix A**, I give additional details on the simulations of the reflections in the MMI-plasmonic slot configuration

In **Appendix B**, I show additional illustrations of the 3D thermal simulations

In **Appendix C**, I give a short summary on the research project that encompassed the first two years of my PhD. In this project, the focus was on performing a CARS-like measurement on 22-nm polystyrene beads that were dropcasted on top of a silicon nitride slot waveguide. The ultimate goal was to demonstrate that the on-chip Raman sensor is capable of detecting the acoustic vibrations of (viral) nanoparticles. Unfortunately, experimental results did not follow and it was concluded that it was no longer fruitful to continue this line of experimental work.

1.3 Publications

This dissertation has led to the following list of publications in international peer-reviewed journals.

1.3.1 Publications in international journals

1. **K. Reynkens**, S. Clemmen, H. Zhao, A. Raza, T. Vanackere, A. Stassen, M. Van Daele, J. Dendooven and R. Baets, "Gold-induced photothermal background in on-chip surface enhanced stimulated Raman spectroscopy", *Optics Letters*, 46(5), pp. 953–956, 2021.

2. **K. Reynkens**, S. Clemmen, A. Raza, H. Zhao, J. S. D. Penaranda, C. Detavernier and R. Baets, "Mitigation of photon background in nanoplasmonic all-on-chip Raman sensors", *Optics Express*, 28(22), pp. 33564–33572, 2020.

2

Introduction to Raman spectroscopy

Raman spectroscopy is a sensing technique that probes the vibrational frequencies of molecules resulting in a distinct spectral fingerprint without the need for extra labeling. In Raman scattering, the incoming pump photon is inelastically scattered by a molecule resulting in an energy loss (or gain) of the scattered photon compared to the incident photon. This type of Raman scattering is called Stokes (anti-Stokes) Raman scattering (see Fig. 2.1). The energy difference between the incident and scattered photon corresponds to the energy of a vibrational mode. The observed frequency shifts in the Raman spectra are then linked to the vibrations of the probed molecules in order to do sensing.

In the rest of this chapter we will provide the theoretical basis of Raman scattering based on classical theory. Firstly, we introduce spontaneous Raman scattering and discuss the Raman scattering cross section and selection rules. Then, we describe coherent Raman scattering and show how it originates from the third-order susceptibility. Afterwards, we detail on two coherent Raman techniques: coherent anti-Stokes Raman scattering (CARS) and stimulated Raman scattering (SRS). Finally, we give an introduction on plasmonic physics and the enhancement mechanism in surface-enhanced (coherent) Raman scattering.

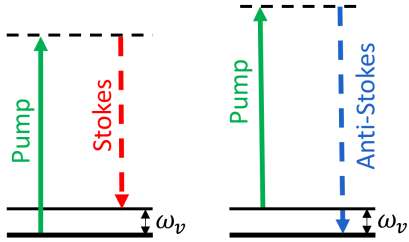


Figure 2.1: Energy level diagram of Stokes and anti-Stokes spontaneous Raman scattering. The vibrational (virtual) energy states correspond to the full (dotted) lines.

2.1 Spontaneous Raman scattering

In this section we make use of the classical theory of light-matter interaction to describe Raman scattering. A complete description involving all the energy levels of the molecules requires quantum theory and is out of the scope of this thesis (a detailed quantum based description of Raman scattering can be found in [20–22]). Firstly, we derive the vibrational modes of a simple biatomic molecule as a harmonic oscillator. Then we describe the Raman response at the Stokes and anti-Stokes wavelength generated by the interaction between the incoming light wave and the probed molecule. The classical theory in this section largely follows [23–25].

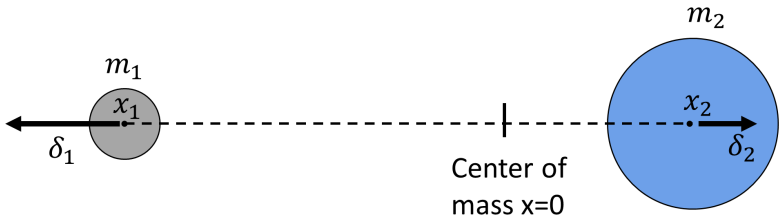


Figure 2.2: The simple diatomic molecule used as a model for the vibrational modes.

Let us now consider the vibration of the diatomic molecule (with mass m_1 and m_2) shown in Fig. 2.2, where the molecular bond can be modeled as a spring with a spring constant K and the origin is placed at the center of mass. When an external (optical) force is applied on these molecules the atoms are displaced by a distance δ_1 and δ_2 respectively while the center of mass is conserved such that:

$$\delta_1 = \frac{m_2}{m_1}\delta_2 \tag{2.1}$$

Let us now assume that the restoring force of the spring (F_s) is linear with the

displacement (which is valid for small displacement from the equilibrium position). Such that the restoring force becomes:

$$\begin{aligned} F_s &= -K(\delta_1 + \delta_2) \\ &= -K\frac{m_1 + m_2}{m_2}\delta_1 \\ &= -K\frac{m_1 + m_2}{m_1}\delta_2 \end{aligned} \tag{2.2}$$

By filling in Newton's second law we get the harmonic oscillator equation:

$$\frac{m_1 m_2}{m_1 + m_2}\left(\frac{d^2\delta_1}{dt^2} + \frac{d^2\delta_2}{dt^2}\right) = -K(\delta_1 + \delta_2) \tag{2.3}$$

$$\zeta\frac{d^2Q}{dt^2} = -K\,Q \tag{2.4}$$

Where $\zeta = \frac{m_1 m_2}{m_1+m_2}$, the displacement $Q = \delta_1 + \delta_2$ and the differential equation can be solved as:

$$Q = Q_0(e^{-j\omega_v t} + c.c.) \tag{2.5}$$

Which is an harmonic oscillation with an angular frequency $\omega_v = 2\pi\nu_v$. Where the fundamental vibrational frequency ν_v corresponds to:

$$\nu_v = \frac{1}{2\pi}\sqrt{\frac{K}{\zeta}} \tag{2.6}$$

The potential energy V for this harmonic oscillator now corresponds to a parabolic well:

$$V = \frac{1}{2}KQ^2 \tag{2.7}$$

Fig.2.3 shows the vibrational energy levels calculated for our diatomic molecule in the harmonic oscillator approximation. For real molecules, the Morse potential is a better approximation since it takes into account the repulsion between nuclei and the fact that the bond between the atoms will break at the dissociation energy (D_e) for too large displacements between the nuclei. We can see that in the harmonic approximation the energy levels are equally spaced ($\hbar\omega_v$) while in the Morse

approximation the energy levels between the higher order modes decreases. The spacing between the ground state ($\upsilon = 0$) and ($\upsilon = 1$) is almost the same in both approximations.

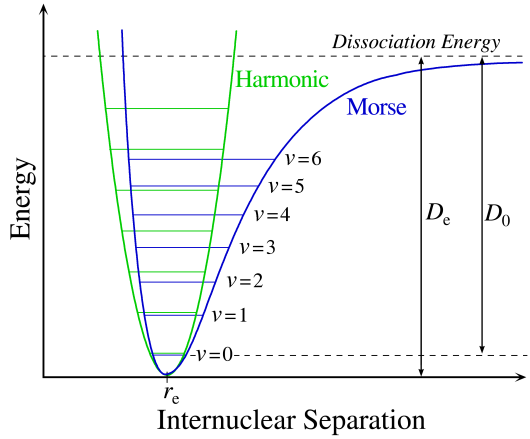


Figure 2.3: Harmonic oscillator potential well (green) and Morse potential well (blue) for a diatomic molecule. The different vibrational energy levels correspond to the tie lines [26].

Next, we consider the effect of the electromagnetic wave on our diatomic molecule, where the electric field corresponds to:

$$\tilde{E} = \tilde{E}_0 e^{(-j\omega_p t)} + c.c. \tag{2.8}$$

Which induces a dipole μ in the molecule corresponding to:

$$\tilde{\mu} = \alpha \tilde{E} \tag{2.9}$$

With α the polarizability, which indicates how easily an external electric field can induce a dipole moment μ in the molecule. Besides the specific molecule, α is also dependent on the frequency of the incoming light field. Since electrons are bound to the nuclei, the presence of vibrational modes will also have an effect on the electron motion. Therefore, the polarizability can be expressed as a Taylor series with a static component that is also present in the hypothetical absence of the nuclear modes (α_0) and a variable first-order component that includes the effect of the vibrational mode(s) on the polarizability:

$$\alpha \approx \alpha_0 + \left(\frac{\partial \alpha}{\partial Q}\right)_0 Q + ... \tag{2.10}$$

Where $\left(\frac{\partial\alpha}{\partial Q}\right)_0$ corresponds to the coupling strength between motion of the electron cloud and the molecular vibrations. The induced dipole can now be expressed as:

$$\begin{aligned}\tilde{\mu} &\approx \left(\alpha_0 + \left(\frac{\partial\alpha}{\partial Q}\right)_0 Q\right)\tilde{E} \\ &\approx \alpha_0\tilde{E}_0 e^{-j(\omega_p t)} \\ &+ \left(\frac{\partial\alpha}{\partial Q}\right)_0 Q\tilde{E}_0 e^{-j(\omega_p-\omega_v)t} + \left(\frac{\partial\alpha}{\partial Q}\right)_0 Q\tilde{E}_0 e^{-j(\omega_p+\omega_v)t} + \text{c.c.}\end{aligned} \tag{2.11}$$

Which thus can be separated in three frequency contributions:

$$\tilde{\mu} = \tilde{\mu}(\omega_p)^R + \tilde{\mu}(\omega_p - \omega_v)^S + \tilde{\mu}(\omega_p + \omega_v)^{AS} \tag{2.12}$$

The first term corresponds to a dipole that oscillates and re-emits radiation at the same frequency as the incident light wave (ω_p). This process is called Rayleigh scattering and is classified as an elastic scattering process since the light is re-emitted at the same energy (frequency) as there is no energy exchange with the molecule. Rayleigh scattering is the most common form of scattering and is typically 3-5 orders of magnitude stronger than Raman scattering. The other two terms in Eq. 2.12 correspond to Raman scattering and are inelastic scattering processes since the light is radiated at a frequency $\omega_p - \omega_v$ (Stokes scattering) and at $\omega_p + \omega_v$ (anti-Stokes scattering). From quantum theory it is shown that Stokes scattering is much stronger than anti-Stokes Raman scattering. This can be understood from the Maxwell-Boltzmann distribution law which states that at room temperature the ratio between Stokes and anti-Stokes scattering intensity corresponds to [27]:

$$\frac{P_s}{P_{as}} = \left(\frac{\omega_p - \omega_v}{\omega_p + \omega_v}\right)^4 e^{\frac{\hbar\omega_s}{k_b T}} \tag{2.13}$$

Where k_b is the Boltzmann constant, $\hbar$ the reduced Planck constant and T is the temperature. Since at room temperature the population ratio of the first excited state ($\nu = 1$) is much smaller than the ground state ($\nu = 0$) of the vibrational states (see Fig. 2.3), Anti-Stokes scattering is less likely to occur than Stokes scattering. At room temperature, the conversion ratio of the scattered intensity to incoming intensity roughly corresponds to 10^{-5},10^{-8},10^{-11} for Rayleigh, Stokes and anti-Stokes scattering respectively assuming a Raman shift of 2500 cm^{-1}.

For more complex molecules Eq. 2.11 can be extended to:

$$\begin{aligned}\tilde{\mu} &= \alpha_0 \tilde{E}_0 e^{-j(\omega_p t)} \\ &+ \tilde{E}_0 \sum_i \left[\left(\frac{\partial \alpha}{\partial Q_i} \right)_0 Q_i e^{-j(\omega_p - \omega_i)t} + \left(\frac{\partial \alpha}{\partial Q_i} \right)_0 Q_i e^{-j(\omega_p + \omega_i)t} \right] + \text{c.c.}\end{aligned} \tag{2.14}$$

Where we can see that complex molecules have multiple vibrational frequencies corresponding to different vibrational modes. For a nonlinear molecule there exist 3N-6 vibrational modes (3N-5 for linear molecules), where N is the number of atoms. Each of these vibrational modes has a distinct frequency associated to it (however there can be degenerate normal modes) and the complex molecular vibration can be resolved into this set of 3N-6 simple harmonic motions. By linking theses vibrational frequencies to the frequency shift observed in the Raman spectra we can thus identify chemical species where the spectra can be seen as a molecular fingerprint. Note however that not all these 3N-6 modes are Raman active (i.e. scatter Raman light) since we can see from Eq. 2.11 that $\left(\frac{\partial \alpha}{\partial Q_i} \right)_0$ needs to be nonzero for there to be a Raman scattering component. This means that for a vibrational mode to be Raman active there needs to be a change in the polarizability during the molecular vibration. Analogously, in IR-absorption spectroscopy a vibrational mode is IR active if the dipole moment of the molecule is changed during vibration. Therefore, certain vibrational modes can be Raman active but not IR active and vice versa (see Fig. 2.4), both spectroscopic techniques can thus be complementary.

Let us now express the total emitted power by the dipole at the Stokes frequency $\omega_s = \omega_p - \omega_v$ as:

$$P_s(\omega_s) = \frac{\omega_s^4}{6\pi c^4 \epsilon_0^2} \left| \frac{\partial \alpha}{\partial Q} \right|^2 Q^2 2c\epsilon_0 |E_0|^2 \tag{2.15}$$

Which can be expressed in function of the intensity I of the incoming beam as:

$$\begin{aligned}P_s(\omega_s) &= \frac{\omega_s^4}{6\pi c^4 \epsilon_0^2} \left| \frac{\partial \alpha}{\partial Q} \right|^2 Q^2 I \\ &= \sigma I\end{aligned} \tag{2.16}$$

Where σ [m^2] is defined as the scattering cross-section. We can see that our classical model correctly predicts that the amount of Raman scattered power scales as $1/\lambda_s^4$. Therefore, shorter wavelengths will result in a higher Raman response.

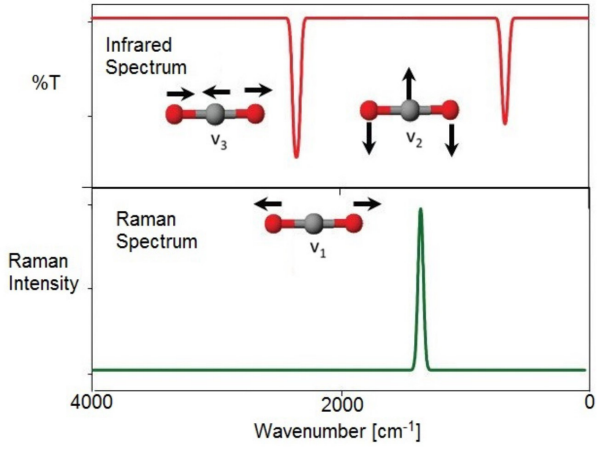


Figure 2.4: Theoretical Raman and IR spectra for a molecule of CO_2*. We can see that different vibrational modes (frequencies) are active in both techniques [28].*

Note that in experiments typically an excitation wavelength of 785 nm is preferred since it provides a sweet spot between on one hand a sufficiently short wavelength (and hence increased signal generation) and on the other hand avoiding parasitic effect such as auto-fluorescence and (bio-)sample degradation which are typically present for shorter wavelengths [29]. Additionally, we see that the scattered power scales with $|\frac{\partial\alpha}{\partial Q}|^2$ again showing the importance of a non-zero change of the polarizability along the nuclear coordinate. Furthermore, we can see that the phase of the emitted Stokes light is determined through Q and is thus dependent on the phase of the vibrational mode. Since this phase will differ between molecules the emitted Raman scattered light will be uncorrelated between the different dipoles. Spontaneous Raman scattering is therefore an incoherent Raman technique. The total scattered power for M molecules thus simply corresponds to: $P_s^{\text{total}}(\omega_s) = M\sigma I$.

It is interesting to note that Raman spectroscopy is a relatively weak sensing technique compared to alternative spectroscopic techniques such as IR absorption spectroscopy and fluorescence spectroscopy ($\sigma_{\text{Ra.}} \sim 10^{-27}[\text{cm}^2]$, $\sigma_{\text{IR.}} \sim 10^{-16}[\text{cm}^2]$, $\sigma_{\text{Fl.}} \sim 10^{-16}[\text{cm}^2]$ for rhodamine 6G [25, 30]). Therefore, typically enhanced Raman techniques are employed to further boost the signal strength like coherent Raman scattering and surface enhanced Raman scattering as will be discussed in Section 2.2 and 2.3. Raman spectroscopy however provides several inherent advantages over the other two techniques. In fluorescence spectroscopy, fluorescent labels are often needed to detect the molecule of interest [31]. However, these labels can change the intrinsic properties of the analyte and furthermore experience photobleaching (loss of the fluorescent properties). In contrast, Raman scattering

is a noninvasive technique that directly probes the molecular vibrations and does not require the use of labels. In IR absorption spectroscopy, the same molecular vibrations are probed as in Raman scattering by using IR radiation that is absorbed when the frequency of the incoming light matches the transition energy between two vibrational modes. Typically, these IR-absorption transitions occur in the mid-infrared range. A drawback of IR-absorption spectroscopy is therefore that it is challenging to probe biological samples since it suffers from the strong absorption of water in the mid-infrared range. For these reasons Raman spectroscopy is preferred over the other two techniques in this work.

2.2 Coherent Raman scattering

Coherent Raman techniques are closely related to spontaneous Raman techniques. In these nonlinear Raman techniques, the Raman active molecules are efficiently driven into resonance by the beat frequency of two incident fields (pump, Stokes) instead of one field like in the spontaneous equivalent (pump). The radiation that follows, maintains a well defined phase relationship with the driving fields leading to coherent Raman signals and thus constructive interference between the contributions of the different Raman scatterers [32,33]. In contrast, for spontaneous Raman scattering the phase of the emitted radiation from different molecules is random relative to each other (and hence it is an incoherent Raman technique). Coherent Raman techniques thus enhance the strength of the light-matter interaction thereby resulting in much stronger Raman signals.

In order to further describe these coherent Raman techniques let us first introduce the electric dipole moment $\mu(t)$ that is formed by the interaction between the electrons of a molecule or material and the electric field of the electromagnetic radiation (in this section we follow the classical theory as given in [24, 34, 35]):

$$\mu = -e\, r(t) \tag{2.17}$$

Where e is the electron charge and r(t) the displacement of the bound electrons from their equilibrium position. By adding up all the N electric dipoles per unit volume we come to the expression of the macroscopic polarization :

$$P(t) = N\, \mu(t) \tag{2.18}$$

In case the externally applied electric field is weak compared to the binding field of the electrons to their nuclei, the polarization can be expressed as being

directly proportional to the electric field such that:

$$P(t) = \epsilon_0 \chi E(t) \tag{2.19}$$

where ϵ_0 is the electric permittivity in vacuum and χ the susceptibility of the material. Hence we can see that in the limit of a weak field there is a linear relationship with the electric field. When applying stronger fields, this linear dependence is not strictly valid as the electron binding potential can no longer be assumed to be harmonic because there are also anharmonic effects that become more important. In case the anharmonic contributions are relatively small, the polarization can be described as a power series of the electric field to include these nonlinear effects:

$$P(t) = \epsilon_0[\chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t) + ...] \tag{2.20}$$

With $\chi^{(1)}, \chi^{(2)}, \chi^{(3)}$ the linear, second- and third-order susceptibility. The coherent Raman effects described here originate from the third-order susceptibility. Note that besides the Raman nonlinearity ($\chi_R^{(3)}$), which emerges from the interaction between the molecular vibrations and external fields, there are also other third-order nonlinear effects such as four-wave mixing [36] that originate from the almost instantaneous electronic non-resonant response to the external fields which we will characterize as Kerr nonlinearities ($\chi_K^{(3)}$) [34]. Second-order nonlinear effects do not play a major role in silicon (nitride) photonics since they are centrosymmetric materials for which the second-order susceptibility disappears [12].

In order to have a more intuitive interpretation of coherent Raman scattering lets us provide a classical description of the scattering process. We know that if the beat frequency of the two incoming fields (pump and Stokes) corresponds to the vibrational resonance frequency of the probed molecule it will induce oscillations of the molecular electron cloud that will result in the nuclear modes being actively driven. These modes lead to a spatially coherent modulation of the refractive index of the medium at the beat frequency. A third beam that propagates through the medium will experience this modulation in the refractive index and form sidebands that are shifted by the modulation frequency. Note that in our experiments this third beam simply corresponds to the pump beam as our coherent Raman experiments are performed in a degenerate configuration.

So now we consider the incoming light fields as:

$$\tilde{E}_i(z, t) = E_i e^{j(k_i.z - \omega_i t)} + c.c. \tag{2.21}$$

Where $\tilde{E_p}, \tilde{E_s}$ correspond to field due to the pump and Stokes beam respectively with frequencies ω_p and ω_s. Note that these frequencies are far from the vibrational resonance frequency (ω_v) such that they will not efficiently drive the nuclear mode. In contrast, we will further assume that their difference frequency $\Omega = \omega_p - \omega_s$ is close to ω_v such that it can actively drive the vibrational mode since it is in resonance. The driving force exerted by the incoming fields on the molecule along the vibrational degree of freedom Q can now be expressed as:

$$\tilde{F}(t) = \frac{1}{2}\left(\frac{\partial\alpha}{\partial Q}\right)_0 \tilde{E}^2 \tag{2.22}$$

If we now fill in $\tilde{E} = \tilde{E}_p + \tilde{E}_s$ we can see that the exerted optical force will consist of many frequency contributions (DC, $2\omega_p$, $2\omega_s$, $\omega_p + \omega_s$ and $\omega_p - \omega_s$). However, only $\omega_p - \omega_s$ will efficiently drive the nuclear mode as the other contributions are too far of resonance to have a considerable influence. Therefore, F(t) can be expressed as:

$$\tilde{F}(t) = \left(\frac{\partial\alpha}{\partial Q}\right)_0 [E_p E_s^* e^{j((k_p - k_s)z - \Omega t)} + c.c.] \tag{2.23}$$

It is interesting to note that the electron motions are coupled to the vibrational modes through a nonzero $(\frac{d\alpha}{dQ})_0$ (the basis of the Raman selection rules). Therefore, a modulation of the electron cloud at a frequency Ω by the time-varying optical force will be felt by the nuclear mode. This force (F(t)) can be linked to the nuclear displacement $\tilde{Q}(t)$ as:

$$\frac{d^2\tilde{Q}}{dt^2} + 2\gamma\frac{d\tilde{Q}}{dt} + \omega_v\tilde{Q(t)} = \frac{\tilde{F}(t)}{m} \tag{2.24}$$

Where m is the reduced mass of the nuclear oscillator, γ is the damping constant, ω_v is the resonance frequency and we again model the molecular vibration as a harmonic oscillator like in the previous section. Such that $\tilde{Q}(t)$ corresponds to:

$$\tilde{Q}(\Omega, t) = Q(\Omega)e^{j((k_p - k_s).z - \Omega t)} + c.c. \tag{2.25}$$

with the amplitude of the molecular vibration:

$$Q(\Omega) = \frac{1}{m}\left(\frac{\partial\alpha}{\partial Q}\right)_0 \frac{E_p E_s^*}{\omega_v^2 - \Omega^2 - 2j\Omega\gamma} \tag{2.26}$$

The electric dipole moment can now be expressed as:

$$\mu(t) = \alpha(t)\tilde{E}(z,t) = \left(\alpha_0 + \left(\frac{\partial\alpha}{\partial Q}\right)_0 \tilde{Q}(\Omega,t) + ...\right)\tilde{E}(z,t) \tag{2.27}$$

where we included the effect of the nuclear modes on the polarizibility. Since the polarizability is a measure of how easily the electron cloud is distorted by an electric field, the presence of the nuclear mode also needs to be accounted for. Varying nuclear coordinates lead to a variation of the adiabatic electronic potential felt by the electron and hence affects the motion of the electrons when an external field is applied. Now we can express the macroscopic polarization as:

$$\tilde{P(t)} = N\left(\alpha_0 + \left(\frac{\partial\alpha}{\partial Q}\right)_0 \tilde{Q}(\Omega,t)\right)[\tilde{E_p}(z,t) + \tilde{E_s}(z,t)] \tag{2.28}$$

We will now consider the terms proportional to $\left(\frac{\partial\alpha}{\partial Q}\right)_0$ since they describe the third-order nonlinear polarization contribution (P_{NL}) due to driven Raman mode (while the terms proportional to α_0 correspond to the linear polarization (P_L)). Such that:

$$\begin{aligned}\tilde{P_{NL}}(t) &= \tilde{P}(\omega_{as}) + \tilde{P}(\omega_{cs}) + \tilde{P}(\omega_s) + \tilde{P}(\omega_p)\\ &= P(\omega_{as})e^{j((2k_p - k_s).z - \omega_{as}t)}\\ &+ P(\omega_{cs})e^{j((2k_s - k_p).z - \omega_{cs}t)}\\ &+ P(\omega_s)e^{j(k_s.z - \omega_s t)}\\ &+ P(\omega_p)e^{j(k_p.z - \omega_p t)} + c.c.\end{aligned} \tag{2.29}$$

With $\omega_{cs} = 2\omega_s - \omega_p$ the coherent Stokes frequency and $\omega_{as} = 2\omega_p - \omega_s$ the anti-Stokes frequency. The amplitude of the polarization at the anti-Stokes frequency corresponds to:

$$P(\omega_{as}) = \frac{N}{m}\left(\frac{\partial\alpha}{\partial Q}\right)_0^2 \frac{1}{\omega_v^2 - \Omega^2 - 2j\Omega\gamma} E_p^2 E_s^* = 6\epsilon_0\chi_R(\Omega)E_p^2E_s^* \tag{2.30}$$

With the nonlinear Raman susceptibility:

$$\chi_R(\Omega) = \frac{N}{6m\epsilon_0}\left(\frac{\partial\alpha}{\partial Q}\right)_0^2 \frac{1}{\omega_v^2 - \Omega^2 - 2j\Omega\gamma} \tag{2.31}$$

which can be split into a imaginary and real component as shown in Fig. 2.5. We can see that the real and imaginary part of the susceptibility only have a strong contribution in the vicinity of the Raman resonance (ω_v). At resonance ($\Omega = \omega_v$), the susceptibility becomes purely imaginary.

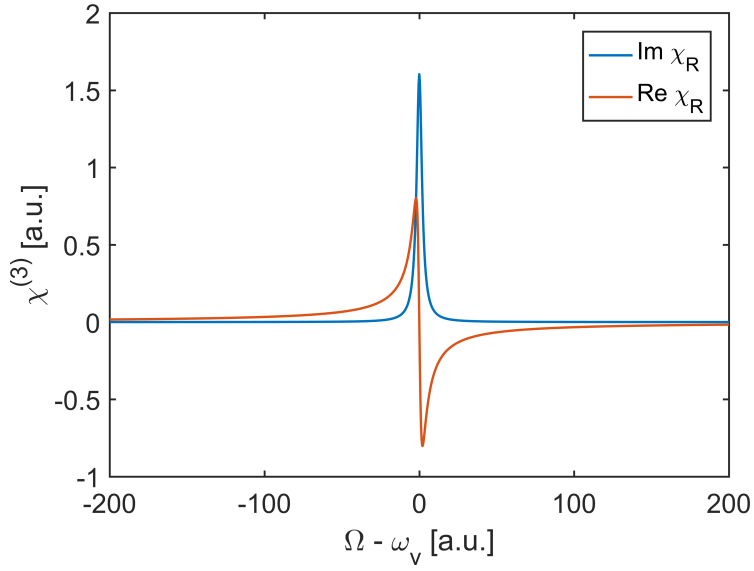


Figure 2.5: Real and imaginary part of the Raman susceptibility $\chi_R(\Omega)$. Where $\Omega = \omega_{pump} - \omega_{Stokes}$ and ω_v the Raman resonance frequency.

The other frequency components can now be written as:

$$P(\omega_{cs}) = 6\epsilon_0 \chi_R^*(\Omega) E_s^2 E_p^* \tag{2.32}$$

$$P(\omega_s) = 6\epsilon_0 \chi_R^*(\Omega) |E_p|^2 E_s \tag{2.33}$$

$$P(\omega_p) = 6\epsilon_0 \chi_R(\Omega) |E_s|^2 E_p \tag{2.34}$$

Where coherent anti-Stokes Raman scattering (CARS) originates from $P(\omega_{as})$, coherent Stokes Raman scattering (CSRS) from $P(\omega_{cs})$ and $P(\omega_s)$, $P(\omega_p)$ are responsible for the Stimulated Raman gain (SRG) and loss (SRL) respectively. We can see that each frequency component is proportional to the cube of the electric field thereby highlighting that they are third-order nonlinear processes. In the following two sections we will use these expression to describe the evolution (with distance)

of the optical fields at the (anti-)Stokes frequency by the CARS and SRS processes. CSRS will not be discusses in this work since it is less common used as CARS as it is an anti-Stokes scattering process that starts in the less populated excited vibrational states which is a disadvantage [37].

2.2.1 Coherent anti-Stokes Raman scattering

In order to come to a workable expression that describes the evolution of the anti-Stokes beam by the coherent anti-Stokes Raman scattering process, let us introduce the nonlinear wave equation for (quasi-)CW excitation (i.e. laser linewidth << linewidth Raman resonance) [34, 35, 38]:

$$\frac{d\tilde{E_i}}{dz} = j\frac{\omega_i}{2cn_i\epsilon_0}\tilde{P}_i^{NL} \tag{2.35}$$

Which links the nonlinear polarization (P_i^{NL}) to the optical field (E_i). We can see that only if $P_{NL} \neq 0$ (hence the light is propagating through a nonlinear medium) there will be variation in the electric field along the propagation direction.

Recall from Eq. 2.29 that for CARS the nonlinear polarization corresponds to:

$$\tilde{P}_{NL}(\omega_{as}, t) = P(\omega_{as})e^{j((2k_p - k_s).z - \omega_{as}t)} + c.c. \tag{2.36}$$

Which we can reorder as:

$$\tilde{P}_{NL}(\omega_{as}, t) = P(\omega_{as})e^{j((2k_p - k_s - k_{as}).z)}e^{j(k_{as}z - \omega_{as}t)} + c.c. \tag{2.37}$$

Such that it can be filled in eq. 2.35 and we get:

$$\frac{dE_{as}}{dz} = j\frac{\omega_{as}}{2cn_{as}\epsilon_0}P(\omega_{as})e^{j((2k_p - k_s - k_{as}).z)} \tag{2.38}$$

which (using Eq. 2.30) can finally be written as:

$$\frac{dE_{as}}{dz} = j\frac{3\omega_{as}}{cn_{as}}\chi_R(\Omega)E_p^2E_s^*e^{j(\Delta k.z)} = \kappa_{as}E_s^* \tag{2.39}$$

With the phase mismatching term $\Delta k = (\mathbf{2_p} - \mathbf{k_s} - \mathbf{k_p}).\hat{z}$ and the nonlinear coupling efficient κ_{as} defined as:

$$\kappa_{as} = j\frac{3\omega_{as}}{cn_{as}}\chi_R(\Omega)E_p^2e^{j(\Delta k.z)} \tag{2.40}$$

Note that in CARS efficient signal generation at the anti-Stokes frequency only occurs when $\Delta k \approx 0$. In case $\Delta k \neq 0$, the spatial dependence of the nonlinear source term (right side eq. 2.39) differs of that of the propagating wave at the anti-Stokes frequency. Therefore, depending on the location in the nonlinear medium, any newly light generated by the nonlinear interaction between the dipole and the driving fields can be in/out of phase with the already propagating wave at the anti-Stokes frequency. The power at ω_{as} will hence increase or decrease depending on the exact location in the medium. In contrast, when there is phase matching, the different contributions in the nonlinear medium will add constructively and the propagating wave will experience gain.

If we now assume that the Pump and Stokes beam are undepleted by the CARS interaction over a distance L we can write:

$$E_{as}(L) = j\frac{3\omega_{as}}{cn_{as}}\chi_R(\Omega)E_p^2E_s^*L\frac{(e^{j(\Delta k.L)}-1)}{j\Delta k} \tag{2.41}$$

and using the fact that the intensity I_i corresponds to:

$$I_i = 2n_i\epsilon_0 cE_i^*E_i = 2n_i\epsilon_0 c|E_i|^2 \tag{2.42}$$

such that the generated intensity at the anti-Stokes frequency corresponds to:

$$I_{as}(L) = \frac{9\omega_{as}^2}{4n_p^2n_{as}n_sc^4\epsilon_0^2}|\chi_R|^2I_p^2I_sL^2\mathrm{sinc}^2(\frac{\Delta kL}{2}) \tag{2.43}$$

Where we again see that the CARS intensity is strongly dependent on the amount of phase mismatch through the term $\mathrm{sinc}^2(\Delta kL/2)$. Besides that, there is also a quadratic dependence on the pump intensity that can be understood from the energy level diagram shown in Fig. 2.7 as there are two pump photons required for the CARS process. Moreover, since the CARS intensity scales quadratic with the susceptibility it also scales quadratic with the analyte concentration and is therefore more sensitive to majority constituents. Furthermore, through this susceptibility it is also prone to a non-resonant background contribution [39]. Up to now we have not considered the Kerr nonlinearity ($\chi_K^{(3)}$) which is also a relevant nonlinear source besides the Raman nonlinearity ($\chi_R^{(3)}$). Since CARS is a parametric process (no energy is transferred to the probed sample but between optical fields, see Fig. 2.7) other four-wave mixing (FWM) processes can also occur without the presence of the resonant molecules. Hence both the Kerr-induced and Raman-induced FWM process will generate a frequency contribution at ω_{as}. In contrast to CARS, the Kerr-induced FWM process is a nonresonant electronic nonlinearity that originates

from the response of the bound electrons to an external optical field [34]. Since the Kerr response only involves virtual processes (see Fig. 2.7) it will be an almost instantaneous process where CARS will have a more delayed resonant response since it involves the excitation of molecular vibrations. Therefore, we also have the include the influence of $(\chi_K^{(3)})$ on I_{as} such that:

$$I_{as} \propto |\chi_R^{(3)}(\Omega) + \chi_K^{(3)}|^2 = |\chi_R^{(3)}|^2 + |\chi_K^{(3)}|^2 + 2\chi_K^{(3)}\mathrm{Re}\{\chi_R^{(3)}\} \tag{2.44}$$

Where $|\chi_K^{(3)}|^2$ is a constant over the considered frequency range since it is the nonresonant (NR) background contribution and $|\chi_R^{(3)}|^2$ has a lorentzian line shape. We can see that there is a third contribution $(2\chi_K^{(3)}\mathrm{Re}\{\chi_R^{(3)}\})$ which is responsible for the asymmetric distorting of the CARS spectra compared to the spontaneous Raman and SRS spectra. The presence of this NR background becomes particular challenging if $\chi_K^{(3)} >> \chi_R^{(3)}$ which is often the case for diluted analytes such that:

$$I_{as} \propto |\chi_K^{(3)}|^2 + 2\chi_K^{(3)}\mathrm{Re}\{\chi_R^{(3)}\} \tag{2.45}$$

And I_{as} becomes proportional to $\mathrm{Re}\{\chi_R^{(3)}\}$ (instead of $\mathrm{Im}\{\chi_R^{(3)}\}$ like in SRS). This will lead to a dispersive lineshape on top of the non-resonant background (see Fig. 2.6) which makes it difficult to interpret the CARS spectra without additional data-analysis [40–43].

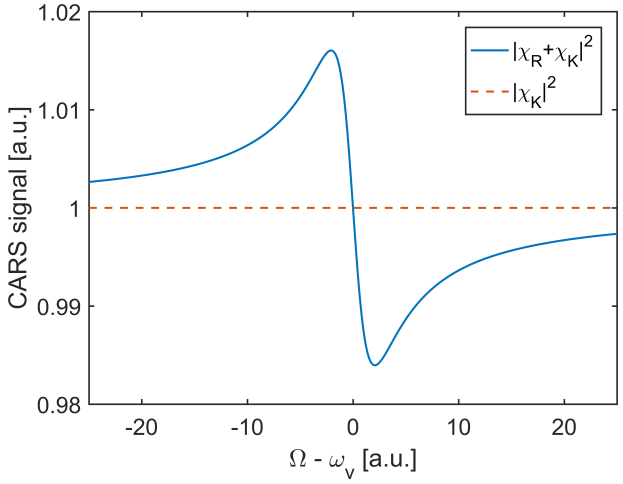


Figure 2.6: Measured CARS spectrum in presence of a strong nonresonant background contribution $(\chi_K^{(3)})$.

Note that through the Kerr nonlinearity there is also an additional contribution to the phase matching term of Eq. 2.43:

$$I_{as} \propto \mathrm{sinc}^2(\frac{\kappa L}{2}) \tag{2.46}$$

Where the net phase mismatch term κ corresponds to [38]:

$$\kappa = \Delta k + 2\gamma_{NL} P_{pump} \tag{2.47}$$

γ is the nonlinear parameter and $k_{NL} = 2\gamma_{NL} P_{pump}$ accounts for the contribution of self-phase modulation and cross-phase modulation to the phase mismatch. In general, a large contribution of the Kerr nonlinearity compared to the signal generation is thus to be avoided since it will make it challenging the separate the relevant Raman information from the Kerr-induced contributions [24].

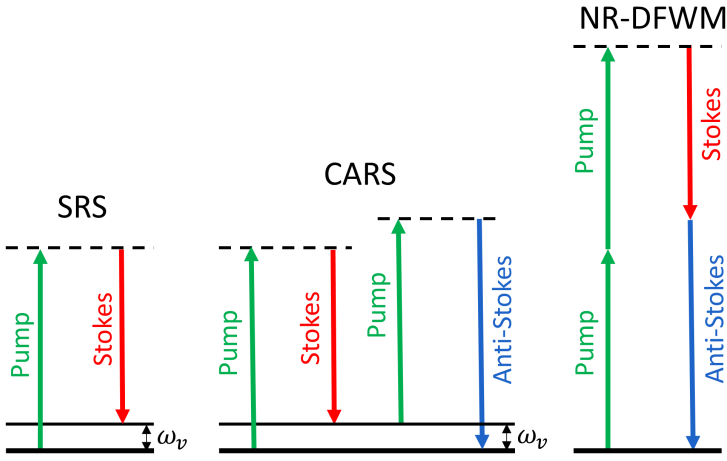


Figure 2.7: Energy level diagram of stimulated Raman scattering (SRS), coherent anti-Stokes Raman scattering (CARS) and the nonresonant degenerate four-wave mixing process (NR-DFWM). The vibrational (virtual) energy states correspond to the full (dotted) lines.

2.2.2 Stimulated Raman scattering

Let us now derive a workable expression for SRS. If we use Eq. 2.29, where we derived that the nonlinear polarization at Stokes frequency corresponds to:

$$\tilde{P}_{NL}(\omega_s, t) = P(\omega_s) e^{j(k_s.z - \omega_s t)} + \text{c.c.} \tag{2.48}$$

Such that eq. 2.35 becomes:

$$\frac{dE_s}{dz} = j\frac{\omega_s}{2cn_s\epsilon_0}P(\omega_s) \tag{2.49}$$

and by using eq. 2.33 we now get:

$$\frac{dE_s}{dz} = j\frac{3\omega_s}{cn_s}\chi_R^*(\Omega)|E_p|^2E_s = -\alpha_s E_s \tag{2.50}$$

Where the gain coefficient of the Stokes wave is defined as:

$$\alpha_s = -j\frac{3\omega_s}{cn_s}\chi_R^*(\Omega)|E_p|^2 \tag{2.51}$$

Since we know that at resonance ($\Omega = \omega_v$), $\chi_R^*(\Omega)$ is a negative imaginary number we can actually see that the Stokes beam will experience gain since α_s will be a negative real number. Therefore, the Stokes beam will grow with distance by the nonlinear Raman interaction and this phenomena is therefore called Stimulated Raman gain (SRG). In contrast, for the pump beam $\chi_R(\Omega)$ at resonance is a positive imaginary number such that the propagating pump beam will experience loss and this phenomena is thus labeled Stimulated Raman loss (SRL).

$$\frac{dE_p}{dz} = j\frac{3\omega_p}{cn_p}\chi_R(\Omega)|E_s|^2E_p = -\alpha_p E_p \tag{2.52}$$

$$\alpha_p = -j\frac{3\omega_p}{cn_p}\chi_R(\Omega)|E_s|^2 \tag{2.53}$$

The difference between the pump and Stokes beam can also be understood from the energy level diagram of Fig. 2.7. Since a pump photon is converted to a Stokes photon through the SRS process the propagating Stokes beam will experience gain while the pump beam will experience loss. Furthermore, the newly generated Stokes photon is generated at a field mode that is already occupied by the incoming Stokes beam such that it will have the same frequency, phase and polarization [24]. Therefore SRS will be automatically phase matched.

Eq. 2.50 can now be solved if we assume that the pump will not be depleted by the SRS process over a distance L. The electric field of the Stokes beam after a distance L then becomes:

$$E_s(L) = E_s(0)\, e^{j\frac{3\omega_s}{cn_s}\chi_R^*(\Omega)|E_p|^2L} \tag{2.54}$$

Since SRS is a weak process this equation can be further approximated to:

$$E_s(L) = E_s(0)(1 + j\frac{3\omega_s}{cn_s}\chi_R^*(\Omega)|E_p|^2 L) \tag{2.55}$$

The Stokes intensity $I_s(L)$ now corresponds to:

$$\begin{aligned} I_s(L) &= 2n_i\epsilon_0 c E_s^* E_s \\ &= I_s(0)\left(1 + \frac{6\omega_s}{n_s c}|E_p|^2 L\ \mathrm{Im}\{\chi_R\} + \left(\frac{3\omega_s}{n_s c}|E_p|^2 L\right)^2 |\chi_R|^2\right) \end{aligned} \tag{2.56}$$

Since the Raman interaction is weak $\mathrm{Im}\{\chi_R\} >> |\chi_R|^2$, the third term can be neglected. The intensity increase of the Stokes beam by the nonlinear Raman interaction can then be expressed as:

$$\Delta I_s = I_s(L) - I_s(0) = \frac{3\omega_s \mathrm{Im}\{\chi_R\}}{n_s n_p c^2 \epsilon_0} I_s I_p L = g_I I_s I_p L \tag{2.57}$$

Where g_I is the stimulated Raman gain [m/W]. We can see that SRS has a linear dependence on the length, pump and Stokes intensity. Furthermore, since ΔI_s is depended on $\mathrm{Im}\{\chi_R\}$ the SRS spectra will have the same lorentzian lineshape as the spontaneous Raman spectra [24]. Therefore, the SRS spectra can easily be compared to spontaneous reference spectra without the need for extra data-analysis unlike the CARS spectra. Moreover, through its linear dependence on $\mathrm{Im}\{\chi_R\}$, SRS has a linear dependence on the concentration. Therefore, SRS still generates reasonable signal levels for low analyte concentrations. In contrast, for CARS the generated signal levels will drop more drastically because of its quadratic dependence. SRS also does not suffer from a nonresonant background contribution like in CARS since SRS is a energy transfer process between the analyte and the optical field (see Fig.2.7). SRS(-like) processes can thus only occur if the beating frequency of the pump and Stokes beams matches the energy difference between the vibrational energy levels. SRS is also automatically phase matched since the intensity gain and loss is generated on top of the incoming Stokes and pump beams. One disadvantage of SRS over CARS is that its detection is more challenging since the intensity increase ΔI_s is usually a couple orders of magnitude smaller than input Stokes beam $I_s(0)$. Therefore, a more complex lock-in detection scheme is required and special care is needed to suppress the shot noise contribution of the Stokes beam in detection.

2.3 Surface-enhanced Raman scattering

In surface-enhanced Raman scattering (SERS) the inherently weak Raman scattering process is enhanced by many order of magnitude by the highly localized field in close vicinity ($\sim$ nm) to the surface of a plasmonic nanostructure. In order for molecules to experience this enhancement they thus need to be in close proximity to the surface (explain the SE in the term SERS). This strong electric field near the metallic surfaces originates from localized surface plasmon resonances (or propagating surface plasmon polaritons) that are optically excited and can be seen as the collective (resonant) oscillations of the free electrons at the frequency of the incident light. The incoming light wave will drive these motions of the electrons which on their turn will generate an electromagnetic field . In this section, we will discuss the theory of plasmonic physics making use of classical theory. We will give more details about the physics of localized surface plasmon resonances (LSPR) and propagating surface plasmon polariton (PSPP) which are responsible for the enhancement of Raman scattering near the surface of metallic nanoparticles (smaller than the wavelength of light) and longer metallic nanostructures respectively. Finally, we give a short discussion about the enhancement factors playing a role in SERS. In this section we largely follow the classical theory as given in [30, 44, 45]

2.3.1 Plasmon physics

Let us first provide the definition of a surface plasmon polarition. A plasmon is a quantized oscillations of the electron density in a metal, while at the metalic surface a surface plasmon polarition can be formed which is a quasi-particle created by the coupling of a photon and a plasmon.

Let us now start with a metal-dielectric interface (see Fig. 2.8) where the plasmon mode is propagating along the x-direction such that its propagation constant K_{sp} is orientated along the interface. For TM polarized light ((E_x,0,E_z) and (0, H_y,0)) the dispersion relation is given by:

$$K_{sp} = \frac{\omega}{c}\sqrt{\frac{\epsilon_D \epsilon_M}{\epsilon_D + \epsilon_M}} \tag{2.58}$$

According to the generalized Drude theory the permitivity of a metal with a perfectly free electron gas corresponds to:

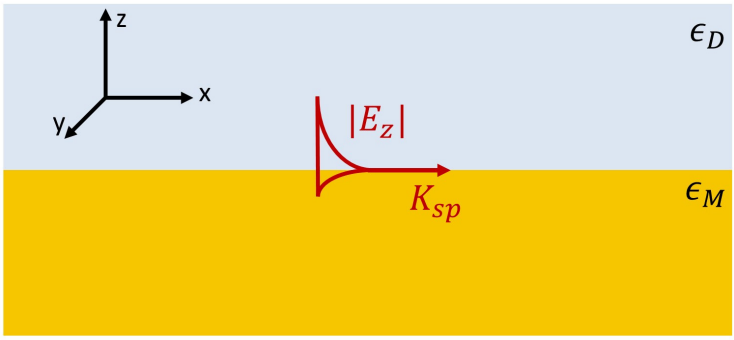


Figure 2.8: Schematic representation of the field of a surface plasmon polariton (SPP) at the metal-dielectric interface. ϵ_D and ϵ_M are the dielectric and metal permittivity respectively.

$$\epsilon_M(\omega) = 1 - \frac{\omega_{pl.}^2}{\omega^2} \tag{2.59}$$

Where $\omega_{pl.}$ is the bulk plasma frequency corresponding to:

$$\omega_{pl}^2 = \frac{n_e e^2}{\epsilon_0 m_e} \tag{2.60}$$

We can thus see that the plasma frequency depends on the electron charge e, the density n_e and the effective mass m_e. The dispersion relation for a metal-air interface ($\epsilon_D = 1$) can now be expressed as:

$$K_{sp}(\omega) = \frac{\omega}{c}\sqrt{\frac{1 - \frac{\omega_{pl.}^2}{\omega^2}}{2 - \frac{\omega_{pl.}^2}{\omega^2}}} \tag{2.61}$$

If we now plot the solution to Eq. 2.61 in Fig. 2.9, we can see that there are two types of solutions. In the $0 < \omega < 0.707\omega_{pl.}$ frequency range the solution is a surface plasmon mode. The other solution in the $\omega > \omega_{pl.}$ frequency range corresponds to the bulk plasmon mode. In the $0.707\omega_{pl.} < \omega < \omega_{pl.}$ frequency range there are no solutions. From the dispersion curve we can see that for low K_{sp}, the SPP acts like a photon but for higher K_{sp} the dispersion curve bend away from the ligth line and reach the surface plasmon frequency ($\omega_{pl.}/\sqrt{2}$). Since the disperison curve of the SPP is to the right of the light line, the SPP has a shorter wavelength than the free-space light.

The field decay length of the SPP now corresponds to :

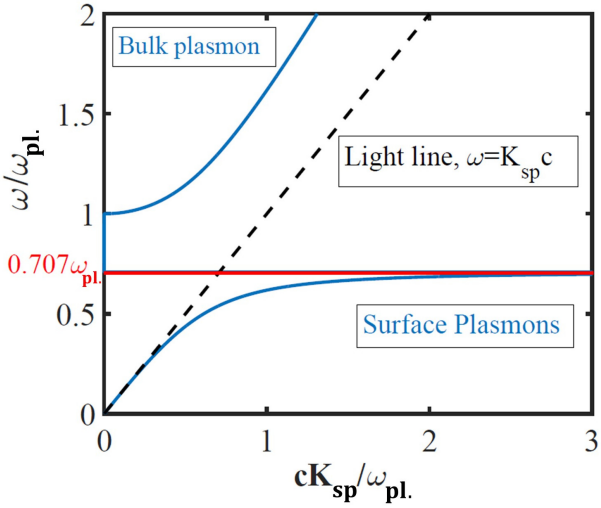


Figure 2.9: Dispersion graph of the surface plasmon polariton for a metal-dielectric interface [45].

$$L_i = \frac{c}{\omega}\sqrt{\frac{\epsilon_D + \epsilon_m}{\epsilon_i^2}} \tag{2.62}$$

Where it can either correspond to the metal or the dielectric (air in the case considered here). Since the dielectric constant of a metal is typically larger than that of the dielectric, the decay length will be shorter for the metal than for the dielectric. As an example, at 600 nm for a silver-air interface L_{air} = 390 nm and L_{silver} = 24 nm. For large K_{sp}, $L_i \propto 1/K_{sp}$ leading to highly concentrated fields near the surface.

From Fig. 2.9 we can see that the light line does not intersect with the surface plasmon polariton line. Thus light incident on the metal surface in a arbitrary direction cannot excite surface plasmons since there is a momentum mismatch between the incoming light and the plasmon. Therefore, special configurations are needed such as using a prism coupler [46] or a periodic corrugation [47]. In the Kretschmann configuration, a prism is used to excite the surface plasmon where one side of the prism is coated with a thin metal layer. The metal-dielectric interface is then illuminated at an angle greater than the total internal reflection angle. The coupling of the incoming light with the surface plasmon occurs when the in plane component of the wavevector matches with K_{sp}:

$$\frac{\omega}{c}\sqrt{\epsilon_p}\sin(\theta_{sp}) = K_{sp} \tag{2.63}$$

At resonance ($\theta = \theta_{sp}$) a sharp minimum will be observed for the reflectively. The glass prism thus acts as a medium through which the wavenumber (and hence momentum) of the incoming light beam is increased to match that of the surface plasmon.

So let us now summarize a bit, surface plasmons polaritons are propagating electromagnetic waves at the metal-dielectric interface coupled to the electron plasma of the metal. They have high field intensities and a compact optical confinement near the surface since their wavelength is shorter than that of the incident optical field. Alternatively to SPP, there are also localized surface plasmon resonances(LSPR). These LSPR are non-propagating and are excited by the interaction between sub-wavelength conductive nanoparticles and the electromagnetic field. The electric field then induces the collective oscillations of the electron cloud of the particles. Since the surface of the particles is spatially confined, a restoring force will act on the electrons that leads to a resonance of the electron cloud. The resonance frequency will depend on the geometry of the particle and its material and that of the environment. One important thing to note is that LSPR is a stationary mode that is confined to a particle and occurs at a single frequency. In contrast, SPP are propagating modes along a metal-dielectric interface and occur in a frequency range as shown in Fig. 2.9.

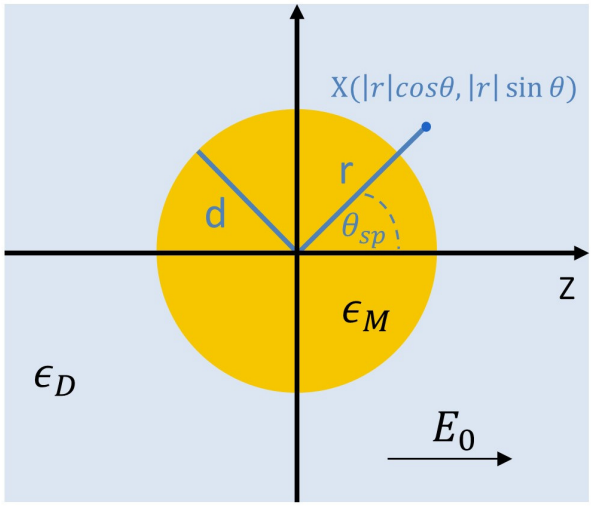


Figure 2.10: Metal particle placed in a electrostatic field (E_0), with ϵ_D the permittivity of the surrounding dielectric.

Let us now assume a conducting nano-sphere (with size d) that is smaller than the wavelength of light $d << \lambda$. The sphere is surrounded by a dielectric medium (ϵ_D) without dispersion and loss (see Fig. 2.10). Furthermore, we assume that the EM wave has a constant phase over the nanoparticle. In extracting the spatial field distribution we can thus apply a quasi-static approximation. For $\nabla \times E = 0$, the electric field can be extracted from the gradient of the electric potential V_E:

$$E = -\nabla V_E \tag{2.64}$$

From Gauss law, we can now write:

$$\nabla E = \nabla(-\nabla V_E) = -\nabla^2 V_E = 0 \tag{2.65}$$

The general solution of Eq. 2.65 now corresponds to [48]:

$$V_E(r,\theta) = \sum_{n=0}^{\infty} [A_n r^n + B_n r^{-(n+1)}] P_n(\cos\theta) \tag{2.66}$$

With $P_n(\cos\theta)$ the Legendre Polynomials of order n and θ the angle between the position vector r and the z-axis. Now since $V_E(r,\theta)$ must remain finite at the origin, Eq. 2.66 can be split in a solution inside the sphere ($r < d$):

$$V_{E,in}(r,\theta) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos\theta) \tag{2.67}$$

and outside the sphere ($r > d$):

$$V_{E,out}(r,\theta) = \sum_{n=0}^{\infty} [B_n r^n + C_n r^{-(n+1)}] P_n(\cos\theta) \tag{2.68}$$

We can now apply the boundary conditions to determine B_n such that for $r \to \infty$, $B_1 = -E_0$ and $B_n = 0$ for $n \neq 1$. Furthermore, at r=d the tangential component of the electric field demands that:

$$\frac{1}{d}\frac{\partial V_{E,in}}{\partial \theta} = \frac{1}{d}\frac{\partial V_{E,out}}{\partial \theta} \tag{2.69}$$

and the continuity of the normal components of the displacement field that:

$$\epsilon_0\epsilon_M \frac{\partial V_{E,in}}{\partial r} = \epsilon_0\epsilon_D \frac{\partial V_{E,out}}{\partial r} \tag{2.70}$$

Such that $A_n, C_n = 0$ for $n \neq 1$ and for n=1 V_E corresponds to [48–50]:

$$V_{E,in}(r, \theta) = -\frac{3\epsilon_D}{\epsilon_M + 2\epsilon_D} E_0 r(\cos\theta) \tag{2.71}$$

$$V_{E,out}(r, \theta) = -E_0 r(\cos\theta) + \frac{\epsilon_M - \epsilon_D}{\epsilon_M + 2\epsilon_D} E_0 d^3 \frac{\cos\theta}{r^2} \tag{2.72}$$

Where it is interesting to note that the second term for the field outside the sphere corresponds the field of the dipole located at the origin. The response of the particle to the static electric field is thus an induced dipole moment. By combining 2.64, 2.71 and 2.72 it can now be seen that the internal and external dipolar field (E_{in} and E_{out}) show a resonance effect where the following condition applies:

$$\mathbb{R}(\epsilon_M) = -2\epsilon_D \tag{2.73}$$

For a metal who follows the Drude model the frequency at which this field enhancement due to the plasmon resonance occurs corresponds to:

$$\omega_{LSP} = \omega_{pl.} \sqrt{\frac{1}{1 + 2\epsilon_D}} \tag{2.74}$$

Where the resonant mode of the nanoparticle is called the dipole localized surface plasmon mode. We can see that ω_{LSP} is depended on the surrounding material and will experience a red-shit for increasing ϵ_D.

Finally, let us have a look at the field enhancement provided by the metal particle. We know that at resonance the dipole mode of the spherical particle is efficiently excited. If we now consider an isolated gold particle such that only the dipole moment can be coupled to the external EM radiation and no higher order modes can be excited because of their vanishing dipole moments. Then the total field enhancement $|E_o|$ of the dipole mode over the electric field of a tightly focused Gaussian beam $|E_G|$ can be expressed as [50]:

$$
\begin{aligned}
F &= \frac{|E_o|}{|E_G|}\left(\frac{d}{d+x}\right)^3 \\
&= \frac{\sqrt{2}}{\sqrt{[Q^{-1}+2\chi_d^3/3\epsilon_D]^2+\delta^2}}\left(\frac{\chi_d}{\chi+\chi_d}\right)^3
\end{aligned}
\tag{2.75}
$$

Where Q is the quality factor which corresponds to $Q = \frac{\omega}{\gamma}$ with γ the non-radiative decay rate. δ is the normalized excitation detuning and χ, χ_d are the normalized radius and distance from the sphere respectively ($\chi_d = 2\pi d/\lambda_d$). Gold particles have a $Q \sim 10$ and therefore their field enhancement corresponds to $F \approx 14$ at resonance near the surface. Q is larger when the non-radiative decay rate γ is small. Gold and silver have a large Q in the NIR range with silver the best plasmonic material in the visible (see Fig. 2.11). However, silver quickly oxidizes in ambient atmosphere and therefore gold is usually preferred for sensing. Note that discussion provided above is only valid for small particles. For larger particles and complex geometries the quasi-static approximation is no longer valid and geometrical parameters need to be accounted for to determine F and ω_{LSP} [51].

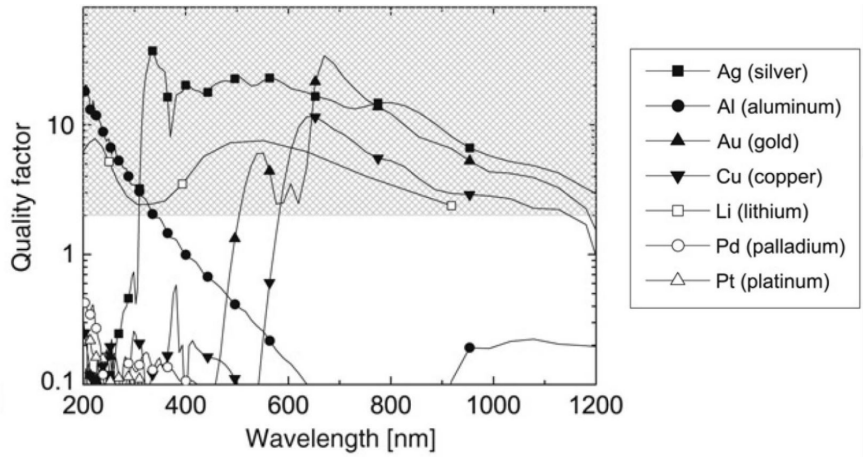


Figure 2.11: Quality factor of different metal nanoparticles interfaced to air [30].

2.3.2 SERS enhancement mechanism

Let us now consider a molecule in vacuum (with angular vibrational frequency ω_s) for which an external EM field (frequency ω_0) will induce a dipole $\tilde{p}$ (see Eq. 2.27) [25]:

$$\tilde{\mu}(\omega_0 + \omega_s) = \alpha_0(\omega_0, \omega_0 + \omega_s)\tilde{E}_0(\omega_0) \tag{2.76}$$

If we now place the molecule in close vicinity of a plasmonic structure where the surface plasmon is efficiently excited in the selected wavelength range, then the induced dipole becomes:

$$\tilde{\mu}_m(\omega_0 + \omega_s) = \alpha_m(\omega_0, \omega_0 + \omega_s)\tilde{E}_m(\omega_0) \tag{2.77}$$

Where the presence of the plasmonic structure increases the electric field intensity at the location of the molecule, changes the electron structure and hence the polarizability or vibrational frequency of the molecule and increases the local density of states which will increase the probability of spontaneous emission. For these reasons the power emitted by the dipole will increase. Furthermore, from the optical reciprocity theorem it is shown that this enhanced emission is proportional to the electric field intensity [52]. Since the power radiated by the dipole (P) is proportional to $|\tilde{\mu}|^2$ the ratio of the total Stokes power scattered by the dipole near the plasmonic structure compared to vacuum corresponds to:

$$EF = \frac{P_m}{P_O} = \overbrace{\left|\frac{\alpha_m}{\alpha_0}\right|^2}^{\text{Chemical factor}} \times \overbrace{\left|\frac{\tilde{E}_m(\omega_0)}{\tilde{E}_0(\omega_0)}\right|^2}^{\text{Field enhancement}} \times \overbrace{\left|\frac{\tilde{E}_m(\omega_0 + \omega_s)}{\tilde{E}_0(\omega_0 + \omega_s)}\right|^2}^{\text{Radiation enhancement}} \tag{2.78}$$

where we can see that there is a chemical factor which is present because the polarizability of a molecule can change due to the chemical binding of the molecule and the plasmonic structure. This effect is called chemical enhancement and the molecule may experience a moderate enhancement to a 100-fold because of it [53]. The second and third term of Eq. 2.78 are usually referred to as the electromagnetic enhancement and are the main contributors to the SERS enhancement. The EF is reported on different positions at the plasmonic structure depending on the application. When talking about single molecule detection usually the EF at the hotspot (maximum electric field point) is reported. For a monolayer absorbed on the surface of the plasmonic structure the average EF is used that averages the enhanced field points over the surface (see Fig. 2.12). Since the frequency difference between the Stokes and pump is usually relatively small compared to the FWM of the plasmonic resonance, the electromagnetic enhancement factor is often written as:

$$EF_{SERS/Sp.R} = \left|\frac{\alpha_m}{\alpha_0}\right|^2 \times \left|\frac{\tilde{E}_m(\omega_0)}{\tilde{E}_0(\omega_0)}\right|^4 \tag{2.79}$$

And we can see that the SERS enhancement scales to the fourth power with the electric field enhancement [54]. Note that SERS itself is however a linear process since it scales with I_{pump}.

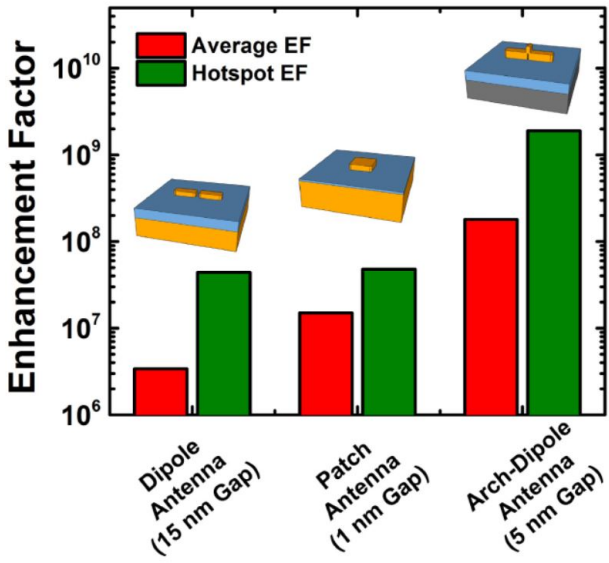


Figure 2.12: Theoretical hotspot and average enhancement factor of SERS for three different plasmonic structures [55].

In order to further increase the molecular detection sensitivity it is also possible to combine plasmonic enhancement (SERS) together with coherent Raman scattering (CARS, SRS). In this way the generated signal provided by coherently driving the molecular vibration at resonance can be further enhanced by the local fields of the excited plasmonic modes (assuming that the input and generated frequencies in CARS and SRS are in resonance with the collective modes of the plasmonic structures). If we consider SECARS, then the electromagnetic EF factor compared to CARS corresponds to [56, 57]:

$$\begin{aligned} EF_{SECARS/CARS} &= \left|\frac{\tilde{E}_m(\omega_p)}{\tilde{E}_0(\omega_p)}\right|^4 \times \left|\frac{\tilde{E}_m(\omega_s)}{\tilde{E}_0(\omega_s)}\right|^2 \times \left|\frac{\tilde{E}_m(\omega_{as})}{\tilde{E}_0(\omega_{as})}\right|^2 \\ &= |g_p|^4 |g_s|^2 |g_{as}|^2 \end{aligned} \tag{2.80}$$

where we see that strong enhancement of SECARS over CARS is due to the eight power dependence. Furthermore, we can also see that there is a additional enhancement of $\sim |g_p|^2|g_{as}|^2$ of SECARS over SERS. Theoretically, $EF_{SECARS/CARS}$

and $EF_{SECARS/SERS}$ in hotspots can reach values of $\sim 10^8 - 10^{24}$ and $10^{10} - 10^{18}$ respectively [58]. In practice, single molecule detection has experimentally been demonstrated together with an enhancement of 11 orders of magnitude over spontaneous Raman scattering [56]. Analogously, the enhancement factor for SESRS can be expressed as [32]:

$$EF_{SESRS/SRS} = \left| \frac{\tilde{E}_m(\omega_p)}{\tilde{E}_0(\omega_p)} \right|^2 \times \left(\frac{\tilde{E}_m(\omega_s)}{\tilde{E}_0(\omega_s)} \right)^2 \tag{2.81}$$

Where intensity enhancements of $\sim 10^4 - 10^6$ compared to normal SRS have been reported [59, 60]. Furthermore, detection sensitivity up to the single molecule level has recently been demonstrated [61]. A more detailed discussion about SECARS and SESRS will follow in Chapter 4 .

3

Waveguide-based Raman spectroscopy

In this chapter I will firstly introduce the concept of waveguide-based Raman spectroscopy on nanophotonic chips (Section 3.1). Afterwards, I will review the general formulas to calculate the generated Raman signal for waveguide-enhanced (stimulated) Raman spectroscopy (which we will use in Chapter 4) (Section 3.2). Finally, I give an overview of the experimental work that has already been done on nanophotonic waveguide-enhanced Raman spectroscopy (NWERS) together with its remaining challenges (Section 3.3). In Chapter 4 we will then go into detail on on-chip SERS, SRS, SE-SRS and SE-CARS as they (partially) try to solve the issues of NWERS, namely the strong photon background contribution and weak signal generation.

3.1 Raman spectroscopy via photonic integrated circuits

As discussed in Chapter 1, Raman spectroscopy typically relies on bulky and expensive instrumentation to overcome the inherently weak nature of Raman scattering. A typical Raman spectroscopic system consists of a (confocal) microscope, a high power laser and deeply-cooled CCD cameras. Since a Raman microscope suffers from small interaction volumes and poor signal collection, their conversion

efficiencies (ratio of the generated Stokes power to pump power) are typically poor. Furthermore, the high cost and bulky instrumentation limits the broader application of Raman spectroscopy outside the lab environment. To overcome these limitations, waveguide-based Raman spectroscopy has been proposed where the evanescent field of a guided waveguide mode excites and collects the Raman signal of an analyte of interest [62]. Optical waveguides consist of a high index core surrounding by a low index cladding region (which contains the analyte). The high index contrast leads to a reduced mode area and a large enhancement of the evanescent field near the waveguide surface. Furthermore, it enslaves the optical mode to the waveguiding structure such that its interaction length can be greatly enhanced beyond the diffraction limit ($\sim$cm). In [14] it is shown that a single-mode waveguide enhances the Raman signal by a factor 10-1000 per cm compared to a confocal microscope system.

Another major advantage of on-chip Raman sensors is that they are CMOS-compatible i.e. they are manufacturable by the standard process technology in a state-of-the-art CMOS fab [12]. This allows for the fabrication of cheap, mass producible Raman sensors with a small footprint (i.e. meaning more chips per wafer). Currently, the most prominent photonic integrated circuits (PIC) platform is the silicon-on-insulator platform (SOI) since it is mature and provides high yields, reproducibility and robustness [63]. However, in this work silicon nitride is preferred over silicon as the waveguiding material since (unlike silicon) it is transparent for wavelengths below 1.1μm (till at least 500 nm). This corresponds to the relevant wavelength range for Raman spectroscopy due to the reduced photo-damage, (aqueous) absorption and fluorescence of biological samples [14]. Furthermore, the pump wavelength at 785 nm allows for a relative strong signal generation (because of the λ^{-4} dependence of Raman scattering) and low-cost, high performing silicon-based detectors are available in the 800 nm-1000 nm wavelength range (unfortunately it is this strong absorption that limits the use of silicon as the waveguiding material in this wavelength range). Just like the silicon-on-insulator technology, Si_3N_4 is also CMOS-compatible allowing for the fabrication in fabs that use manufacturing technology that already exists for electronic integrated circuits. Silicon nitride provides a moderate index contrast (compared to silica), while silicon has a very high index contrast ($\Delta n_{SiN} = 0.55$ vs. $\Delta n_{Si} = 2.25$). This leads to a less tight confinement of the mode (and hence a reduced enhancement of the evanescent field and larger circuits footprints) but this effect is largely offset by the use of shorter wavelengths which leads to a comparable size of the single-mode waveguides of both materials [25]. Furthermore, a less high index contrast leads to reduced scattering loss and thus a reduced propagating loss (which is on the order of 2 – 3 dB/cm for Si_3N_4 waveguides [12, 64]). A range of high-performance photonic components have been developed on the Si_3N_4-platform such as: spectrometers [18, 65, 66], lasers [67, 68], spectral filters necessary to

remove the strong excitation radiation [17] and Raman sensors that boost the Raman response (as will be discussed in Section 3.3). These open the door to a complete Raman spectroscopic system on a single chip.

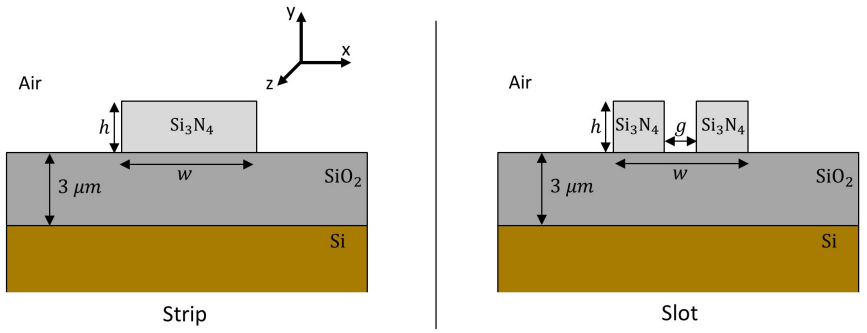


Figure 3.1: Air-cladded strip and slot waveguides sitting on top of a thick buried oxide layer [BOX]. Typical waveguide dimensions correspond to w × h: 700 nm × 220 nm (with a gap g: 150nm).

In this thesis we will use two types of waveguides: strip waveguides and slot waveguides (i.e. a double strip waveguide, see Fig. 3.1). Strip and rib waveguides are the most widely used waveguide structures because they display low propagation loss, tolerance to fabrication errors and efficient mode coupling [45, 69]. The waveguide core is made of silicon nitride, which is deposited on top of a thick silica layer (often referred to as the buried oxide layer [BOX]). Waveguides are defined using reactive ion etching and deep-UV lithography [13]. Typical layer thicknesses of the silicon nitride layer correspond to 150nm, 220nm or 300 nm. Waveguide widths typically vary from 300 nm-1,5 μm depending on the application and wavelength. The BOX layer needs to be $\sim$ 3μm thick in order to prevent leakage of the waveguide mode in the silicon substrate. Both the slot and strip waveguide can support a multiple set of discrete modes. We can separate these modes into two categories. The modes of the first category are called TE-modes and refer to those modes where the electric field is mainly polarized along the horizontal-direction (x-axis). The modes of the second category are called TM-modes and refer to those modes where the electric field is mainly polarized along the vertical-direction (y-axis). Fig 3.2 shows an example for both types of modes in a strip waveguide. Note that the guided modes in this strip waveguide are not strictly TE-or TM-polarized but are hybrid modes. The guided modes are however nearly linear and therefore they are referred to as quasi-TE and quasi-TM modes. For simplicity we will refer to them as TE-and TM-modes for the remainder of this thesis. In the rest of this chapter we will first provide the theory on waveguide-based Raman scattering before summarizing the work that already has been done on nanophotonic waveguide-enhanced Raman spectroscopy.

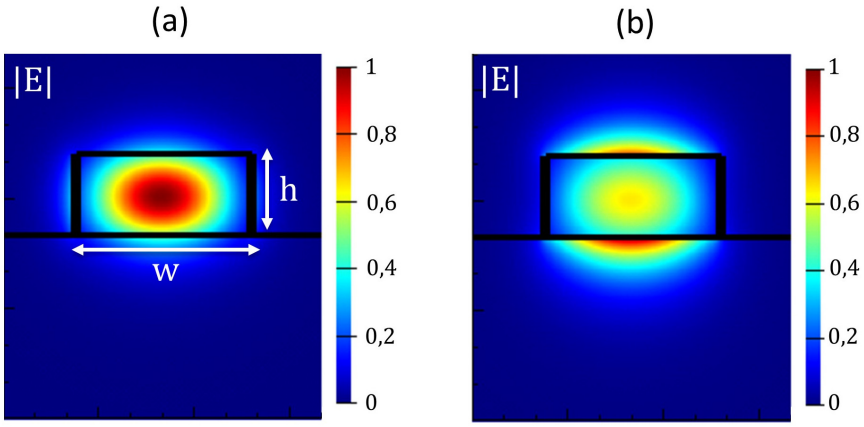


Figure 3.2: Simulated (a) TE fundamental mode and (b) TM fundamental mode for a Si_3N_4 *strip waveguide with an ethanol top cladding. The strip waveguide has dimensions of width* $\times$ *height: 550 nm* $\times$ *220 nm .*

3.2 Waveguide-enhanced Raman scattering

In Chapter 2, we used classical theory in order to describe Raman scattering which fails to recognize the inherent similarities between spontaneous and stimulated Raman scattering. From a quantum perspective, spontaneous Raman scattering can be seen as a type of stimulated Raman scattering that is stimulated by vacuum fluctuations. By Heisenberg's uncertainty principle, even in vacuum there will be some electromagnetic radiation present which stimulates the atom to relax and emit a Stokes photon. Therefore, spontaneous emission can be described as a form of stimulated emission [70–72]. A lot of work has already been done on waveguide-based spontaneous Raman scattering [14, 25, 73] and therefore it will be interesting to come to an expression of the signal evolution for waveguide-based spontaneous and stimulated Raman scattering that follows a unified treatment. In this section we largely follow the theory presented in [35, 45].

3.2.1 Raman scattering in free-space

Before starting we can first make a simplification on our theoretical model. Since Raman scattering is a complex 3D process, in principle we need to keep track of the polarization and propagation directions of the optical waves involved. Luckily we only consider isotropic materials (e.g. silicon nitride), which means that for Raman scattering in these materials, the scattering field intensity will be independent of the propagating direction.Therefore, it is sufficient to consider the angle between the

polarisations of the participating optical fields to determine the optical power of the scattered field and we do not have to worry about most of the vectorial character. In the rest of this chapter the denotations $||$ and $\perp$ refer to the polarisation of light parallel and perpendicular to the pump beam.

Let us now consider spontaneous Raman scattering in a bulk medium with free-space excitation and collection. We can assume that a pump beam (with power P_z) is propagating in the +z direction, has a optical frequency ω_p and polarisation along a direction e_p. The total Raman scattered photons (with a total power P_s) from a unit volume dV = dAdz that propagates within a solid angle $d\vartheta$ and a frequency range $[\omega_s, \omega_s + d\omega_s]$ then corresponds to:

$$P_s = \frac{\omega_s}{\omega_p} P_p R d\vartheta dz d\omega_s \tag{3.1}$$

Where R is the proportionality factor which is given by:

$$R = \frac{1}{32\pi^2 \hbar \epsilon_0^2 c^4} \frac{n_s}{n_p} \omega_p \omega_s^3 [M_{||}(|\Omega|)\cos^2\theta + M_{\perp}(|\Omega|)\sin^2\theta)] h(\Omega, T) \tag{3.2}$$

With ω_p and ω_s the pump and Stokes frequency and $\Omega = \omega_p - \omega_s$, $\hbar$ is the reduced Planck's constant, n_p and n_s are the refractive indices at the pump and Stokes frequency and θ is the angle between the pump beam polarization e_p and Stokes beam polarization e_s. $M_{||}(|\Omega|$ and $M_{\perp}(|\Omega|$ are the materials Raman response for the different polarizations and $h(\Omega, T)$ is the Boltzman factor that is described as:

$$h(\Omega, T) = \begin{cases} (1 - e^{\frac{-\hbar\Omega}{kT}})^{-1} \text{ if } \Omega > 0 \text{ (Stokes)} \\ (e^{\frac{-\hbar\Omega}{kT}} - 1)^{-1} \text{ if } \Omega < 0 \text{ (anti – Stokes)} \end{cases} \tag{3.3}$$

We can see that Eq. 3.2 scales with ω^4 just as in the classical case. Furthermore the Stokes scattered power is also stronger then the anti-Stokes power by a factor $e^{\frac{-\hbar\Omega}{kT}}$ which can be obtained from $h(\Omega, T)$.

We can now consider stimulated Raman scattering in a bulk medium where a pump and Stokes beam are incident on the Raman medium. Both the pump and Stokes beam are assumed to be infinite plane waves with an electric field polarization e_p and e_s respectively. We assume the Stokes beam to be propagating in the z+ direction and no assumption is made about the pump beam propagation direction. We can decompose the Stokes beam polarization (e_s) in a component parallel $e_{||}$ and perpendicular $e_{\perp}$ to the pump. If we now consider the change in

Stokes intensity $\Lambda_{||}(z)$ (with the same polarization as the pump beam and within the frequency range $[\omega_s, \omega_s + d\omega_s]$) then its evolution can be described as:

$$\frac{d\Lambda_{||}}{dz} = g_{||} I_p \Lambda_{||}(z) \tag{3.4}$$

Where I_p is the pump intensity and the stimulated Raman gain $g_{||}$ defined as:

$$g_{||} = \frac{\Omega}{|\Omega|} \frac{\pi}{4\hbar^2 \epsilon_0^2 c^2} \frac{\omega_s}{n_p n_s} M_{||}(|\Omega|) \tag{3.5}$$

The parameters are the same as those defined in Eq. 3.2. An equivalent solution is obtained for the perpendicular polarised Stokes component (simply replacing the subscript $||$ by $\perp$). We can see from Eq. 3.5 that there is no limitation on the sign of $|\Omega|$ and therefore for $\Omega > 0$ the signal is Stokes-shifted from the pump and experiences stimulated Raman gain and vice versa for $\Omega < 0$ where the signal will experience stimulated Raman loss. This is equivalent to the result obtained using classical theory.

Since the laser linewidth is usually much smaller then the linewidth of the Raman modes (assuming a CW or quasi-CW condition) we can see $g_{||}$ and I_p as being roughly constant over the linewidth of the Stokes beam $\Delta\omega_s$. Therefore, Eq. 3.4 can be written as:

$$\frac{dI_s}{dz} = \int_{\Delta\omega_s} g_{||} I_p \Lambda_{||}(z) = g_{||} I_p I_s \tag{3.6}$$

With I_s the Stokes beam intensity. It is interesting to note that both spontaneous Raman and stimulated Raman scattering are connected through $M_{||}$ ($M_\perp$). In this way the stimulated response can be estimated from the spontaneous Raman response of an analyte.

3.2.2 Raman scattering near waveguides

We will now consider SRS and spontaneous Raman scattering when using waveguides. Unlike the free-space description of Raman scattering, for waveguide-based Raman the excitation field and collected field are confined to the waveguide. Therefore, not all of the guided pump light will interact with the analyte as a part of it is confined to the waveguide core and does not overlap with the Raman medium (as an example see Fig. 3.2). This is thus a difference between free-space and waveguide-based Raman.

If we now assume that the waveguide mode is confined to the xy-plane, that the Stokes beam is propagating in the +z direction of the waveguide (in the forward direction, see Fig. 3.4) and that its polarization is parallel to that of the pump beam, then the evolution of the Stokes intensity $\Gamma_{||}(z)$ (within the frequency range $[\omega_s, \omega_s + d\omega_s]$) can be described as:

$$\frac{d\Gamma_{||}(z)}{dz} = \underbrace{\frac{\omega_s}{\omega_p} R_{||} P_p}_{\text{spontaneous}} + \underbrace{g_{||}\Gamma_{||}(z) P_p}_{\text{stimulated}} \tag{3.7}$$

We can see that there is a contribution for both spontaneous Raman scattering and Stimulated Raman scattering. With $R_{||}$ the probability that a pump photon is converted to a Stokes photon within a distance dz and where the scattered photon will preserve its original polarization, will have a frequency in the $[\omega_s, \omega_s + d\omega_s]$ range and propagates along the +z direction. For the SRS-term, $g_{||}P_p$ describes the gain of the Stokes beam over dz. The gain coefficients of the (stimulated) Raman scattering processes correspond to:

$$R_{||} = \frac{1}{8\hbar\epsilon_0^2 c^2} \frac{\omega_p \omega_s}{n_p n_s} h(\Omega, T) \iint_{\text{Raman region}} M_{||}(|\Omega|)|f_p|^2 |f_s|^2 dxdy \tag{3.8}$$

$$g_{||} = \frac{\Omega}{|\Omega|} \frac{\pi}{4\hbar^2\epsilon_0^2 c^2} \frac{\omega_s}{n_p n_s} \iint_{\text{Raman region}} M_{||}(|\Omega|)|f_p|^2 |f_s|^2 dxdy \tag{3.9}$$

Where n_p, n_s are the effective indices and $f_p(x, y)$, $f_s(x, y)$ are the normalized transverse mode profiles of the pump and Stokes beam. Note that to obtain $g_\perp$ and $G_\perp$, not only $M_\perp$ needs to be replaced but also the effective indices and their mode profiles. We can also see that the gain coefficients of waveguide based Raman mainly differ with the free-space alternative through the integral:

$$\iint_{\text{Raman region}} M_{||}(|\Omega|)|f_p|^2 |f_s|^2 dxdy \tag{3.10}$$

Where this integral is determined by the waveguide geometry as will be discussed in the following subsection. Since we will mainly be interested in TE-modes for both the pump and Stokes beam in this work (as they provide the best signal enhancement in a slot waveguide with low loss) we will use $g_{||}$ in the remainder of this chapter which we will name *g* for simplicity.

Before further discussing the effect of the waveguide geometry on the gain factors let us first connect the Raman response function to the scattering cross-section σ as defined in Eq. 2.16. In this way $M(\Omega)$ can be obtained from the

experimentally captured σ. In [35] it is shown that the material response function can be written as:

$$M_{||}(|\Omega|) = \left(\frac{\partial^2\sigma}{\partial\omega\partial\vartheta}\right)\frac{\rho N_a}{M_u}\frac{32\pi^2\hbar\epsilon_0^2c^4}{\omega_s^4}\frac{n_p}{n_s} \tag{3.11}$$

Where M_u is the molar mass of the analyte, ρ the density of the analyte, N_a Avogadro's number and $\left(\frac{\partial^2\sigma}{\partial\omega\partial\vartheta}\right)$ the spectral density of the differential Raman scattering cross-section. If we now assume that we have a Raman mode with a Lorentzian lineshape then the $\max(M_{||})$ can be written as a function of $d\sigma/d\vartheta$:

$$\max(M_{||}) = \frac{2}{\pi\Delta_R}\left(\frac{d\sigma}{\partial\vartheta}\right)\frac{\rho N_a}{M_u}\frac{32\pi^2\hbar\epsilon_0^2c^4}{\omega_s^4}\frac{n_p}{n_s} \tag{3.12}$$

where Δ_R is the full-width-half-maximum (FWHM) of the Raman peak. Note that differential cross section $\frac{d\sigma}{\partial\vartheta}[cm^2/sr]$ is often reported over σ in literature since in most configurations only a fraction of the scattered light is captured and furthermore the radiation pattern of dipoles is not uniform and depends on its environment and the orientation of the molecule compared to the light polarization. Therefore, it is more correct and practical to report $\frac{d\sigma}{\partial\vartheta}$ [25].

3.2.2.1 Geometrical factor

We can now have a look of the effect of the waveguide geometry on the spontaneous and stimulated Raman gain coefficients. We can simplify Eq. 3.10 by assuming that the analyte is homogeneous such that $M_{||}(|\Omega|)$ is a constant in the cladding region:

$$M_{||}(|\Omega|)\iint_{\text{Raman region}}|f_p|^2|f_s|^2dxdy \tag{3.13}$$

If we now assume that the mode profile of the pump and Stokes beam are similar such that $|f_p| \approx .|f_s|$ and:

$$\iint_{\text{Raman region}}|f|^4dxdy = \frac{\kappa}{A_{eff}} \tag{3.14}$$

Where A_{eff} $[m^2]$ is the effective mode area and is defined as:

$$\frac{1}{A_{eff}} = \iint_{-\infty}^{+\infty}|f|^4dxdy \tag{3.15}$$

And κ is the confinement factor:

$$\kappa = \frac{\iint_{\text{Raman region}} |f|^4 dxdy}{\iint_{-\infty}^{+\infty} |f|^4 dxdy} \tag{3.16}$$

We can see that A_{eff} characterizes the size of the guided mode and a smaller mode area will hence lead to higher signal generation. The confinement factor on its turn is dimensionless and can be seen as an indication of the percentage of light of the waveguide mode that effectively is propagating in that analyte. Therefore, we can see from Eq. 3.14 that the Raman signal is maximised if the mode is tightly-confided and has a good overlap with the probed analyte. Furthermore, if we include the effective indices at the pump and Stokes frequency we can define the geometrical coefficient:

$$GC = \frac{1}{n_p n_s} \frac{\kappa}{A_{eff}} \tag{3.17}$$

For a silicon nitride strip waveguide covered with IPA (width $\times$ height = 400 nm $\times$ 300 nm), the geometrical coefficient for the fundamental TE-mode corresponds to GC_{TE} = 0.25 and GC_{TM} = 0.45 for the fundamental TM-mode [35]. The waveguide width (and height) are typically designed to maximize GC such that it provides a trade-off between having a small mode area A_{eff} and a large confinement factor κ. Making the waveguide width bigger then the optimal value (for a fixed height) will result in an increase of A_{eff} and less overlap with the analyte and hence a decrease in κ. In contrast, if the width is too small the waveguide mode will not be properly guided which results in an increase of A_{eff} which offsets any improvement in κ. For a strip waveguide, the TM-mode has a larger geometrical coefficient since it has a more extensive overlap with the analyte while A_{eff} is comparable for both polarisation modes. It is important to note that the TM-mode typically also has a larger propagation loss and therefore not always leads to better results than the TE-mode [74].

Finally, the Raman gain coefficients can now also be expressed in function of A_{eff} and κ such that we get:

$$R_{||} = \frac{1}{8\hbar\epsilon_0^2 c^2} \frac{\omega_p \omega_s}{n_p n_s} h(\Omega, T) M_{||}(|\Omega|) \frac{\kappa}{A_{eff}} \tag{3.18}$$

$$g_{||} = \frac{\Omega}{|\Omega|} \frac{\pi}{4\hbar^2 \epsilon_0^2 c^2} \frac{\omega_s}{n_p n_s} M_{||}(|\Omega|) \frac{\kappa}{A_{eff}} \tag{3.19}$$

It is interesting to note that besides the geometrical factors, the other parameters of the gain coefficients are constant or set by the used analyte and technology platform.

3.2.2.2 Linking the spontaneous and stimulated gain

It is now interesting to link the spontaneous Raman gain (G), as will be defined below, to the stimulated Raman gain (g) since a considerable amount of work has already been performed on calculating G for a diverse set of waveguide geometries and materials [14].

So in order to come to a workable expression, we will first rewrite Eq. 3.7 such that it describes the evolution of the Raman scattered power over the whole Raman mode. The signal evolution in the spontaneous Raman case then corresponds to:

$$\frac{dP_s}{dz} = G_{||} P_p \tag{3.20}$$

With $G_{||}$ the spontaneous Raman gain coefficient for P_s which corresponds to:

$$G_{||} = \int_{-\infty}^{+\infty} R_{||} d\Omega = \frac{\omega_s}{\omega_p} \frac{1}{8\hbar\epsilon_0^2 c^2} \frac{\omega_p \omega_s}{n_p n_s} h(\Omega, T) \frac{\kappa}{A_{eff}} \int_{-\infty}^{+\infty} M_{||}(|\Omega|) d\Omega \tag{3.21}$$

For simplicity we will use G instead of $G_{||}$ in the remainder of this chapter.

If we now assume that we are using (quasi-) CW lasers, such that their linewidths are much smaller than the Raman linewidth, then we can assume that the stimulated Raman gain g is constant over the laser linewidth. We can then describe the signal evolution in the stimulated Raman case as:

$$\frac{dP_s}{dz} = g P_p P_s \tag{3.22}$$

where g corresponds to (see Eq. 3.19):

$$g = \frac{\Omega}{|\Omega|} \frac{\pi}{4\hbar^2 \epsilon_0^2 c^2} \frac{\omega_s}{n_p n_s} M_{||}(|\Omega|) \frac{\kappa}{A_{eff}} \tag{3.23}$$

And if we now assume that we have a Lorentzian lineshape for the Raman response function $M_{||}(|\Omega|)$, we can find that:

$$\frac{G}{\max(g)} = h(\Omega, T)\frac{\hbar\omega_p}{2\pi \max(M_{||})}\int_{-\infty}^{+\infty} M_{||}d\Omega = \frac{\hbar\omega_p}{4}h(\Omega, T)\Delta_R \tag{3.24}$$

With Δ_R the full-width-half-maximum of the Raman response and where $h(\Omega, T)$ is close to unity at room temperature.

3.2.2.3 Alternative definition of the spontaneous gain

In [75] and [64] an alternative definition of the spontaneous Raman gain coefficient G is given namely:

$$G = \frac{1}{2} \times \rho \times \sigma \times \eta(w, h, \lambda_p, \lambda_s) \tag{3.25}$$

With ρ, σ the molecular density and differential Raman cross section of the analyte and η the Raman conversion efficiency which depends on the modal properties of the waveguide (a.o. the width (w) and height (h) of the waveguide). Eq. 3.21 and 3.25 are basically two sides to the same coin as the two definitions of G calculate the same numerical value for the spontaneous Raman gain factor. In the rest of this work we will use both definitions depending on the situation to make estimates on the spontaneous Raman gain.

A complete derivation of Eq. 3.25 is out of the scope of this thesis and can be found in [45, 73]. We will quickly give a bit more detail on η since it encompasses the effect of the waveguide on spontaneous Raman scattering.

$$\eta = \frac{n_g^2\lambda_s^2}{n}\iint \frac{\tilde{\epsilon}|\tilde{E}(\tilde{r_0}, \lambda_p)|^2}{\iint \tilde{\epsilon}|\tilde{E}(\tilde{r}, \lambda_p)|^2 d\tilde{r}}\frac{1}{\tilde{A}'_{eff}(\tilde{r_0}, \lambda_s)}d\tilde{r} \tag{3.26}$$

Where the outer integral runs over the analyte (e.g. to estimate the amount of signal generation for an analyte on top of the waveguide this area corresponds to the top cladding of the waveguide). Furthermore, we can see that η for a given waveguide geometry depends on the model profile where A'_{eff} (assuming a dispersionless medium) is defined as:

$$A'_{eff} = \frac{\iint \left(2\epsilon_0\tilde{\epsilon}|\tilde{E}(\tilde{r}, \lambda_s)|^2\right) d\tilde{r}}{\epsilon_0\tilde{\epsilon}|\tilde{E}(\tilde{r_0}, \lambda_s)|^2} \tag{3.27}$$

If we now assume that a molecule can be approximated as a dipole emitter, then the enhanced ratio of the spontaneous emission rate for a dipole emitter near

a waveguide (where part of the emission will radiate into the waveguide mode at a rate γ_{wg} while the rest will decay into radiating modes) compared to free-space (γ_0) corresponds to:

$$\frac{\gamma_{wg}(\tilde{r},\omega)}{\gamma_0(\tilde{r},\omega)} = \frac{3}{4\pi}\frac{c}{v_g n_i^3}\frac{\lambda_s^2}{A_{eff}^{'}} \tag{3.28}$$

Where v_g is the group velocity of the waveguide mode, λ_s the Stokes wavelength and n_i the refractive index of the surrounding medium in which the dipole emitter is placed. We can thus see from Eq. 3.28 that in order to enhance the spontaneous emission rate, the waveguide needs to support a mode with a small $A_{eff}^{'}$ (increases local intensity) and v_g (increases local density of states). Therefore, typically slot waveguides are designed with small slot gaps because the high field confinement in the narrow gaps is attributed to an efficient excitation and collection of the scattered signal (for a dielectric slot in air with a 40 nm gap $\frac{\gamma_c}{\gamma_0} \approx 4$). Plasmonic slots on their end, combine high field intensities in the gap together with a large group index of the gap plasmonic mode. Therefore, $\frac{\gamma_c}{\gamma_0}$ for the plasmonic slot is about two orders of magnitude stronger than a dielectric slot with the same gap width. Eq. 3.28 can also be written as:

$$\frac{\gamma_{wg}}{\gamma_0} = \frac{P_{wg}}{P_0} \tag{3.29}$$

Where P_{wg} is the power coupled into the waveguide mode and P_0 the power radiated in a homogeneous medium. And the dipole strength of the dipole excited by a waveguide $\tilde{E}(r,\tilde{\lambda}_p)$ with power P_p corresponds to:

$$|\tilde{d}(\tilde{r},\lambda_s|^2 = \alpha^2 n_g P_p \frac{\tilde{\epsilon}|\tilde{E}(\tilde{r_0},\lambda_p)|^2}{\iint \tilde{\epsilon}|\tilde{E}(\tilde{r},\lambda_p)|^2 d\tilde{r}} \tag{3.30}$$

Where α is the molecular polarizability, in [73] it is shown that by using $\sigma = \frac{\pi^2\alpha^2}{\epsilon_0^2\lambda^4}$ the total power coupling into a waveguide mode in backward direction from dipoles lying uniformly over a waveguide of length L corresponds to:

$$\frac{P_{wg}^{tot}}{P_p} = \overbrace{\frac{1}{2}\times\rho\times\sigma\times\eta}^{G}\times LF \tag{3.31}$$

We thus see the appearance of our gain factor of Eq. 3.25 and where LF is the length factor which will be discussed in the next section.

Typically values for η_A (where A indicates that we integrated over the analyte cladding region) corresponds to $\eta_A^{TE} = 0.075$ and $\eta_A^{TM} = 0.16$ for the strip waveguides shown in Fig. 3.2. The fundamental TM-mode for the strip waveguide has a higher η_A due to the higher modal overlap with the analyte and higher field intensity at the analyte location which originates from the discontinuity of the electric field at the core-top cladding interface (owing to the high width-height aspect ratio). Note that for the TM mode, since the electric field is only discontinuous in the perpendicular direction, only varying the waveguide height will have a substantial effect on the field enhancement near the waveguide surface. For the slot waveguides shown in Fig. 3.3, $\eta_A^{TE} = 0.31$ and $\eta_A^{TM} = 0.13$. This shows that the slot waveguide with a excited fundamental TE-mode has an improved η_A due to the tightly confined slot mode. The light for the TE-mode is mostly confined to the gap and thus provides a good overlap with the analyte together with a large field enhancement due to the discontinuity of the refractive index at the core-slot interface leading to a better η_A compared to the TM-mode. In contrast, for the TM-mode the electric field is continuous in the gap and therefore there will be a much smaller enhancement. For decreasing slot gaps, the enhancement of the TE-modes will further increase and thus lead to a better η_A^{TE}. In general, small slot gaps and a high refractive index of the waveguide material will lead to high Raman conversion efficiencies [45]. Higher index contrasts lead to smaller mode area‘s, higher field intensities of the mode and high field discontinuities which results in stronger evanescent fields and also lead to a reduced overlap with the waveguide core which lowers its background contribution. In [75], it is experimentally shown that slot waveguides have about a six-times more efficient signal generation compare to the strip waveguide, where in both cases the fundamental TE-mode is excited and the same guided pump power is assumed. Note that the effective improvement of a slot waveguide over a strip waveguide is not as big, as slot waveguides typically experience more loss than strip waveguides depending on the analyte and the fabrication which also needs to be accounted for (with upper bounds in the propagation loss of 10 dB/cm) [35].

3.2.3 Collection Schemes

Let us now consider the different collection schemes to collect the spontaneous (stimulated) Raman scattered light. Afterwards, we can come to workable expressions that describes the evolution of the generated Raman power as a function of the gain factors as defined in the previous section, the interaction length L and propagation loss α. We will not take into account other contributions to the signal beam such as the Kerr effect, free carrier effects and dispersive effects since these contributions are expected to be negligible in the (quasi-) CW power regime considered here.

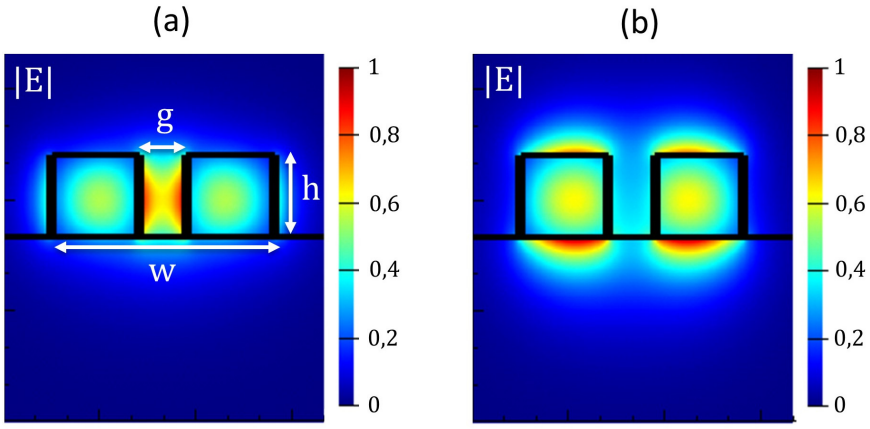


Figure 3.3: Simulated (a) TE fundamental mode and (b) TM fundamental mode for a Si_3N_4 slot waveguide with an ethanol top cladding. The slot waveguide has dimensions of width × height: 700 nm× 220 nm and a 150 nm gap.

3.2.3.1 Spontaneous Raman scattering

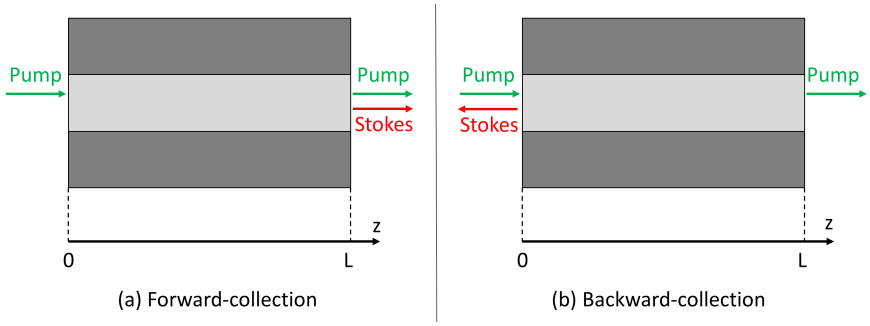


Figure 3.4: (a) Forward-collection scheme and (b) backward-collection scheme used to collect the Raman scattered light in waveguide-based spontaneous Raman spectroscopy.

Let us first start with spontaneous Raman scattering. We can now consider a waveguide where the pump light is injected at z=0 such that it propagates in the +z direction (see Fig. 3.4). Due to the isotropic nature of the dipole emission an equal amount of the Stokes light will couple in the forward direction (+z) as in the backward direction (-z). Therefore, we a make distinction in the collection scheme that collects the forward-propagating Stokes light (i.e. the forward-collection schemes) and the backward-propagating Stokes light (i.e. the backward-collection scheme). It is interesting to note that although the generated forward- and backward-propagating Stokes light will experience the same gain G over the interaction

length L, the amount of Stokes power collected at the opposite end-facets might differ. In order to better understand this let us first consider the forward-collection configuration (3.4a). Since the generated Stokes beam (s) is propagating in the same direction as the pump beam (p) (namely in the +z direction) the evolution of the Stokes beam can be expressed as:

$$\frac{dP_s}{dz} = GP_p(z) - \alpha_s P_s(z) \tag{3.32}$$

With α_s the propagation loss at the Stokes frequency (due to material absorption and surface roughness) and G the spontaneous Raman gain. Since the Raman interaction is only a weak phenomena its effect on the pump beam can be neglected and the evolution of the pump beam corresponds to $P_p(z) = P_p(0)e^{-\alpha_p z}$ with α_P the propagation loss at the pump frequency. If we now use the boundary boundary condition that $P_s(0) = 0$, the generated Stokes beam after a length L can be expressed as:

$$P_s(L) = GP_p(0)\overbrace{e^{-\alpha_p L}\frac{1 - e^{-(\alpha_s - \alpha_p)L}}{\alpha_s - \alpha_p}}^{LF_F} \tag{3.33}$$

Where we introduce the length factor (LF_F). In case the propagation loss at the pump and Stokes frequency is the same ($\alpha_s = \alpha_p$) this length factor will reduce to $LF_F = L \times e^{-\alpha_p L}$.

If we now consider the alternative backward-collection, where the pump is propagating in the opposite direction (+z) as the Stokes beam (-z), the evolution of the Stokes beam can be expressed as:

$$\frac{dP_s}{dz} = -GP_p(z) + \alpha_s P_s(z) \tag{3.34}$$

If we again use the boundary condition that there is no Stokes beam present at z= L such that $P_s(L) = 0$, the generated Stokes beam after a length L can be expressed as:

$$P_s(0) = GP_p(0)\overbrace{\frac{1 - e^{-(\alpha_p + \alpha_s)L}}{\alpha_p + \alpha_s}}^{LF_B} \tag{3.35}$$

The signal generation in both detection schemes thus differs through the length factors. If we now plot $LF_{Forward}$ & $LF_{Backward}$ assuming a loss of $\alpha = \alpha_p = \alpha_s =$

2 dB/cm (see Fig. 3.5) we can see that LF_F first increases, reaches a maximum and then decays after a certain waveguide length ($L_s = 1/\alpha$), while LF_B will increase and then saturate. For the forward-collection scheme there thus exist an optimal waveguide length which needs to be accounted for when doing experimental work. This becomes especially problematic when the analyte is the main loss contributor, such that this optimal length changes between analytes and different waveguides need to be used. In contrast, for the backward-collection scheme L_s corresponds to the threshold length for optimal signal generation. Furthermore, we can see that LF_B is approximately 1.4 times stronger than LF_F. The difference between both configurations lies in the fact that in the forward-configuration, the generated Stokes signal at the beginning of the waveguide (where the signal generation is the strongest signal due to the limited effect of α_p on the pump power) has to propagate over the whole waveguide length before it is collected. Therefore, the generated signal in the backward-configuration will be stronger (and will saturate after a certain length) since it does not experience this extra propagation loss. Note that Eq. 3.34 and 3.35 have been experimentally verified in [13] and [16] and show good agreement with theory.

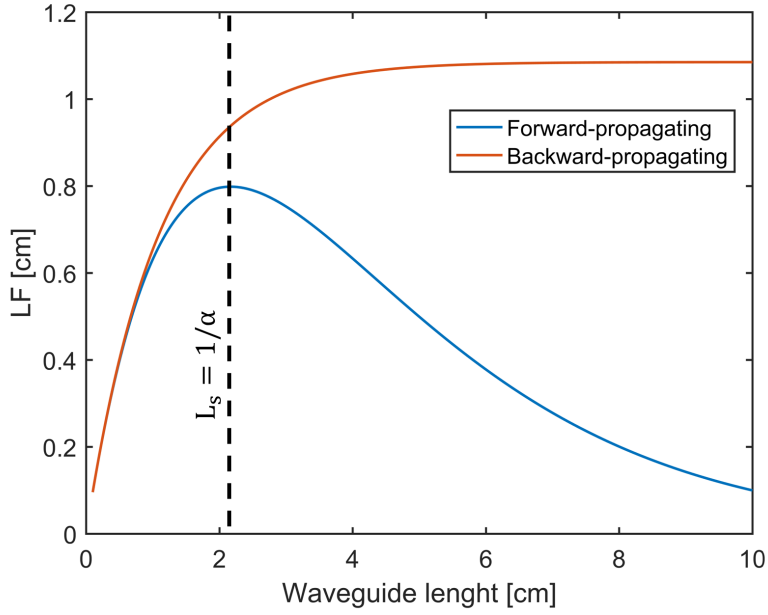


Figure 3.5: Comparison of the length factor (LF) for a forward- (blue) and backward- (red) propagating Raman signal.

It is also interesting to compare Eq.3.35 to the pump to Stokes conversion efficiency of a diffraction limited Raman microscope [76]:

$$\frac{P_s}{P_p} = 2\rho\sigma\left(\frac{\lambda_e}{n}\right) \tag{3.36}$$

Where we can see that for a diffraction limited system, the pump to Stokes conversion is only dependent on the excitation wavelength, the Raman cross section and molecular density of the analyte (where n is the refractive index of the immersion medium). In contrast, for waveguide enhanced Raman spectroscopy the pump to Stokes conversion scales with the waveguide length L. Therefore, the interaction length can be increased by many-fold and combined with the high index contrast leads to strong signal enhancements compared to the conventional Raman microscope. In [76] it is shown that for a lossless waveguide the ratio of the generated Stokes signal for the waveguide compared to diffraction limited microscopy corresponds to:

$$\Theta = \frac{P_{wg}}{P_{DL}} = \frac{\eta}{2}\left(\frac{L}{\lambda_e/n}\right) \tag{3.37}$$

Where for a 1-cm long strip and slot (with λ_e = 785nm) this ratio theoretically corresponds to Θ_{strip} = 500 and Θ_{slot} = 2000. In practice this ratio is an order of magnitude smaller since (a.o. the waveguide loss also needs to be included).

3.2.3.2 Stimulated Raman scattering

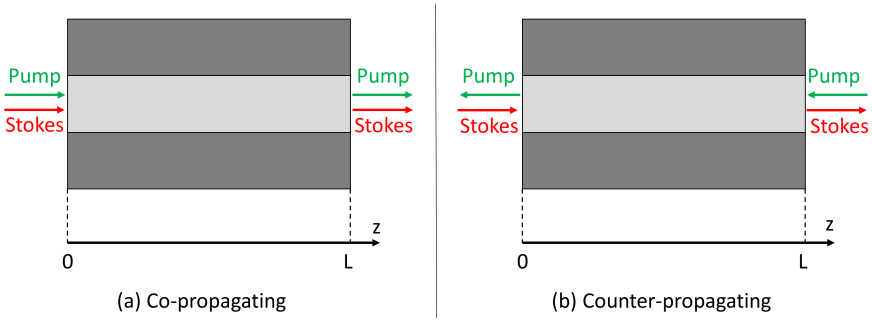


Figure 3.6: (a) Co-propagating configuration and (b) Counter-propagating configuration used to collect the Raman scattered light in waveguide-based stimulated Raman spectroscopy.

Analogously, we can make a comparable analysis for stimulated Raman scattering. In SRS we couple in two lasers (i.e. the pump and Stokes beam, see Section 4.2 for more details). If we assume that the Stokes beam is coupled in at z=0 and

propagates in the z+ direction then there can be two different configurations for SRS. In the first configuration the pump beam is injected at the same facet (z=0) and co-propagates along the Stokes beam in the z+ direction, this configuration is therefore called the co-propagating configuration (see Fig. 3.6a). In the second configuration the pump beam is injected at the opposite facet (z=L) and is counter-propagating to the Stokes beam in the z- direction, this configuration is therefore labelled as the counter-propagating configuration (see Fig. 3.6b).

If we now consider the co-propagating configuration, then the evolution of the Stokes beam can be expressed as:

$$\frac{dP_s}{dz} = gP_s(z)P_p(z) - \alpha_s P_s(z) \tag{3.38}$$

Where we can again assume that the pump power will not be affected by the Raman interaction such that its evolution is expressed as $P_p(z) = P_p(0)e^{-\alpha_p z}$ and the generated Stokes beam after an interaction length L can be solved as:

$$P_s(L) = P_s(0)\exp\left(g\,P_p(0)\,\frac{1-e^{-\alpha_p L}}{\alpha_p} - \alpha_s L\right) \tag{3.39}$$

Where we can define the effective length as:

$$L_{eff} = \frac{1-e^{-\alpha_p L}}{\alpha_p} \tag{3.40}$$

Such that Eq. 3.39 becomes:

$$P_s(L) = P_s(0)\exp\left(g\,P_p(0)\,L_{eff} - \alpha_s L\right) \tag{3.41}$$

If we now consider the alternative counter-propagating configuration, then the Stokes beam will still be injected at z=0 (propagating in the z+ direction) while the pump beam is injected at z=L (propagating in the z- direction). The evolution of the Stokes beam then corresponds to:

$$\frac{dP_s}{dz} = gP_s(z)P_p(z) - \alpha_s P_s(z) \tag{3.42}$$

With the evolution of the pump beam (ignoring Raman loss) corresponding to:

$$P_p(z) = P_p(L)e^{-\alpha_p(L-z)} \tag{3.43}$$

The generated Stokes beam after an interaction length L, then can be solved as:

$$P_s(L) = P_s(0)\exp\left(g\,P_p(L)\,\frac{e^{-\alpha_p L}(e^{\alpha_p L}-1)}{\alpha_p} - \alpha_s L\right) \tag{3.44}$$

Which can be written as

$$P_s(L) = P_s(0)\exp\left(g\,P_p(L)\,L_{eff} - \alpha_s L\right) \tag{3.45}$$

Where L_{eff} corresponds to the definition given in Eq. 3.40. We can thus see that (unlike spontaneous Raman scattering), for stimulated Raman scattering the collected Stokes power at the opposite end-facets is the same for both configurations. Therefore, in experiments we can choose the most practical configuration without being confronted by a less optimal signal generation. Note that the co-propagating configuration might have a slightly higher Stokes beam intensity when other nonlinear effects such as the Kerr effect are taken into account [77].

Since Raman is only a weak effect Eq. 3.45 can be further approximated as:

$$\begin{aligned} P_s(L) &= P_s(0)(1 + g\,P_p(L)\,L_{eff})\exp(-\alpha_s L) \\ &= P_s(0)\exp(-\alpha_s L) + P_s(0)\,g\,P_p(L)\,\underbrace{L_{eff}\exp(-\alpha_s L)}_{L_{eff,tot}} \end{aligned} \tag{3.46}$$

Where we used $\exp(x) \approx 1 + x$ for $x \to 0$. We can thus see that the Stokes beam at the output facet consists of a strong term on the left side which does not contain any Raman information (and simply corresponds to the attenuated input Stokes beam) and a second much weaker term on the right-side where the Raman information is embedded in the stimulated Raman gain factor (g). This result is to be expected since, unlike in spontaneous Raman, in stimulated Raman the signal generation occurs on top of the already present Stokes beam. Therefore, special detection techniques such as a lock-in detection scheme are needed to extract the Raman information in SRS (see Section 4.2 for a detailed discussion).

3.3 Nanophotonic waveguide enhanced Raman spectroscopy

Nanophotonic waveguide enhanced Raman spectroscopy (NWERS) was for the first time demonstrated in our group by Dhakal, et. al. [13], where a bulk liquid (IPA) was detected on top of a single mode PECVD silicon nitride waveguide. Since

then detection of alternative analytes (e.g. organic compounds [78, 79], gasses [80], DNA [14] and hemoglobin [81]), alternative waveguide materials (e.g. tantalum oxide waveguides [82] and titania waveguides [83]) and waveguide structures (e.g. slot waveguides [75]) have been demonstrated.

Despite the successful demonstration of NWERS, two main challenges still remain when performing spontaneous Raman spectroscopy on dielectric waveguides. First off, the signal enhancement provided by the waveguides is still too low to avoid the use of deeply-cooled detectors. This hinders the demonstration of a fully integrated Raman-on-chip spectroscopic system [19]. As an example, when measuring a bulk liquid (IPA) on 4-cm a long Si_3N_4 slot waveguide, the conversion efficiency (i.e. ratio of the generated Stokes power to the incoming pump power) corresponds to $\zeta = \frac{P_s}{P_p} = 3.4 \times 10^{-10}$ [16]. If we assume a guided pump power of 6 mW [75], signal levels in the order of 3×10^{-13}W need to be detected by the detector after chip (6×10^{-3}[W] $\times\, 3.4 \times 10^{-10} \times 10^{-0.8}$, where the last term corresponds to the -8 dB coupling loss from the waveguide to free-space). Unfortunately, for those signal levels ($< 10^{-12}$W) the use of expensive deeply-cooled camera cannot be avoided. Furthermore, the generated signal power can easily drop two or more orders of magnitude when detecting biologically relevant concentrations instead of bulk liquids [84].

Secondly, there is a background contribution due to the Raman-like scattering of the waveguide core itself (since part of the optical mode overlaps with it). This undesired photon background will ultimately reduce the signal-to-background (and signal-to-noise) ratio of any acquired Raman spectrum. On our Si_3N_4-platform, the present background is broadband and decreases with longer Raman-shifts as is shown in Fig. 3.7. Note that this background changes between different types of Si_3N_4 and even depends on the exact deposition parameters within the same Si_3N_4-type [85]. Furthermore, it scales linearly with the intensity of the excitation beam and with the propagation length (just like the generated Stokes signal from the probed analyte) and is generated in both the forward and backward propagation direction. There is evidence that this background originates from localized thermal fluctuations [86]. The presence of this background will strongly increase the shot noise since the shot noise is proportional to the square root of the total intensity (of background and analyte signal combined) [25]. Since the shot noise leads to random fluctuations of the incoming intensity it will compromise the minimal detectable signal levels. The background shot noise contribution will be the dominant noise source for sufficiently strong pump powers and weak signals relative to the background [87]. For long integration times ($\sim$ min) and high pump power powers, this shot noise contribution will however be sufficiently low (since the signal-to-noise ratio scales with $\sqrt{P_p t_{int}}$ [88]) such that it will not be the main limiting factor on the signal sensitivity. In that case, the strength of the

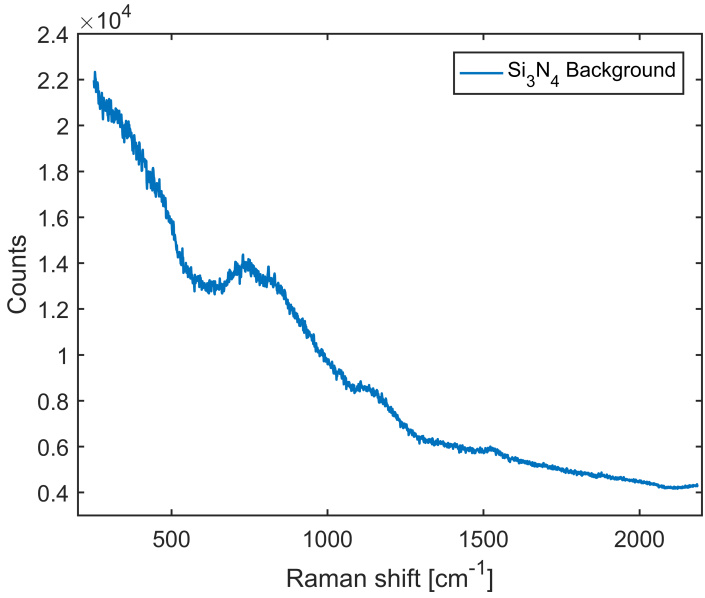


Figure 3.7: Raman spectrum of a Si_3N_4 *strip waveguide.*

background itself will be the limiting factor since there is a practical limit in the accuracy of the background subtraction as the the background needs to be estimated from algorithmic extrapolation or from reference measurements [87]. In practice getting an exact knowledge of the background is challenging because changes in the cladding, small variations in the optical alignment or waveguide alter the overall slope and small features of the background. Furthermore, the background is inherent to the material and does not improve with increased power or integration time. Therefore, in the remainder of this thesis we will use the signal-to-background ratio (SBR) as the figure of merit (over the signal-to-noise ratio) since it is does not depend on the measurement conditions (i.e. integration time and power) and sets the sensitivity limit on the overall strength of the background instead of its shot noise.

One possible solution to get rid of this background contribution could be to change the waveguiding material itself [64, 86, 89]. Let us now consider the two main parameters of the NWERS sensor, i.e. the Raman conversion efficiency and the amount of Raman background generation from the guiding material itself. In [64] it is shown that TiO_2 waveguides posses a high conversion efficiency (because of high index contrast) but suffers from a large background which deteriorates its SBR. In contrast, Al_2O_3 waveguides suffer from a poor conversion efficiency (because of a low index contrast) but exhibit a weak Raman background. Si_3N_4 and Ta_2O_5

waveguides on their turn seem to be nearly equally good as the guiding material since they both exhibit a rather low Raman background while still proving a strong conversion efficiency. Therefore, the Si_3N_4 platform considered here is already a good choice for the guiding material so the solution for now must be found elsewhere. Another solution is to change the waveguide structure itself such that there is less overlap of the optical mode with the waveguide core and it is pushed into the cladding. In [75] it is shown that a TE-mode slot waveguide exhibits a five (two) times lower background generation compared to the TE-mode (TM-mode) strip waveguide, furthermore in Section 3.2.2.3 we already discussed that the conversion efficiency of a TE-mode slot waveguide is a factor six (two) times bigger than a TE- (TM-) mode strip waveguide. Despite these improvements, dielectric slot waveguides still suffer from the cm-long interaction length which inevitably will result in the generation of a strong background contribution. Therefore, the obvious solution is to decrease the amount of background generation by reducing the interaction length and overlap of the optical mode with the dielectric core, while still maintaining the same conversion efficiency (and hence a good signal generation).

In the next chapter, we show that the plasmonic slot waveguide provides exactly these qualities as it provides strong signal levels, supports a optical mode with a limited overlap of the waveguide core and the interaction length is reduced to mere $\sim$ μm's. Furthermore, since the enhancement provided by the plasmonic slot is still not enough to avoid the use of deeply-cooled detectors we will discuss the potential of combing plasmonic enhancement together with coherent Raman scattering (which will enhance signal levels to room temperature detection levels). By performing surface-enhanced coherent Raman scattering on the plasmonic slot we can thus overcome the main challenges of NWERS on dielectric waveguides (i.e. a limited signal generation and a strong background contribution).

4

Comparing waveguide-based Raman techniques

In the last chapter, we discussed the improved Raman signal brought by NWERS but also the challenges it poses in terms of background noise and low signal levels. Alternative waveguide based Raman spectroscopic techniques have been demonstrated to address these challenges. Waveguide based surface-enhanced Raman spectroscopy (SERS) provides similar broadband signal levels as NWERS while drastically reducing the interaction length (and hence the photoemission background [16]). Stimulated Raman spectroscopy (SRS) on its turn provides a sufficient enhancement of the Raman response such that cooled detection is no longer needed [19]. However, further improvement are still possible as on-chip SERS still requires cooled detectors and on-chip SRS makes use of dielectric waveguides with a non-negligible photon background. Surface-enhanced Stimulated Raman spectroscopy (SE-SRS) & surface-enhanced Coherent anti-Stokes Raman spectroscopy (SE-CARS) are interesting candidates to achieve this goal as they combines the advantages of both techniques.

In this chapter we will compare the performance of these alternative on-chip Raman techniques (i.e. on-chip SERS, SRS, SE-SRS and SE-CARS) to NWERS. I first start by reviewing the work on waveguide-based SERS & SRS together with their reported experimental performance and remaining challenges (Section 4.1 & 4.2). Then, I give an overview of the existing literature on SE-SRS using free-space

excitation/collection (Section 4.3.1), followed by my own theoretical estimates on the expected strength and photoemission background of on-chip SE-SRS (Section 4.3.2). Afterwards, I compare those results to that of on-chip SERS & SRS (Section 4.3.3 & 4.3.4). Then, I discuss the existing literature on SE-CARS using free-space excitation/collection followed by my own phase matching calculation for the nanoplasmonic slot (Section 4.4.1). Next, I make my own calculation of the expected signal strength and background for on-chip SE-CARS and discuss why I prefer to explore SE-SRS over SE-CARS in this work (Section 4.4.2 & 4.4.3). Finally, I provide an overview table of my theoretical comparison (Section 4.5).

Section 4.3 of this chapter has been adapted from my published paper [90].

4.1 Surface-enhanced Raman spectroscopy (SERS)

4.1.1 Introduction to on-chip SERS

Despite its advantages, the provided enhancement of NWERS is still relatively low compared to what is possible in SERS [88, 91–93]. Furthermore, they suffer from the presence of a non-negligible photon background due to scattering in the waveguide core which ultimately sets a limit on detection (as discussed in Section 3.3). Therefore, it is tempting to combine plasmonic enhancement with photonic integrated circuits (PICs) as the plasmonic structure leads to large Raman scattering enhancements while the use of PICs to excite and collect the Raman response from the plasmonic structures allows for the integration with other Raman spectroscopic components. This combination has been explored using nanoplasmonic gold bowties antennas together with a single mode Si_3N_4 waveguide which excites and collects the SERS response [73, 94–97]. However, the patterning of these antennas required electron beam lithography which is both time-consuming, resource-intensive and inhibits the exploitation of the full functionality of the matured integrated photonics platform. A later study made use of an alternative nanosphere lithography approach to fabricate gold nanotriangles and showed greater SERS enhancement compared to the bowtie antennas [98]. Yet the signal-to-background ratio (SBR) of both the integrated bowties and antennas suffer from the background of their underlying Si_3N_4 waveguides [88]. Alternative structures have been reported such as the combination of PICs with metallic nanoparticles but they lack in their capability of quantitatively reproducing their Raman spectra [99, 100]. In that context, nanoplasmonic waveguides have recently emerged [16] (see Fig. 4.1) which are capable of further increasing the SERS enhancement of the waveguide based SERS platforms [84, 88]. These nanoplasmonic slots rely on Si_3N_4 photonic circuits and are fabricated by making use of deep-UV lithography and

atomic layer depositions which are mass-scale fabrication techniques. They display a non-resonant (broadband) enhancement [101, 102] (unlike the antennas and bowties) without suffering from a significant photon background from a waveguide dielectric like in NWERS. Furthermore, they have a rather long interaction length (in comparison to the interaction length in common SERS substrates) which leads to a large Raman scattering enhancement so that the required pump and Stokes fields can be of very low power and hence integration on a photonic integrated circuit (PIC) becomes easier. They therefore appear as the ideal option for a full Raman integrated system [88].

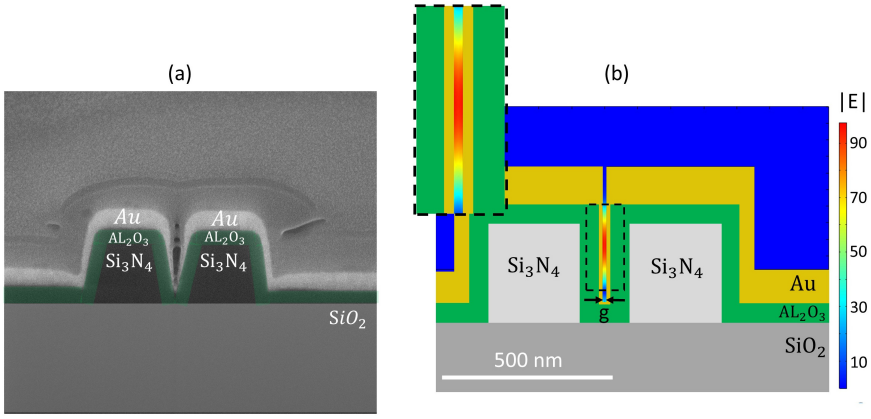


Figure 4.1: (a) 2D cross-section of the nanoplasmonic slot (using SEM) with the different material layers indicated. (b) Simulation (using COMSOl multiphysics) of the highly confined propagating plasmon polariton gap mode, where the slot gap (g) corresponds to 19 nm.

In this work we aim to achieve a gold thickness of $t_{slot} \approx 5.5$ nm in the plasmonic slot together with an alumina thickness of $t_{Al_2O_3} \approx 60$ nm, where the two Si_3N_4 cores are 150 nm apart (see Section 5.1.2 & 6.1.1 for additional details on the plasmonic slot dimensions). In [16] it is shown (through simulations using COMSOL multiphysics) that the ratio $\zeta = P_{SERS}/P_{pump}$ of the generated SERS power (P_{SERS}) to the incoming pump power (P_{pump}) is optimized for those layer dimensions (see Fig. 4.2). Note that ζ is the highest (and nearly constant) for $t_{slot} + t_{Al_2O_3} = 65$ nm indicating on one hand that the nanoplasmonic slot has the highest field enhancements for narrow slots and on the other hand that there is flexibility in choosing the thickness of the gold and alumina layer as long as their combined thickness is 65 nm. Here we target a thicker alumina layer ($t_{Al_2O_3} \approx 60$nm) and a thin gold layer $t_{slot} \approx 5$nm because the thickness of the alumina layer is better controlled (as it is deposited with ALD vs sputtering for the gold layer). Furthermore, a thin gold layer also results in less back-reflection when

transitioning from the access waveguide to the plasmonic slot waveguide (compared to a thick gold layer and a thin alumina layer, resulting in a -16 dB reduction in reflection [45]). Alumina is preferred as a filler layer since it only has a very limited background contribution and can be well controlled since it is fabricated with ALD. Note that the thickness of the gold layers on top and next to the plasmonic slot are less relevant since most of electric field enhancement is situated in the plasmonic slot itself as can be seen from Fig. 4.1. Therefore, varying those gold layers does not have much impact on the pump to Stokes conversion.

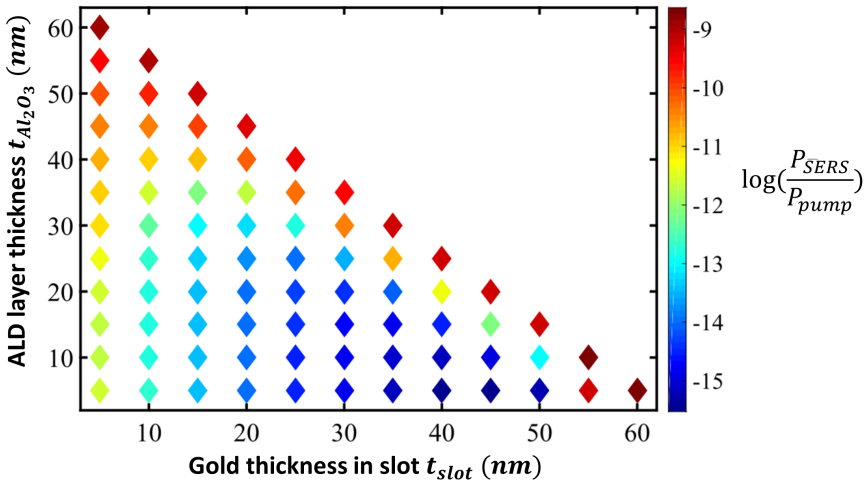


Figure 4.2: Simulated dependence of $\zeta = P_{SERS}/P_{pump}$ *on the alumina thickness* $t_{Al_2O_3}$ *and gold thickness in the plasmonic slot* t_{slot}. *Image adapted from [16].*

4.1.2 Operation

Figure 4.3 illustrates the operation of the nanoplasmonic waveguide as used in the experimental demonstration of [16] where a fraction ($P_{SERS}/P_{pump} = 10^{-9}$) of the incoming pump beam is converted in the Stokes beam through the SERS process. The measurement is done in back-reflection (i.e. the exit waveguide from which the Stokes light is collected is the same as the access waveguide that provides the pump beam to the plasmonic slot). The back-reflection configuration is preferred over the transmission configuration (i.e. the access waveguide is not the same as the exit waveguide and the forward-scattered Raman light is collected behind the plasmonic structure) since the backward-scattered Raman light can be more easily separated from the excitation beam. Due to the isotropic nature of the dipole emission an equal amount of the Raman scattered light couples in the backward as the forward

direction. In this way, the Raman signal can be efficiently collected in the backward collection while the the pump light keeps on propagating in the forward direction. The pump light can therefore not generate any additional photon background in the analyzing circuit after the Raman sensor (which would otherwise deteriorate the SBR, Chapter 5 will give an in depth discussion on this topic).

As mentioned in Section 3.2.3.1, the (plasmonic) slot length should at least corresponds to $L_{plas} = 1/\alpha_{plas} = 1.8$ μm (where $\alpha_{plas} = -2.4$ dB/μm [16] is the plasmonic slot loss) in order to generate the strongest possible backward-propagating Raman signal in the plasmonic slot. In experimental demonstrations the plasmonic slot length is fabricated to multiples of L_{plas} as the signal strength of the backward captured Raman signal saturates and is not additionally effected by further increasing the plasmonic slot length. This is because the strongest Raman signal generation occurs in the beginning of the plasmonic slot while the extra propagation loss offsets any additional Raman signal generation further away. This eases fabrication and assures that the pump excitation field and any associated background from the access waveguide are completely absorbed in the plasmonic structure.

For other on-chip Raman techniques that make use of the plasmonic slot such as SE-SRS and SE-CARS (discussed in Section 4.3 & 4.4) the generated signal needs to be collected in the transmission-configuration (in CARS, to satisfy the phase matching condition as Stokes, pump and anti-Stokes need to propagate in the same direction and in SRS because the signal generation is on top of the already present Stokes beam). Therefore, in those cases the plasmonic slot length is fabricated with a length of $L_{plas} = 1.8$ μm as in the transmission configuration the forward-collected Raman signal will deteriorate instead of saturate above this length (as shown in Section 3.2.3.1). This can be understood from the fact that in the transmission-configuration the signal collection occurs in the exit waveguide which is situated behind the plasmonic sensor (e.g. see Fig. 4.7) such that if the plasmonic slot length is $> L_{plas}$ any signal generated in the beginning of the plasmonic slot will be offset by the propagation loss if the plasmonic slot length is too long. On the other hand, the signal generated within a length L_{plas} of the exit waveguide will be efficiently captured by the exit waveguide but the excitation field will be attenuated before it reaches this part of the plasmonic slot and hence results in a lower signal collection in the forward direction.

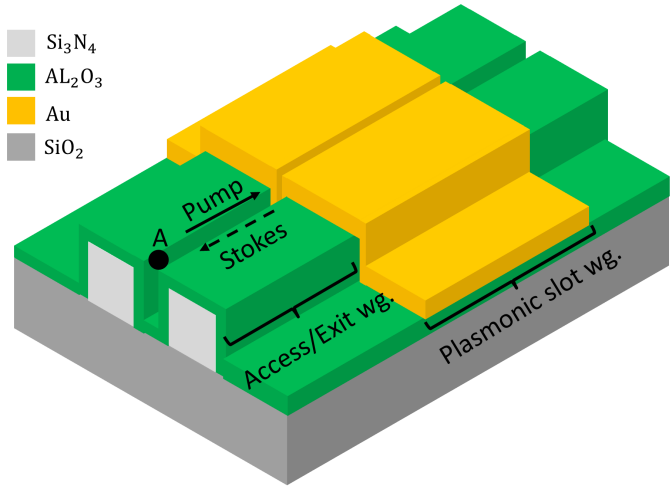


Figure 4.3: Schematic of the nanoplasmonic slot were a fraction of the incoming pump beam is converted in the Stokes beam through the SERS process. The measurement is done in back-reflection (i.e. the exit waveguide from which the Stokes light is collected is the same as the access waveguide that provides the pump beam to the plasmonic slot).

4.1.3 Performance

Table 4.1 (obtained from [16]) shows the simulated Raman conversion efficiency $\eta = P_{Stokes}/P_{pump}$ of the nanoplasmonic slot for a 4-Nitrothiophenol (NTP) monolayer and for a bulk isopropyl alcohol (IPA) cladding. These values are numerically compared to the two NWERS demonstrations on a standard Si_3N_4 strip and slot waveguide (gap=150 nm) together with the effective lengths needed to achieve that result [75]. Note that NTP does not bind to Si_3N_4 such that the presented value is only artificial and serves the purpose of making a comparison with the plasmonic slot. The NTP values for NWERS are calculated by taking the experimental results obtained with IPA and adjusting those for the Raman cross section and the spatial overlap with the analyte.

Structure	P_{Stokes}/P_{pump}	Effective length
Metal slot	$\mathbf{1 \times 10^{-9}}$ 1.2×10^{-10}	≈ 1.8 μm
Si_3N_4 slot [75]	$\mathbf{7 \times 10^{-11}}$ 3.4×10^{-10}	≈ 4 cm
Si_3N_4 strip [75]	$\mathbf{5 \times 10^{-11}}$ 1.1×10^{-10}	≈ 5 cm

Table 4.1: Comparison of the simulated conversion efficiencies (P_{Stokes}/P_{pump}) for SERS on the nanoplasmonic slot waveguide to a Si_3N_4 strip and slot waveguide sensor for an 4-Nitrothiophenol (NTP) monolayer (bold) and for a bulk isopropyl alcohol cladding (italic). Data from [16].

We can see that the nanoplasmonic slot provides an improved (comparable) conversion efficiency for the NTP monolayer (bulk IPA) compared to the two Si_3N_4 waveguides, while drastically reducing the interaction length. This reduction is important for two reasons: firstly, because of their long interactions length the dielectric waveguides cannot accommodate lossy materials as they will induce too much loss. Secondly, the interaction length also has a major effect on the spurious background generation in the Si_3N_4 waveguide core which ultimately sets the detection limit of the Raman sensor. The nanoplasmonic slot itself only has a modest spurious background contribution due to its small interaction length (this effect is partially offset by a larger background generation in the plasmonic slot per unit length due to the presence of higher field intensities). In [45], it is shown that the spurious background contribution of the plasmonic slot configuration is generated in the access waveguide (which can be minimized since it does not contribute to the signal generation unlike the dielectric waveguide Raman sensors) and in the plasmonic slot itself. For access waveguide lengths $\leq$ 60 μm, the

background is mostly generated in the plasmonic slot itself and its background strength is equivalent to the background generated in a 200 μm access waveguide. Therefore, in our calculations we will represent the background contribution of the plasmonic slot as a 200 μm access waveguide, while we assume that the contribution due to the actual access waveguide itself is small enough to not have a major influence. Note that the experimentally captured conversion efficiency for a NTP-coated plasmonic slot corresponds to the simulated value reported in Table 4.1, while the experimentally captured value for IPA on the plasmonic slot is an order of magnitude smaller ($P_{Stokes,IPA}/P_{pump} = 3 \times 10^{-11}$) [45]. This discrepancy could originate from a imperfect fabrication and/or surface wetting issues resulting from the small sputtered slot gap (g= 19 nm) and the hydrophobic nature of the gold surface due to the accumulation of a.o. organic contaminants [103]. In the future, oxygen plasma [104], two-component droplets (with different surface tensions) [103] or ALD plasmonic slots (which have lower aspect ratios) [45] can be employed to ensure a better wetting of the slot. Special care thus needs to be taken with liquid analytes on the plasmonic slot to avoid a mismatch between theoretical estimates and practically measured values.

Now, we can compare the background generation between on-chip SERS and NWERS. Using eq. 3.2.3.1 we can calculate the background ratio of the plasmonic slot compared to the two dielectric waveguide demonstrations of Table 4.1 :

$$\frac{BG_{plas}}{BG_{slot}} = \frac{\eta_{slot}\beta_{SiN} \times P_{pump} \times L_{acc.}}{\eta_{slot}\beta_{SiN} \times P_{pump} \times LF_{slot}} = \frac{L_{acc.}}{LF_{slot}} \approx 0.08 \tag{4.1}$$

$$\frac{BG_{plas}}{BG_{strip}} = \frac{\eta_{slot}\beta_{SiN} \times P_{pump} \times L_{acc.}}{\eta_{strip}\beta_{SiN} \times P_{pump} \times LF_{strip}} = \frac{\eta_{slot} \times L_{acc.}}{\eta_{strip} \times LF_{strip}} \approx 0.013 \tag{4.2}$$

Where β_{SiN} corresponds to the scattering efficiency of the particles in the waveguide core . $\eta_{slot}, \eta_{strip}$ are the conversion efficiencies of background generated in the waveguides core(s) of a Si_3N_4 slot and strip waveguide respectively ($\eta_{strip} = 3 \times \eta_{slot}$ [105]).$L_{acc.}$(= 200 μm) represents the background contribution due to the plasmonic slot and LF_{strip}, LF_{slot} are the length factors that include the effect of the waveguide loss, α_{strip} and α_{slot}, on the background generation (where $LF = e^{-\alpha_p L}(1 - e^{-(\alpha_s - \alpha_p)L})/(\alpha_s - \alpha_p)$). Since the access waveguide length is much shorter the effect of the loss on the background generation is ignored.

From eq. 4.1 & 4.2, we can clearly see that the amount of spurious background due to the waveguide core is reduced in the nanoplasmonic slot waveguide compared to their dielectric waveguide alternatives.

4.1.4 Improvements

Firstly, while the scattering enhancement provided by these plasmonic waveguides is nearly on par with the best surface enhanced Raman spectroscopy (SERS) substrates, this is still not enough to avoid the use of a bulky and expensive deeply cooled camera. Therefore, alternative Raman techniques such as surface-enhanced Stimulated Raman spectroscopy (SE-SRS) & surface-enhanced Coherent anti-Stokes Raman spectroscopy (SE-CARS) can be used as they have the potential to boost the Raman signal by many orders of magnitude which enables room temperature detection. Furthermore, by making use of the nanoplasmonic slot waveguide these techniques avoid a significant spurious background generation in the waveguide core unlike other coherent Raman techniques as SRS which still needs cm-long dielectric waveguides in order to generate a strong Raman response (as will be discussed in Section 4.2).

Secondly, when integrating the plasmonic sensor with a circuit capable of analyzing the Raman scattered light it demands that no further inelastic scattering or photoluminescence occurs in the circuit surrounding the sensor. Unfortunately, the Raman-like background from the silicon nitride waveguides becomes significant over propagation distances that can easily exceed millimeters for spectral filters [17] or AWG. This ultimately limits the concentration of analytes that can be probed because of the shot noise associated to it. Therefore, in practical SERS experiments the access waveguide length before the plasmonic slot is minimized and the measurement is done in back-reflection (see Fig. 4.3). In [16] it is clearly shown that the signal-to-noise ratio was critically dependant on the length of that access waveguide so that it could not be combined with a ”conventional” spectral notch filter such as Bragg gratings [17] placed after the sensor. Furthermore, integrating such a notch filter on the access waveguide will also prevent the excitation field (pump) to reach the plasmonic sensor. In Chapter 5 we will provide an alternative circuit based on a multi-mode interferometer (MMI) to overcome this issue.

4.2 Stimulated Raman spectroscopy (SRS)

4.2.1 Introduction to on-chip SRS

Coherent Raman scattering (CRS) is an alternative Raman technique to spontaneous Raman scattering which holds the potential of sufficiently increasing the Raman signal such that deeply cooled detectors are no longer needed. As discussed in Section 2.2, CRS makes use of two laser sources, a pump laser with a fixed wavelength and a Stokes laser with a varying wavelength such that the difference

frequency between both lasers can resonantly excite the molecular vibrations of the probed analyte. Specifically, in Coherent anti-stokes Raman scattering (CARS) the newly generated anti-Stokes (blue-shifted) frequency is probed while in Stimulated Raman scattering (SRS) the intensity increase (decrease) on the Stokes (pump) laser is detected (see Fig. 4.4). Both techniques have been extensively studied in literature for Raman microscopy [33, 106–112]. The potential of waveguide-enhanced SRS has been explored on hollow-core photonic crystal fibers (HCPCFs) [113–116] using milliwatt CW lasers with the hollow core containing gas (H_2, CO_2). Such a hollow core provides a good modal overlap with the analyte resulting in a strong signal but compared to nanophotonic waveguides, the probing of liquid analytes is complex. Furthermore, on-chip Raman allows for a much more convenient integration with microfluidics [117] and the detection of thin films and monolayers.

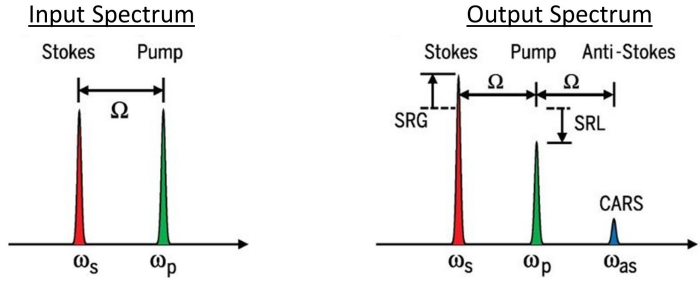


Figure 4.4: In CARS, the newly generated anti-Stokes frequency (ω_{as}) is detected, while in SRS we detect the intensity increase on the already present Stokes laser (ω_s). ω_p is the pump frequency, $\Omega = \omega_p - \omega_s$ and SRG, SRL corresponds to the stimulated Raman gain and loss respectively. [118].

SRS on the waveguide platform has lead to applications such as optical amplification and lasing in silicon waveguides [119–123] (a complete overview of SRS in micro (nano)structures can be found in [124]). Later, SRS enhanced by a nanophotonic integrated circuit has been for the first time demonstrated using a 8 mm long dielectric strip waveguide covered with a droplet of dimethyl sulfoxide (DMSO) and sub-10 mW excitation beams on the chip [19] (see Fig. 4.5 for an illustration). It shows an enhancement of the Raman signal close to five orders of magnitude compared to the NWERS using the same waveguide and pump power ($P_{SRS}/P_{NWERS} \approx 10^5$). The generated Stokes signal is sufficiently strong ($\sim$ nW) to avoid the use of cooled detection and instead is detected by employing a lock-in detection scheme which follows a detection protocol as described in Fig. 4.6 (a detailed discussion on the lock-in setup is presented in Section 6.2.1). The modulation of the pump laser is imprinted as a modulation of the signal beam, in this way the SRS signal can be separated from other varying parameters (e.g. coupling loss to the chip, laser noise, etc.) that influence the Stokes power. The lock-in

amplifier can be regarded as extremely narrow band RC filter that can efficiently reject the local oscillator (DC Stokes contribution) while only detecting the SRS signal (RF Stokes contribution). The lock-in is sensitive to variations in the Stokes power of $\Delta P_S/P_S = 10^{-8}$ with the SRS modulation depth in the order of $\sim 10^{-6}$ (while fluctuations on the DC Stokes power are in the order of $\sim 10^{-3}$, but unlike SRS they are broadband). We can thus see that the improved signal generation of SRS comes at the cost of a more complex detection scheme compared to SERS & NWERS. In Section 4.5, we will give an overview table of the most important advantages and disadvantages of each technique.

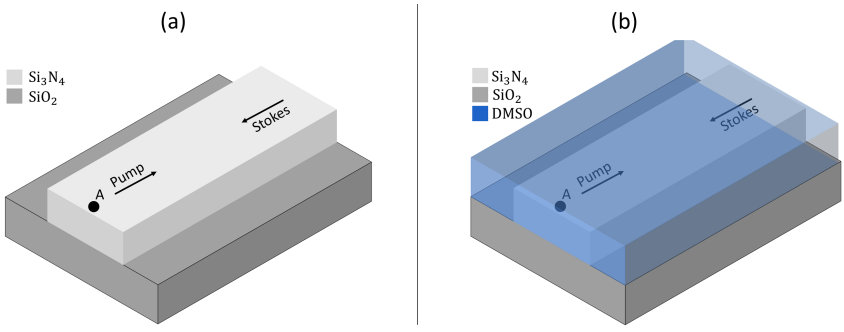


Figure 4.5: (a) Schematic of the SRS interaction on the dielectric strip waveguide were an intensity increase is generated on the Stokes beam in the strip through the SRS process b) illustration with the DMSO top cladding included.

In the case of a rib waveguide, on-chip SRS is preferred over CARS due to its automatic phase-matching while CARS needs complex dispersion engineering in order to achieve phase-matching over a large bandwidth [35]. Silicon nitride waveguides possess a strong chromatic dispersion over the therapeutic window (700 nm-900 nm) which makes it difficult to efficiently generate the anti-Stokes signal for a cm-long waveguide. Even then, challenging fabrication of the strip waveguide is needed (i.e. underetching the waveguide) to tune the waveguide dispersion such that the CARS signal can be efficiently generated. However, this solution is only valid for a gaseous top cladding while efficient phase matching cannot be achieved for liquid top claddings [35]. Since detection of bio-samples in a liquid environment is essential for a Raman sensor, this limits the application of on-chip CARS. The most straightforward solution to satisfy the phase matching condition is to reduce the interaction length (i.e. waveguides of much shorter length) but this inevitably leads to a loss of the generated signal. This is the reason why a technique such as SE-CARS holds potential, as plasmonic slot waveguides only have a short interaction length ($\sim$ µm) while still providing reasonable signal levels. Of course the added dispersion of the plasmonic waveguide can offset any advantage of reducing the interaction length. In Section 4.4, we will make an estimate on the amount of

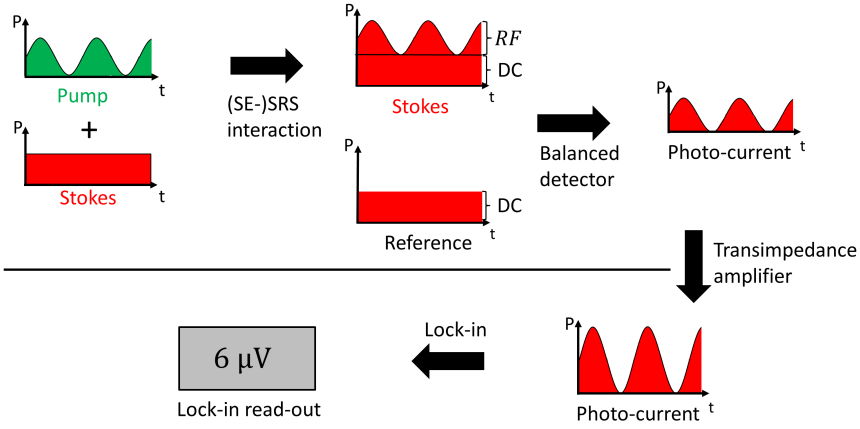


Figure 4.6: Basic lock-in detection scheme used to detect the (SE-)SRS signal. The (SE-)SRS interaction between the analyte, the modulated pump beam and Stokes beam imprints a modulated RF increase on the Stokes beam. A balanced detector combined with a reference beam is then used to remove the majority of the DC contribution of the Stokes beam. Afterwards, the remaining Stokes wave is send to a transimpedance amplifier. Finally, the lock-in filters out the remaining DC contribution and detects the (SE-)SRS signal.

phase-mismatch encountered when performing SE-CARS on the nanoplasmnic slot.

4.2.2 Improvements

The on-chip SRS demonstration makes use of a dielectric strip waveguide which is accompanied by a significant photoemission background that ultimately hinders the detection of low analyte concentrations or analytes with weak Raman scattering cross-sections [35, 64, 87]. This background can be reduced by putting a larger amount of the optical field into the analyte rather than in the core material while still maintaining proper guiding. Alternative waveguide geometry such as slot or a plasmonic waveguides (see Section 4.3) can on one hand provide an enhancement in the signal level while on the other reducing the Si_3N_4 background by having a reduced interaction length/less overlap with the waveguide core.

Note that the Si_3N_4 background in SRS does not set a detection limit through its shot noise contribution but through the difficulty to distinguish it from the Raman signal. To better understand this, let us consider that in SRS the stimulation with a Stokes beam allows for uncooled detection but also induces an extra shot noise term

to the system. It is this shot noise due to the Stokes laser that is the fundamental noise source for SRS instead of that due to the dielectric waveguide core [35]. This is an advantage of SRS over NWERS (of course coming at the cost of a more complex detection scheme) since it results in a stronger scaling of the signal-to-noise ratio (SNR) with power (P) (for SRS the SNR scales as $P^{3/2}$ compared to $P^{1/2}$ for NWERS, with $P_{Stokes} = P_{pump} = P$) [35]. This allows for a greater improvement of the SNR when optimizing the coupling efficiency, the modulation depth and waveguide geometry. However, the dielectric Raman background is still important since for low Raman signal levels this background is a few orders of magnitude stronger than the signal itself. This makes it very challenging to separate the signal from the background as it is difficult to build a statistical model for a strong background such that it can be accurately subtracted to re-discover the Raman spectra of the analyte. Since there are multiple noise sources i.e. laser shot noise, small sample variations, inaccuracy in normalization which causes a difference between the predicted background and the real value, they all need to be accounted for in the background estimate which is challenging. Therefore, the presence of the dielectric waveguide background still sets a limit on the minimal detection levels which can only be improved by reducing the background itself.

4.3 Surface-enhanced Stimulated Raman spectroscopy (SE-SRS)

4.3.1 Introduction to on-chip SE-SRS

In Section 4.1, we discussed that the nanoplasmonic waveguide turned out to be advantageous over dielectric waveguides as they do not suffer from the significant photoemission background that is characteristic of dielectric materials. But while the scattering enhancement provided by these plasmonic waveguides is nearly on par with the best SERS substrates, this is still not enough to avoid the use of an expensive deeply-cooled camera. Stimulation with a Stokes beam can boost the Raman signal by many orders of magnitude. This has been demonstrated on plasmonic nanoparticles using SE-SRS. The Van Duyne group was the first to successfully perform a SE-SRS measurement using gold nanoparticles embedded in a silica shell (for thermal stability) and ultrafast pulses. They showed an intensity enhancement of about 10^6 compared to normal SRS [59]. However, they suffered from the use of high energy pulses causing photodegradation of their analyte (DPE) and metallic surfaces. In order to overcome this issue, alternative detection schemes using higher modulation frequencies, chirped pulses, CW beams and data analysis have been proposed [60, 125–127]. Recently, detection of single molecules on silica coated gold nanoparticles has been reported while avoiding sample photodamage [61]. At the moment of this writing, SE-SRS studies have been mainly focus around gold nanoparticles and not yet been demonstrated on a waveguide platform. In comparison to on-chip SRS, the use of a nanoplasmonic waveguide instead of a dielectric one to perform SE-SRS promises major advantages. Indeed, the combination of plasmonic enhancement and a rather long interaction length (in comparison to the interaction length in common SERS substrate) leads to a large Raman scattering enhancement [16] so that the required pump and Stokes fields can be of very low power and therefore integration on a Photonic Integrated Circuit (PIC) becomes easier. Moreover, we know that dielectric waveguides suffer from a broandband photoemission that superposes itself onto the Raman spectrum while gold plasmonic waveguides suffer far less from it [15, 16, 105, 128]. Therefore, the obvious next step in expanding the Raman on-chip portfolio is to explore waveguide based SE-SRS.

4.3.2 Estimating the SE-SRS signal

First, we start by making a theoretical estimate on the strength of the SE-SRS signal. Equivalent to the on-chip SRS demonstration we assume a modulated pump beam

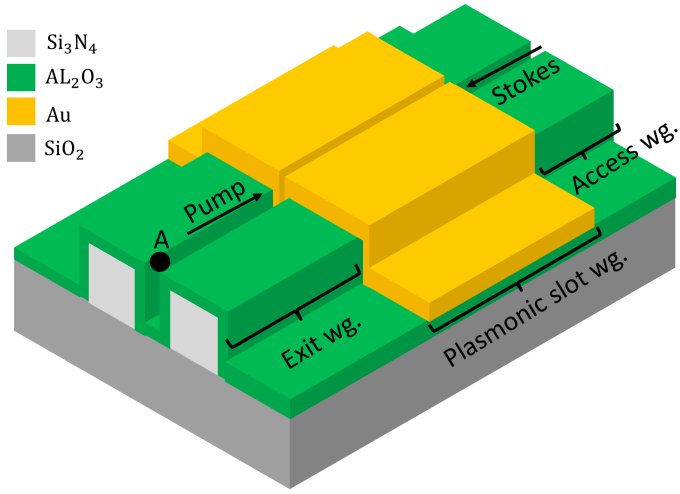


Figure 4.7: Schematic of the SE-SRS interaction on the nanoplasmonic slot waveguide. An intensity increase is generated on the Stokes beam in the plasmonic slot (i.e. the gold-coated slot) through the SE-SRS process were a monolayer (4-NTP in present estimate) is coated on top of the plasmonic slot. The pump and Stokes beam are counterpropagating in order to ease the separation of both beams prior to detection.

that imprints a shallow modulation on the Stokes beam through the SE-SRS gain process which is picked up by a lock-in amplifier (see Fig. 4.6). Since we try to measure this intensity increase on the Stokes beam, it needs to be transmitted through the plasmonic slot such that it can be detected by the lock-in. From Section 4.1 we know that the optimal length for the plasmonic slot in that case corresponds to $\approx$ 1.8 µm which is the length we assume in our estimate. We take a plasmonic slot with a gap of 19 nm, a gold thickness in the gap of 5.5 nm and a total width of 0.88 µm and a monolayer of 4-Nitrothiophenol (NTP) as the analyte (Section 6.1.1 provides a detailed discussion and illustration on these plasmonic slots). We assume a 2 mm long dielectric slot waveguide prior and after the plasmonic slot that functions as an access/exit waveguide and the low absorption loss in these dielectric waveguides is neglected. Note that this length of 2 mm on both sides is the practical length used in the experimental SE-SRS measurement (see Chapter 6) and is set by i.a. the sample holder dimensions. A guided pump and Stokes power of 1 mW and 4 mW are chosen respectively (with 1 mW being the modulated pump power) and both beams being CW with power levels sufficiently low to prevent damage to the sample and the analyte [45]. Just like the on-chip SRS demonstration we take the pump and Stokes beams to be counterpropagating in order to ease the separation of both beams prior to detection.

In order to make a fair comparison to the other methods of on-chip Raman spectroscopy (i.e. SERS, SRS, & SE-CARS), the difference in generated signal between them is evaluated at the end of the dielectric exit waveguide before the light is coupled out into free-space (Point A in Fig. 4.7). In all cases we will assume the same analyte (NTP) and a guided pump and Stokes power of 1 mW and 4 mW (not present for SERS). The plasmonic slot length in the SERS calculation is equivalently chosen as 1.8 μm-long and a lossless 8 mm-long dielectric strip is taken for the SRS calculation.

Now using Eq.3.46, the generated SE-SRS power in the plasmonic slot waveguide and coupled out to the dielectric exit waveguide (Fig. 4.7) can be written as:

$$P_{SESRS} = g_{plas} \times (\gamma_p P_P) \times (\gamma_s^2 P_S) \times L_{eff} \tag{4.3}$$

where g_{plas} is the stimulated Raman gain (g_{stim}) for the plasmonic slot specifically, (P_P, P_S) are the guided pump and Stokes power in the dielectric exit/access waveguide prior to the plasmonic slot and (γ_p, γ_s) are the respective coupling losses from dielectric to plasmonic waveguide for both lasers. Since the frequency of the generated SE-SRS signal P_{SESRS} corresponds to that of the Stokes laser, the loss γ_s is squared in order to account for the coupling-loss P_{SESRS} experiences (due to the mode-mismatch) when being coupled out from to plasmonic slot to the exit waveguide (see Fig. 4.7). The effective interaction length L_{eff} accounts for the loss experienced in the nanoplasmonic waveguide for both the pump and the Stokes beam $L_{eff} = (1 - \exp(-\alpha_p L_{plas}))/(\alpha_p) \times \exp(-\alpha_s L_{plas})$ with (α_p, α_s) the linear absorption losses in the plasmonic slot and L_{plas} is the plasmonic slot length. Recall from eq.3.24 that g_{plas} [$(Wm)^{-1}$] (expressed at the maximum of the Raman resonance) is related to the spontaneous Raman gain G_{plas}:

$$\frac{G_{plas,NTP}}{g_{plas,NTP}} = \frac{\hbar\omega_s}{4}\Delta_{R,NTP} \tag{4.4}$$

with ω_s the Stokes frequency, $\Delta_{R,NTP}$ the FWHM of the Raman response of NTP and $G_{plas,NTP} = \eta_{plas}\sigma_{NTP}\rho_{NTP}$ with σ_{NTP} the differential Raman cross section, ρ_{NTP} the molecular density of NTP and η_{plas} the Raman conversion efficiency of the plasmonic slot (which accounts for the plasmonic enhancement at the Stokes and pump frequency).

From [16], we know that $G_{plas,NTP} = 0.0082\ m^{-1}$ for the spontaneous Raman measurement of a monolayer of NTP covering the plasmonic slot waveguide.

Now combining Eq. 4.3 and Eq. 4.4, we can calculate $P_{SESRS,NTP}$. Filling in $L_{plas} = 1.8$ μm, $\alpha_p = \alpha_s = -2.4$ dB/μm, $\gamma_p = \gamma_s = -4$ dB and $P_p = 1$mW, $P_s = 4$

mW, $\Delta_{R,NTP} = 2\pi \times 660$ GHz and $\hbar\omega_{s,NTP} = 1.44$ eV gives $g_{plas,NTP} = 3.5 \times 10^4$ $(Wm)^{-1}$ and an estimated SE-SRS signal of $P_{SESRS,NTP} = 3.52$ nW.

Let us now make a comparison of the background generation in the different Raman techniques by calculating the signal-to-background ratio (SBR).

In Chapter 6, we demonstrate a successful SRS measurement on a 4 mm long Si_3N_4 slot waveguide. From that measurement we can extract the stimulated Raman gain of this dielectric slot $g_{slot} = 0.15$ $(Wm)^{-1}$ which we can now use to calculate the amount of background generation in the 2 mm long exit waveguide. Note that we do not consider the background generation in the access waveguide since any background generated there will have to pass through the plasmonic slot before it can couple out. Therefore, it will have an additional 12 dB of extra loss compared to the background generated in the exit waveguide of the same length. We can now fill in:

$$P_{BG',SESRS'} = g_{slot} \times P_P \times (\gamma_{slot} P_S) \times L_{exit} = 7.13 \times 10^{-11} W \tag{4.5}$$

Where $P_p = 1$mW, $P_s = 4$ mW, $\gamma_{slot} = -12$dB and $L_{exit} = 2$mm. We can now calculate that:

$$SBR_{SESRS'} = \frac{P_{SESRS,NTP}}{P_{BG',SESRS'}} = 49 \tag{4.6}$$

This SBR value is artificially low because of the 2 mm long access/exit waveguides that we assumed in our calculations. In a practical experiment, these long waveguides are required since the chip dimensions need to be large enough (> 4 mm) for the laser light to efficiently couple into the waveguide on both sides and not onto the sample holder (in waveguide based SE-SRS, the light needs to be coupled/-collected on both sides of the chip since it is transmitted trough the plasmonic slot). Furthermore, the chip facet cannot be placed close to the plasmonic slot such that the exit waveguide (which generates most of the background) can be minimized, as the access waveguide length on the other side of the plasmonic slot is only 2 mm long in our current design which would make the chip too small to fit on the sample holder. In the future, an improved chip design and/or a custom made sample holder can be made to further ease the requirements of the chip dimension. Since the access/exit waveguides do not have any sensing contribution and are only there for practical reasons, we will also calculate SBR_{SESRS} for $L_{exit} = 10$ μm (corresponding to the smallest length in practical SERS measurements on the plasmonic slot) which presents a better estimate for SBR_{SESRS}. In that regime, the background due to the plasmonic slot itself ($P_{SESRS,BG}$) will be the main background contributor, such that:

$$SBR_{SESRS} = \frac{P_{SESRS,NTP}}{P_{SESRS,BG}} = \frac{g_{plas,NTP}}{g_{plas,BG}} = \frac{G_{plas,NTP} \times \Delta_{R,BG}}{G_{plas,SiN} \times \Delta_{R,NTP}} \times \frac{4}{2\pi} \quad (4.7)$$

$$= SBR_{SERS} \times \frac{\Delta_{R,BG}}{\Delta_{R,NTP}} \times \frac{2}{\pi} = 76 \quad (4.8)$$

where $SBR_{SERS} = 18$ (see Section 4.3.3) and $\Delta_{R,BG} = 2\pi \times 4500$ GHz, corresponding to the full width of the NTP mode. In the remainder of this chapter we will use this value to compare to the other Raman techniques.

4.3.3 Comparing on-chip SE-SRS to SERS

Comparing SE-SRS to the situation without stimulation (SERS), the amount of light scattering $P_{SERS,NTP} = 1 \times 10^{-12}$W. This value is calculated from $\zeta = P_{SERS,NTP}/P_P = 1 \times 10^{-9}$ for 1 mW of pump power (see Section 4.1). The increase in Stokes power brought by the stimulation is therefore $P_{SESRS,NTP}/P_{SERS,NTP} = 3521$.

Next, we calculate SBR_{SERS} where we include the amount of background generation in the plasmonic slot (which is equivalent to the background generated in a 200 μm access waveguide before the plasmonic slot, as discussed in Section 4.1.3) and assume that the actual background contribution of the access waveguide (~ 10 μm) is small enough to be ignored:

$$SBR_{SERS} = \frac{P_{SERS,NTP}}{\frac{1}{2} \times \beta_{SiN} \times \eta_{slot} \times P_p \times L_{acc}} = 18 \quad (4.9)$$

Where $\beta_{SiN} = 0.55 \times 10^{-9}[sr^{-1}][cm^{-1}] \times \frac{120[cm^{-1}]}{2[cm^{-1}]} = 3.3 \times 10^{-8}[sr^{-1}][cm^{-1}]$ [105] corresponds to the scattering efficiency of the particles in the waveguide core over the full NTP mode, $\eta_{slot} = 0.168$ [64] corresponds to the slot waveguide Raman background conversion efficiency and L_{acc} is the access waveguide length. Note that the SBR_{SERS} calculated here is an upper bound, since the experimental captured SBR_{SERS} is closer to unity [88].

4.3.4 Comparing on-chip SE-SRS to SRS

In Section 4.2, we discussed that on-chip SRS has been reported on dielectric strip waveguides using a liquid analyte of DMSO. In that case $P_{SRS,DMSO} = 0.92$ nW

for a guided pump and Stokes power of P_P = 1 mW and P_s = 4 mW respectively and a 8 mm long dielectric strip waveguide.

A direct comparison between on-chip SE-SRS and on-chip SRS is challenging since we cannot perform a SRS measurement using a NTP-covered dielectric waveguide because the binding between NTP and silicon nitride is difficult. Instead, we make an theoretical estimate by calculating the SRS signal that would occur if we replaced bulk DMSO by a monolayer of NTP ($P_{SRS,NTP}$).

From [16], we know that for a strip waveguide the ratio between the spontaneous Raman gain for liquid Isopropanol (IPA) (G_{IPA}) and the spontaneous Raman gain for a monolayer of NTP (G_{NTP}) is equal to:

$$\frac{G_{IPA}}{G_{NTP}} = 2.2 \tag{4.10}$$

We can now trivially compare the DMSO and IPA spontaneous gains as they are both liquid analytes filling the upper optical cladding of the dielectric waveguide so that $\eta_{DMSO} = \eta_{IPA}$. The ratio of their spontaneous Raman gains is then simply calculated from the ratio of their corresponding Raman cross sections and molecule densities $\sigma_{DMSO} = 3.1 \times 10^{-30}$ [cm^2/sr], $\sigma_{IPA} = 8.0 \times 10^{-31}$ [cm^2/sr] and ρ_{DMSO} = 1, 1 [g/cm^3] and ρ_{IPA} = 0.786 [g/cm^3].

$$\frac{G_{DMSO}}{G_{IPA}} = \frac{\sigma_{DMSO}\rho_{DMSO}}{\sigma_{IPA}\rho_{IPA}} = 5.42 \tag{4.11}$$

That leads to $G_{DMSO} = 11.924 \times G_{NTP}$ for the strip waveguide.

Using Eq. 3.24, the ratio between the stimulated Raman gain for DMSO and a hypothetical monolayer of NTP covering the dielectric strip is:

$$\frac{g_{NTP}}{g_{DMSO}} = \frac{G_{NTP}}{G_{DMSO}} \frac{\omega_{s,DMSO}\Delta_{R,DMSO}}{\omega_{s,NTP}\Delta_{R,NTP}} = 0.0539 \tag{4.12}$$

where we use $\lambda_{s,DMSO}$ = 828.5 nm and $\Delta_{R,DMSO} = 2\pi \times 400$ GHz. The resulting decrease in Stokes power by using Eq. (4.3) is $P_{SRS,NTP}$ = 0.05 nW. This allows for a direct comparison between dielectric waveguide-based SRS and plasmonic waveguide-based SE-SRS for the same analyte: $P_{SESRS,NTP}/P_{SRS,NTP}$ = 70.4.

We can now simply calculate SBR_{SRS} as the ratio of:

$$SBR_{SRS} = \frac{g_{NTP}}{g_{strip}} = 0.01 \tag{4.13}$$

Where $g_{strip} = 3 \times g_{slot} = 0.45$ [Wm]$^{-1}$ and $g_{NTP} = 0.005$ [Wm]$^{-1}$.

4.4 Surface-enhanced Coherent anti-Stokes Raman spectroscopy (SE-CARS)

4.4.1 Introduction to on-chip SE-CARS

Just as SE-SRS, on-chip SE-CARS holds the potential of generating a sufficiently strong signal level such that deeply-cooled detection is not needed while at the same time reducing the photoemission background as the sensing is done by only a short ($\sim$ µm) plasmonic waveguide instead of a cm-long dielectric waveguide (see Fig. 4.8 for an illustration). SE-CARS has been extensively studied in literature in order to increase the detection sensitivity of coherent Raman techniques [57, 58, 61, 129–135] and single molecule detection levels has been reported [56, 136, 137]. At the moment of this writing, SE-CARS has not yet been demonstrated on a nanoplasmonic waveguide. As mentioned in Section 4.2, on-chip CARS is limited by the need to fulfill the phase matching condition which is challenging for a cm-long silicon nitride waveguides. Recall that the (SE-)CARS signal is proportional to the phase mismatch factor [34]:

$$I(\omega_{as}) \propto \mathrm{sinc}^2(\frac{\Delta kL}{2}) \tag{4.14}$$

Where $\Delta k = 2k_p - k_s - k_{as}$, L is the interaction length and the assumption is made that we work in the low gain region such that the linear phase mismatch is stronger than the nonlinear phase mismatch contribution ($\Delta k >> \gamma P_{pump}$ [38]). SE-CARS hold the potential of relaxing this condition by using plasmonic waveguides which reduces the interaction length by four orders of magnitude but on the other hand also increases Δk.

In order to get an idea about the impact of the nanoplasmonic slot on the phase matching condition let us first simplify Δk by expressing the wave-vectors of the (anti-)Stokes beams as a Taylor expansion at the pump frequency (ω_p) from [35]:

$$k(\omega_s) = k(\omega_p) + \beta_1(\omega_s - \omega_p) + \frac{\beta_2}{2}(\omega_s - \omega_p)^2 + ... \tag{4.15}$$

$$k(\omega_{as}) = k(\omega_p) + \beta_1(\omega_{as} - \omega_p) + \frac{\beta_2}{2}(\omega_{as} - \omega_p)^2 + ... \tag{4.16}$$

with $\beta_i = \frac{\partial^i k}{\partial \omega^i}$ the i-th order chromatic dispersion of k at the pump frequency such that:

$$-\Delta k = \beta_2 \Delta\omega^2 + \frac{\beta_4}{12}\Delta\omega^4 + ... \tag{4.17}$$

where $\Delta\omega = \omega_{as} - \omega_p = \omega_p - \omega_s$ and in a first-approximation we can assume that the disperisive property of a waveguide is described by $\beta_2\Delta\omega^2$ such that:

$$I(\omega_{as}) \propto \mathrm{sinc}^2(\frac{|\beta_2|\Delta\omega^2 L}{2}) \tag{4.18}$$

For most Raman modes of practical bio-analytes $\Delta\omega$ varies from 500 cm^{-1} to 1500 cm^{-1} (15 THz-45 THz) such that we calculate how well the phase matching condition is fulfilled for the plasmonic slot at the maximum $\Delta\omega_{max}$ = 45 THz. Since β_2 corresponds to the group velocity dispersion at the pump frequency we can calculate it by using a simulation mode of the plasmonic slot (using Lumerical mode solutions) and extract the phase matching factor from it:

$$\mathrm{sinc}^2(\frac{|\beta_2|\Delta\omega_{max}^2 L}{2}) = 0.998 \tag{4.19}$$

with $\beta_2 = 4.57 \times 10^4$ ps^2/km, L = 1.8 μm. We can see that by using the plasmonic slot (and hence reducing the interaction length) SE-CARS signal is efficiently generated since we are no longer limited by a large phase mismatch. In the following subsections we make a calculation on the expected signal strength of SE-CARS and compare the results to SE-SRS.

4.4.2 Estimating the SE-CARS signal

In Chapter 2, we derived the following formula for (SE-)CARS signal generation in terms of intensity (I) (Eq. 2.43):

$$I(\omega_{as}) = \frac{9\omega_{as}^2}{4n_p^2 n_{as} n_s c^4 \varepsilon_0^2}|\chi_R|^2 I_p^2 I_s L^2 \mathrm{sinc}^2(\frac{\Delta k L}{2}) \tag{4.20}$$

Let us now define the effective area as:

$$A_{eff} = \frac{(\int |E|^2 dA)^2}{\int |E|^4 dA} \tag{4.21}$$

such that we can calculate the amount of $P(\omega_{as})$ generated in the plasmonic slot and coupled out to the exit waveguide as:

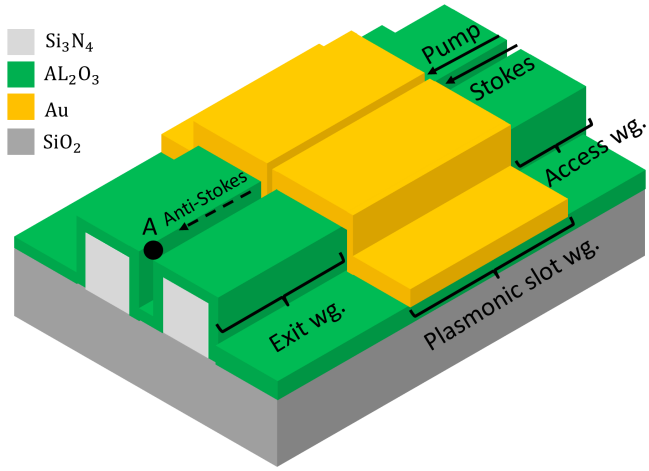


Figure 4.8: Schematic of the SE-CARS interaction on the nanoplasmonic slot waveguide were an anti-Stokes beam is generated in the plasmonic slot through the SE-CARS process.

$$P_{SECARS} = \gamma_{as} \frac{9\omega_{as}^2}{4n_p^2 n_{as} n_s c^4 \varepsilon_0^2 A_{eff,plas}^2} |\chi_{R,plas}|^2 (\gamma_P P_p)^2 (\gamma_s P_s) L_{plas}^2 \text{sinc}^2\left(\frac{\Delta k_{plas} L_{plas}}{2}\right) \tag{4.22}$$

Which results in $P_{SECARS} = 8.86 \times 10^{-14}$W with $\gamma_p \approx \gamma_s \approx \gamma_{as} = -4$dB, $A_{eff,plas} = 0.0078\mu m^2$, $L_{acc} = 1.8\mu$m, $\text{sinc}^2(\frac{\Delta k_{plas} L_{plas}}{2}) = 0.9$, $P_s = 4$ mW, $P_p = 1$ mW and $\chi_{R,plas}(= \Delta\text{Im}\{\chi_{plas}\}$ at resonance) is calculated by filling in Eq. 2.57:

$$\Delta\text{Im}\{\chi_{plas,NTP}\} = \frac{g_{I,plas} \times n_{s,plas}^2 \times c^2 \times \varepsilon_0}{3 \times \omega_s} = 3.8 \times 10^{-19} m^2/V^2 \tag{4.23}$$

Where we use $g_{I,plas} = g_{plas,NTP} \times A_{eff,plas} = 2.73 \times 10^{-10}$[m/W].

We can see that SE-CARS on the plasmonic slot only generates a limited signal $P_{SECARS} = 8.86 \times 10^{-14}$W. One alternative to further enhance the signal generation is by using a pulsed pump laser since the signal power scales quadratically with the pump which leads to an additional enhancement while the average power remains reasonable . In our research lab, a Onefive Katana pulsed laser is available with a pulse duration of $t_{pulse} = 50$ ps and a pulse repetition rate of $f_{rep} = 1$MHz. From experimental verification it is found that an average power of $P_{avg} = 1$mW can be coupled into the access waveguide which corresponds to a peak power of $P_{peak} = 18$ W (while P_s is still CW). Such that we can calculate:

$$P_{SECARS,pulsed} \propto t_{pulse} f_{rep} |\chi_{R,plas}|^2 (\gamma_P P_{peak})^2 (\gamma_s P_s) \tag{4.24}$$

Resulting in $P_{SECARS,pulsed} = 1.9 \times 10^{-9}$W which is in the same order of magnitude as P_{SESRS} (and has a more straightforward detection since it is generated at a new anti-Stokes frequency). However, note that in Eq.4.22, 4.24 for simplicity we didn't include the effect of the absorption loss in the plasmonic slot (α_{plas}) on the generated SECARS signal. This loss could easily introduce $>$ 5dB extra penalty on both P_{SECARS} & $P_{SECARS,pulsed}$.

Next, we calculate the amount of background generation in the 2-mm long access waveguide. Using $g_{slot} = 0.15$ Wm^{-1} and $A_{eff,slot} = 0.36$ μm^2 it follows that $g_{I,BG'} = g_{slot} \times A_{eff,slot} = 5.4 \times 10^{-14}$[m/W] and we can calculate that:

$$Im\{\chi_{slot}\} = \frac{g_{I,slot} \times n_{s,slot}^2 \times c^2 \times \varepsilon_0}{3 \times \omega_s} \tag{4.25}$$

Which corresponds to $Im\{\chi_{slot}\} = 1.7 \times 10^{-23}$m^2/V^2. Where we calculate the background generation as:

$$P_{BG'}(\omega_{as}) = \gamma_{slot} \frac{9\omega_{as}^2}{4n_p^2 n_{as} n_s c^4 \varepsilon_0^2 A_{eff,slot}^2} |Im\{\chi_{slot}\}|^2 P_p^2 P_s L_{Acc}^2 sinc^2(\frac{\Delta k_{slot} L_{acc}}{2}) \tag{4.26}$$

Which results in $P_{BG'} = 5 \times 10^{-23}$W with $\gamma_{slot} = -12$dB, $A_{eff,slot} = 0.36\mu m^2$, $L_{acc} = 2$mm, $sinc^2(\frac{\Delta k_{slot} L_{acc}}{2}) = 6 \times 10^{-6}$, $P_s = 4$ mW and $P_p = 1$ mW. We can see that the inefficient CARS generation in dielectric waveguides limits the amount of background generation in the dielectric core. Additionally, we can also calculate the non-resonant background generation in the access waveguide $\chi_{NR} = 1.96 \times 10^{-21}$m^2/V^2 [138] which corresponds to $P_{BG',NR} = 6.5 \times 10^{-19}$W. Finally, we calculate SBR_{SECARS} using the background generated in the plasmonic slot itself at the frequency where the NTP peak is generated ($P_{SECARS,BG}$). This background contribution will be dominant, since there will be an efficient signal generation in the short plasmonic slot. Using Eq. 4.22 and 4.23, we now get:

$$SBR_{SECARS} = \frac{P_{SECARS,NTP}}{P_{SECARS,BG}} = \frac{|\Delta Im\{\chi_{plas,NTP}\}|^2}{|Im\{\chi_{plas,BG}\}|^2} \tag{4.27}$$

$$= \frac{g_{plas,NTP}^2}{g_{plas,BG}^2} = SBR_{SESRS}^2 = 5.78 \times 10^3 \tag{4.28}$$

Note, that we did not include the effect of $\mathrm{Re}\{\chi_{BG}\}$ and the non-resonant background which might lead to a further background contribution and distortion of the captured signal. We can thus see that $SBR_{SECARS} = SBR^2_{SESRS}$, which leads to an improved SBR for SE-CARS when probing high concentration analytes ($SBR_{SECARS} > 1$), but to an improved SBR for SE-SRS when probing low concentration analytes ($SBR_{SECARS} < 1$). The latter case is more relevant for us as biological samples typically have low concentrations.

4.4.2.1 Two pulsed lasers sources

Until now in our discussion about SE-CARS on the nanoplamonic slot we have set ourselves the limit of the practical available pulsed laser source in our research lab. However, in order to discuss the full potential of SE-CARS on the nanoplasmonic slot we can further expand our signal estimation to the case of two ps-pulsed laser sources (i.e. a pulsed pump and Stokes lasers). We will set the pulse length to $t_{pulse,2} = 10$ ps (in order to ensure that the pulse linewidth is smaller than that of the vibrational mode), take a pulse repetition rate of $f_{rep,2} = 1$ MHz and assume that both pulsed lasers have an average power of $P_{avg} = 1$ mW (corresponding to the power range for CW lasers for which no thermal damage is induced on the plasmonic slot and the analyte). Therefore, the peak pulse power corresponds to $P_s = P_P = P_{Peak,2} = 100$ W ($P_{Peak,2} = P_{avg}/(f_{rep,2}t_{pulse,2})$) and we can calculate that:

$$P_{SECARS,pulsed2} \propto t_{pulse,2} f_{rep,2} |\chi_{R,plas}|^2 (\gamma P_{peak,2})^3 \tag{4.29}$$

such that $P_{SECARS,pulsed2} = 2.4 \times 10^{-4}$W. We can thus see that by making use of two pulsed laser (with an optimized pulse time) there is an additional five orders of magnitude increase in signal generation compared to $P_{SECARS,pulsed}$ with only one pulsed laser. Analogously, we can make a comparable calculation for SE-SRS on the nanoplasmonic slot using the same two pulsed lasers sources such that:

$$P_{SESRS,pulsed2} \propto t_{pulse,2} f_{rep,2} g_{plas} (\gamma P_{peak,2})^2 L_{eff} \tag{4.30}$$

and we can calculate that $P_{SESRS,pulsed2} = 1 \times 10^{-4}$W. Showing that we can get an additional five orders of magnitude signal enhancement for both SE-SRS and SE-CARS by making use of two 10 ps-pulsed lasers. As mentioned before, the estimate of $P_{SECARS,pulsed2}$ is an overestimate since the loss in the plasmonic slot is not included in the calculation and could easily introduce an extra > 5 dB loss penalty.

Note that the calculation above does not account for the fact that pulsed lasers will induce thermal damage (in the plasmonic slot) at lower average power levels than the CW lasers. This can be understood from the fact that for the same average

power, pulsed lasers will reach much higher maximum optical power levels than CW lasers. The heat absorption in the plasmonic slot is quasi-instantaneous and mostly happens in the tiny gold layer inside the plasmonic slot (since the optical mode is the strongest there, see Section 6.4). Therefore, over the pulse time ($\sim$ 10 ps) there will be a strong temperature increase in this gold layer (with the max at the end of the pulse) since there is not enough time for this thermal energy to dissipate in the surrounding materials (there will only be a very local heat dissipation). Where the thermal time constants for heat dissipation in the plasmonic slot correspond to τ_1 = 0.023 µs and τ_2 = 46.9 µs (see Section 6.4.4 for additional details). Therefore, for the same average power, pulsed lasers will induce more thermal damage in the plasmonic slot compared to CW lasers since they reach a higher maximum temperature in the plasmonic slot. The first preliminary measurements with the 50 ps Katana pulsed laser on the plasmonic slot (with a 1 MHz rep rate) seem to confirm this theory as thermal damage was induced to the slot (resulting in a huge increase in its transmission loss) at lower average power levels as those determined for CW lasers ($<$ 10 mW). Where an exact power vs. transmission experiment still needs to be performed. Furthermore, thermal damage on both the plasmonic substrate and analyte have been reported for picosecond pulses in plasmon enhanced coherent Raman experiments [59]. Additionally, multi-Watt 10 ps wavelength tunable (fourier transform) laser sources are pretty uncommon.

The value of P_{avg} = 1 mW in the calculation above is based on the practically determined power limit for CW lasers beyond which photoreduction of the probed analyte (NTP in this case) starts to occur. Therefore, in experimental SE-CARS (SE-SRS) measurements with pulsed laser sources the practical attainable signal levels will be lower than the estimates made here. Since the generated signal levels are critically depended on that pulse power in the plasmonic slot, a reduction in P_{avg} by an order of magnitude (or more) would result in a reduction of $P_{SECARS,pulsed2}$ ($P_{SESRS,pulsed2}$) by three (two) orders of magnitude because of its cubic (quadratic) power dependence. Therefore, when using pulsed laser sources there will be a trade-off between an increased signal generation on one hand and the additional thermal damage induced by those sources on the other hand. In the future, it will be best to first determine the practical power limit for pulsed lasers on the plasmonic slot (for both the analyte and the plasmonic slot itself) before proceeding to the actual SE-CARS (SE-SRS) measurements with pulsed lasers. Furthermore, SE-SRS on the plasmonic slot suffers from a thermal background originating from a thermo-optic effect (as will be discussed in Chapter 6). Therefore, any additional temperature increase induced in the plasmonic slot by using pulsed lasers needs to be handled with care in order to ensure that it does not lead to an increase in the strength of this spurious thermal background.

4.4.3 Comparing on-chip SE-SRS to SE-CARS

We first summarize the calculations made in Section 4.4.2.

- When using CW-lasers on the nanoplasmonic slot, the generated SE-CARS signal levels are predicted five order of magnitude smaller than the SE-SRS signals ($P_{SECARS} = 8.86 \times 10^{-14}$W vs. $P_{SESRS} = 3.52 \times 10^{-9}$W). For such low signal levels, deeply-cooled detectors are still required and the added complexity of SE-CARS does not provide much advantage over waveguide-based SERS (as they reach comparable signal levels).
- If we employ a pulsed pump laser for SE-CARS, the generated SE-CARS signals are in the same order as P_{SESRS} ($P_{SECARS,pulsed} = 1.9 \times 10^{-9}$W). Although (as discussed in the previous section), we did not include the effect of the absorption loss in the plasmonic slot on the generated SE-CARS signal which could easily introduce $>$ 5dB extra penalty on $P_{SECARS,pulsed}$.
- If we employ a pulsed pump and Stokes laser for SE-CARS and SE-SRS (with only a single pulsed laser present in our research lab), the generated SE-CARS and SE-SRS signals can be enhanced by an additional five orders of magnitude ($P_{SECARS,pulsed2} = 2.4 \times 10^{-4}$ W & $P_{SESRS,pulsed2} = 1 \times 10^{-4}$ W). However, such high signal enhancements are not attainable in practice since the calculation does not account for the fact that pulsed lasers will induce thermal damage (to the plasmonic slot and analyte) at lower average power levels then the CW lasers. Therefore, when using pulsed laser sources there will be a trade-off between an increased signal generation on one hand and the additional thermal damage induced by those sources on the other hand (which will limit the practical attainable enhancement). In future projects, firstly the practical power limit for pulsed lasers on the plasmonic slot needs to be determined in order to make a fair estimate on the actual attainable enhancement for pulsed plasmonic enhanced coherent Raman scattering.
- Comparing the SBR for both plasmonic-enhanced coherent Raman techniques, we can see that $SBR_{SECARS}/SBR_{SESRS} = 76$. SE-CARS has an improved SBR compared to SE-SRS, but this will only hold as long as $SBR_{SECARS} > 1$ (since $SBR_{SECARS} = SBR_{SESRS}^2$), such that for low analyte concentrations (and biologically relevant-molecules) SE-SRS will have the best SBR, which is more relevant for us. Furthermore, if we would have considered a bulk liquid (IPA) as the analyte instead of the NTP monolayer, the difference between both SBRs would only have been an order of magnitude. Therefore, we can conclude that both on-chip SE-CARS and SE-SRS have a comparable signal generation (with an edge for SE-SRS) and SBR (with an edge for SE-CARS as long as $SBR_{SECARS} > 1$).

Since the performance of both techniques is comparable it would be tempting to choose SE-CARS over SE-SRS as it requires a less complex detection (compared to the lock-in detection scheme of SE-SRS). In SE-CARS, the newly generated anti-Stokes frequency can be separated from the Stokes and pump lasers by optical filters and afterwards be detected by a cheap single pixel detector. However:

- An important difference between both techniques is the quadratic dependence of SE-CARS on the concentration compared to the linear dependence of SE-SRS. For biologically relevant analytes, the concentration to be detected could easily lead to a drop of $\chi^{(3)}$ by three orders of magnitude or more compared to the monolayer of NTP considered here [79]. In that case $P_{SECARS,pulsed}/P_{SESRS} \approx 10^{-3}$ ($SBR_{SECARS}/SBR_{SESRS} \approx 0.07$) such that SE-SRS will have a much stronger Raman response than SE-CARS. The linear dependence of SE-SRS makes it more sensitive to diluted analytes and better equipped to detect low analyte concentrations [79, 139]. Since the detection of biologically relevant concentrations will be very important for an integrated Raman sensor in a lab-on-chip context, waveguide based SE-SRS provides a clear advantage over SE-CARS in this aspect.

- While SE-CARS has a straightforward detection scheme our calculations also show that the SE-CARS signal is inherently weak such that the use of high power pulsed lasers will be inevitable. The use of higher intensity lasers in SE-CARS will make it more likely that it will run into damage of the sample and analyte compared to the low-power CW beams required for the SE-SRS experiment as the induced temperatures in the plasmonic slot will be much higher (see Section 4.4.2.1). Where thermal damage to both the plasmonic substrate and analyte have been reported for picosecond pulses in plasmon enhanced coherent Raman experiments [59]. Furthermore, preliminary measurements with a pulsed laser on the plasmonic slot show that thermal damage is induced to the plasmonic slot at lower average power levels as those determined for CW lasers.

- At the moment, high power integrated mode-locked lasers (I-MLL) are not yet available in the visible which will hurdle the demonstration of SE-CARS on a full Raman-on chip sensor. Note that such I-MLL in the visible might become available in the (near-)future as I-MLL at 990 nm have been demonstrated [140, 141] and recent advances in modeling allow for a better control of the targeted device parameters [142]. In our SE-SRS estimate we assume CW lasers [67, 68], (balanced) detectors [143, 144] and a lock-in amplifier [145] which all have been demonstrated on integrated platform with decent performance. Furthermore, by integrating an array of lasers onto the chip, where each laser is dedicated to a specific Raman mode, we can probe

the Raman spectrum without the need for tuning of the laser wavelength.

Considering all these elements (and especially the greater sensitivity of SE-SRS over SE-CARS for biologically relevant analytes), it is decided to focus on SE-SRS in this work.

4.5 Overview table of theoretical comparisons

Raman platform	$P_{(a)S}/P_p$	SBR	**Type**	**Effective length**
SERS [16]	1×10^{-9}	18	plas. wg.	1.8 μm
SRS [19]	5×10^{-8}	0.01	strip wg.	8 mm
	1×10^{-6}			
SE-SRS	3.52×10^{-6}	0.76×10^{2}	plas. wg.	1.8 μm
SE-CARS	1.9×10^{-6}	5.78×10^{3}	plas. wg.	1.8 μm

Table 4.2: Overview table on the performance of the different on-chip Raman platforms calculated for a guided pump, Stokes power of 1 mW and 4 mW respectively and a NTP monolayer as the analyte. $P_{(a)S}/P_p$ *is the ratio of the generated (anti-)Stokes power over the injected pump power. For SRS, the Raman response for a liquid analyte (DMSO) top cladding is also shown (italic). SBR corresponds to the signal-to-background ratio.*

Raman platform	**Laser source**	**Detection**	**Spectral resolution**
SERS	$***$	$*$	$**$
	one CW laser	deeply-cooled CCD	Grating
SRS	$**$	$**$	$***$
	two CW lasers	BD+lock-in	Laser tuning
SE-SRS	$**$	$**$	$***$
	two CW lasers	BD+lock-in	Laser tuning
SE-CARS	$*$	$***$	$***$
	one CW+one Pulsed	Single pixel detector	Laser tuning

Table 4.3: Comparison between the required instrumentation for the different Raman techniques.CCD=Charge-coupled device, BD = balanced detector. $***$*=good ,* $**$*=moderate,* $*$*=bad.*

Table 4.2 gives an overview of the theoretical comparison made in this chapter and Table 4.3 gives a qualitative comparison on the instrumentation of the different

Raman platforms. We can derive from Table 4.2 that on-chip SE-SRS and SE-CARS provide a clearly improved Raman response compared to SERS and signal levels that are sufficiently strong such that deeply-cooled detection is no longer needed ($> 10^{-12}$W). In Table 4.3, we see that this increased sensitivity of the Raman sensor (and lack of cooled detection) comes at the cost of an additional tunable laser, a pulsed pump laser for SE-CARS and a more complex lock-in detection scheme for SE-SRS. On the other hand, the resolution of SE-SRS and SE-CARS is higher because they use tunable lasers (resolution of $\sim$ 1 pm) while SERS uses a grating (resolution of $\sim$ 0.1 nm). In terms of detection, we can see that SE-CARS is the most straightforward technique since it make uses of a room-temperature single pixel detector (see Section 4.4.3). Compared to SRS, we can see that SE-SRS and SE-CARS provide a significant improvement on the SBR by reducing the effective interaction length. Note that SRS provides a comparable Raman response to both plasmonic enhanced coherent Raman techniques when probing a liquid analyte such as DMSO. The decreased Raman response of SRS for a monolayer of NTP is to be expected as it suffers from the reduced field overlap with a monolayer compared to a full liquid cladding as dielectric waveguides provide a lower electric field enhancement near the waveguide surface than the plasmonic slot. As both SE-SRS and SRS use comparable instrumentation, it is clear that SE-SRS is preferable because of its improved performance. When comparing SE-CARS and SE-SRS, we can see that both have a similar performance (where the improved SBR for SE-CARS only holds for high concentration analytes, see Section 4.4.3). In terms of instrumentation, SE-CARS has a more straightforward detection, while SE-SRS does not require the use of pulsed lasers (with high power integrated mode-locked lasers not yet available in the near-visible). SE-SRS is preferred over SE-CARS in this work since it has a linear dependence on the analyte concentration (compared to quadratic for SE-CARS) and will therefore have a higher Raman response (and better SBR) for biologically-relevant analytes (as discussed in detail in Section 4.4.3).

Overall, we can conclude that SE-SRS is the best candidate for the application envisioned in this thesis (i.e. identifying a Raman technique that does not require deeply-cooled detection and does not suffer from a large background contribution to enable the integration of the full Raman spectroscopic system on-a-chip). Of course, other Raman techniques such as SERS on the plasmonic slot might be preferable if for example room-temperature detection is not the main requirement as they provide a much more straightforward excitation and detection.

5

Photon background mitigation using nanoplasmonic slot waveguides

To further extend the use of Raman spectroscopy systems, reducing its cost and size is essential. Therefore, the obvious answer lies in the nanophotonic integration of the different spectroscopic elements on a single photonic chip. All the elements have indeed been demonstrated individually: spectrometers [18, 65, 66], lasers [67, 68], spectral filters necessary to remove the strong excitation radiation [17] and Raman sensors that boost the Raman response [13, 14, 80]. As explained in Section 4.1, for the Raman sensor the nanoplasmonic slot is the ideal candidate. These plasmonic waveguides show a level of Raman enhancement similar to the nanophotonic dielectric waveguides and share the same non-resonant (broadband) enhancement but without suffering from a significant photon background from the waveguide dielectric. It is this photon background that ultimately reduces the signal-to-background ratio of any acquired Raman spectrum [64]. Proceeding to integrate this sensor with a circuit capable of analyzing the Raman scattered light demands that no further inelastic scattering or photoluminescence occurs in the circuit surrounding the sensor. Unfortunately, the Raman-like background from the silicon nitride waveguides becomes significant over propagation distances that can easily exceed millimeters for spectral filters [17] or AWG. The resulting background ultimately limits the concentration of analytes that can be probed because of the shot noise associated to it. Therefore, in practical SERS experiments, the

measurement is done in a back-reflection configuration (see Fig.5.1a) where a minimal access waveguide is used as the only interface between the facet of the chip (where an external excitation beam is focused) and the Raman plasmonic sensor itself. In [16] it is clearly shown that the signal-to-noise ratio was critically dependant on the length of that access waveguide so that it could not be combined with a "conventional" spectral notch filter such as Bragg gratings placed after the sensor [17]. Furthermore, integrating such a notch filter on the access waveguide will also prevent the excitation field (pump) to reach the plasmonic sensor. In order to separate the Raman scattered light from the excitation beam without inducing a spurious photon background, a multi-mode interferometer (MMI) [146] can be used together with backward Raman collection. In Fig. 5.2, we see that the MMI lets the excitation beam propagate from the input facet to the sensor and then lets the Raman back-scattered light propagate from the sensor to an output waveguide, while the excitation beam and associated background from the entrance waveguide keep on forward propagating in the plasmonic sensor and are finally absorbed. Of course, this approach induces an intrinsic 3dB loss for the Raman scattered light from the sensor. Despite this, the architecture is beneficial because the MMI is short (112 μm) (therefore inducing little Raman-like background) and it is also highly multimode so that most of that spurious background radiation generated in the MMI actually does not couple to the output port. In theory, using a circulator over an MMI would be even better since it would collect 100% of the backscattered Raman light. However in practice, integrated circulators are difficult to implement and they typically have a power penalty of $\sim$ 3dB associated with them [147].

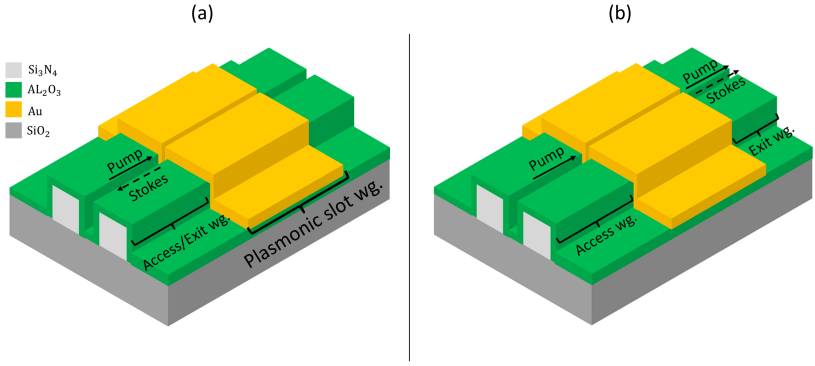


Figure 5.1: a) SERS measurement in the back-reflection configuration (i.e. access waveguide = exit waveguide). The excitation (pump) field is propagation in the opposite direction as the generated Raman signal which allows for the SERS signal to be collected without an additional background penalty of the excitation field in the exit waveguide. b) transmission-configuration (i.e. access waveguide ≠ exit waveguide). The excitation field propagates through the plasmonic slot together with the generated SERS signal where it will generate an additional background in the exit waveguide.

In the remainder of this Chapter, we will firstly discuss the MMI-design and its fabrication, followed by details about the spontaneous Raman setup and the measurement procedure. Then, we discuss the experimental results obtained and determine the origin of the extra background generation compared to a common SERS measurement on the plasmonic slot. Finally, we compare our MMI-configuration to alternatives architectures such as a simple transmission-configuration (see Fig. 5.1b) and a direction coupler-configuration and discuss possible improvements to the device .

The text/figures in this chapter have been adapted from my published paper [128].

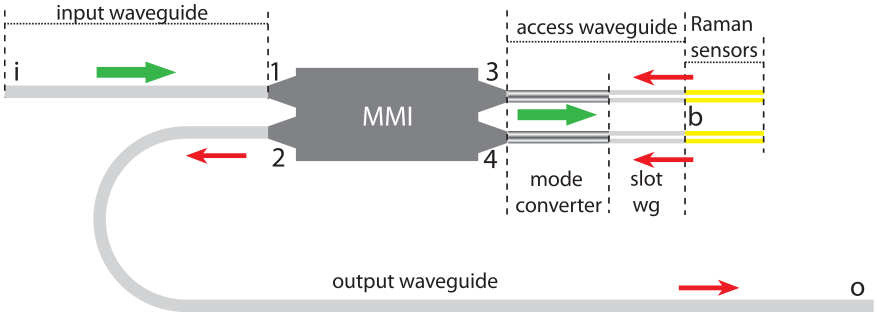


Figure 5.2: Schematic of the configuration relying on two nanoplasmonic Raman sensors and a MMI separating the forward-propagating excitation beam (green arrow) from the backward-propagating Raman scattered light (red arrows). Grey elements are made of Si_3N_4 while yellow elements are primarily made of gold. The MMI and the Raman sensing plasmonic slot waveguides are linked by an access waveguide made of a mode converter (grey gradient section), and a dielectric slot waveguide (grey and white). The access waveguide + the plasmonic Raman sensor correspond to the schematic shown in Fig. 5.1 a.

5.1 Methods

5.1.1 MMI design

Our architecture relies on a 2 × 2 multimode interferometer (MMI) [148]. The two left ports (labeled 1 and 2 in Fig. 5.2) are the input and output of our system. As an MMI is a linear and reciprocal element, having a single output port (labeled 3) necessarily translates into extra loss either from port 1 to 3 or from port 3 to 2 (at a given wavelength). However two output ports (labeled 3 and 4), having a splitting ratio of 50 : 50 maximizes the overall transmission from port 1 to ports 3 and 4, and then from those two ports to port 2. In that case, the input excitation power is optimally used if we place a Raman sensor at each of the two output ports of the MMI. Our MMI is designed via eigenmode expansion using Lumerical 's commercial solver and by taking the thickness (300 nm) and the refractive index of the Si_3N_4 layer (n = 1.89 at a wavelength of 830 nm) as fixed parameters. The designed in-plane dimensions of the balanced MMI are w × l = 8.9 × 112.5 μm^2. This results in a balanced MMI designed at a central wavelength of 830 nm (with a transmission of 49.8% at each port) and a 3 dB bandwidth of 220 nm. The MMI is complemented at each port by a 16.5 μm long taper allowing a good mode matching between the mode of the strip waveguide and the mode of the MMI (taper width of 2.45 µm at the output), thereby reducing spurious reflections. The MMI is preceded by a 300 μm-long strip waveguide (input waveguide) having a cross-section w × t = 550 × 300 nm^2. The MMI and plasmonic Raman sensors are connected together via a mode converter matching the rib waveguide mode to the slot waveguide mode and a short section of dielectric slot waveguide. In an ideal situation, the mode converter, dielectric slot waveguide and plasmonic slot are the only ones that contribute to the photon background as they are expected to scatter photons backward in the same way the nanoplasmonic sensors generates the Raman (signal) light. The mode converter is 20 μm long and the short section of dielectric slot waveguide varies from sample to sample from 23.5 to 116 μm. The width of the slot waveguide is 700 nm including a gap of 150 nm. In the rest of this work the mode converter and dielectric slot waveguide will be taken together and described as one single waveguide: the access waveguide. While the entire circuit (input/output waveguide, MMI and access waveguide) is cladded with 2 μm of silica on the top, the Raman sensor and a fraction of the dielectric slot waveguide are open to the air. To avoid a poor mode matching at the silica-air interface, the air is replaced by water during our measurements.

As our configuration collects backward-scattered Raman light, there is no optimal length for the plasmonic waveguides and the Raman signal simply saturates as the excitation beam decays due to propagation loss. Therefore, we choose to

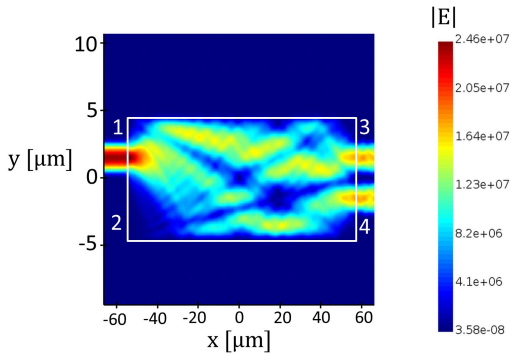


Figure 5.3: Simulated field profile of the balanced MMI with in-plane dimensions of width × length = 8.9 × 112.5 μm^2.

have a plasmonic slot waveguide length of 55 μm-long which far exceeds the decay length of the pump field (1.8 μm). The nanoplasmonic slot has a nominal gap width of 19 nm as reported previously [16]. The output waveguide has the same cross section as the input waveguide. This waveguide is taken purposely long (11 mm) to demonstrate the possibility offered by our architecture. In the future, that waveguide can include spectral functionalities e.g. a pump rejection filter [17].

5.1.2 Fabrication

The photonic circuits used for experiments are fabricated on 200 mm silicon wafers. 300 nm Si_3N_4 is deposited through a low temperature plasma enhanced chemical vapor deposition process [74] on a 3.3 μm SiO_2 layer (imec's BioPIX platform [149]). Patterning is performed by deep UV-lithography (193 nm) and fluorine based inductive coupled plasma reactive ion etching. A SiO_2 top-cladding of 2 µm is deposited over the photonic circuits except for an open-cladded region of w × l = 200 × 3000 µm^2 over the access waveguide in order to fabricate the plasmonic slots. The nanoplasmonic slots are fabricated by conformally covering the Si_3N_4 slot waveguides with 60 nm Al_2O_3 using atomic layer depostion (ALD), this deposition is performed by the CoCooN Research Group at Ghent University. The ALD deposition is performed by a home-build reactor with a base pressure of 10^{-6} mbar. Subsequent the ALD depostion is done by using a thermal ALD process with a substrate temperature of 120°C using trimethylaluminum and water pulses of 5s at 5×10^{-3} mbar [150]. This alumina layer serves the purpose of reducing the dielectric slot gap of 150 nm to 30 nm. Al_2O_3 is chosen as it only has very limited background generation and does not induce an additional background penalty [16]. After this deposition, on a short section of the narrowed slot waveguide, an open

window is defined by using AZ 5214 photo-resist and contact lithography. This window is 55 µm along the direction of the waveguide and 3.9 µm wide. The followed procedure is the following: The AZ 5241 resist is spin-coated over the chip with 3500 RPM and a spinning time of 40s, followed by a bake of 3 min at 100°C. After aligning the contact mask, the resist is exposed to DUV for 11s followed by a post-bake of 3 min at 120°C and DUV flood exposure (image reversal step) for 51s. After this exposure the resist development step is performed where the resist is exposed for 21s to AZ400 K:DI-water (1:3) solution followed by deep rinsing with DI water. Finally a 30s oxygen plasma step (PVA-TEPLA GIGAbatch 310 m, 6000 SCCM O2, 600 W, 750 mTorr) is performed to remove any remaining residues. Afterwards a 2 nm Ti adhesion layer is sputtered followed by the Au-sputtering of a 5.5 nm gold layer along the side walls of the slot. Finally our 55 µm long plasmonic slot is defined during metal lift-off.

Optimizing the Au-sputtering step turned out to be a challenging endeavour as the gold gap kept being closed (see Fig. 5.4a). Controlling this gold layer in the slot is difficult due to the random nature of the sputtering process and the sputtered layers not being conformal. A meticulous trial and error process was needed (i.e. varying the sample orientation and sputtering time) in order to achieve the required gold thickness of 5.5 nm in the plasmonic slot (See Fig. 5.4b). It was found that the best sputtering results are obtained when the sample is orientated at 90 degrees compared to the spin direction. We can see that sputtered gold does not completely reach the bottom of the slot but this only has a limited effect on the device performance [84].

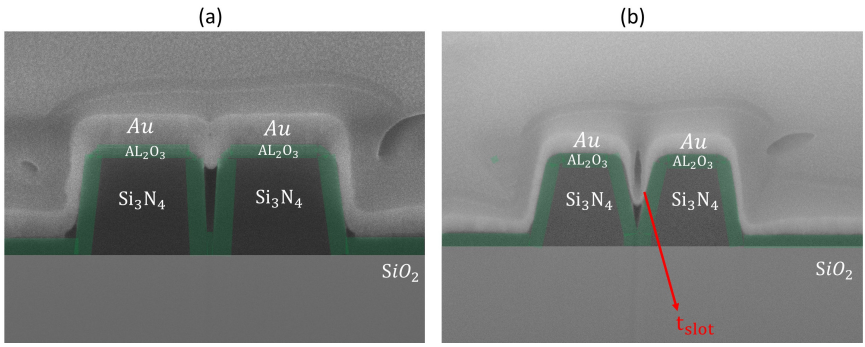


Figure 5.4: (a) 2D cross-section of the nanoplasmonic slot (using SEM) with a closed gold gap and the different material layers indicated. (b) Plasmonic slot with an open gold gap and the slot gold thickness (t_{slot}=5.5 nm) indicated.

Samples have been prepared that differ in the length of the dielectric slot waveguide between the MMI and the Raman sensor. This results in access waveguide lengths of 43.5 and 136 μm. Two structures having the same length are fabricated

on the same die for a total of four samples. The samples are functionalized with an analyte of 4-Nitrothiophenol (NTP) that binds selectively to the gold sensor. The NTP is bound to gold through a standard binding process, i.e. the chemisorption of thiols from an ethanolic solution to form the self-assembled monolayer of NTP [16, 151]. In order to achieve this, the chips with the plasmonic sensors are first cleaned using a combination of acetone and isopropyl alcohol followed by a 30s oxygen plasma clean. The chips are then put in a beaker of 1 mM of NTP-ethanol solution (solution made by mixing dry 4-NTP powder with ethanol) and left overnight. Afterwards, they are cleaned with a pure ethanol solution by softly rinsing this solution over the chip surface and facets to remove any NTP residues that are not bound to the gold. Finally, the chips are blow dried with a N_2 dry clean. In this way the NTP-monolayer is formed on the gold through a gold-sulfur bond. During measurements a droplet of water is applied on top of the plasmonic slot and NTP. This allows a better mode matching at the interface of the cladded and uncladded parts of the circuit.

5.1.3 Setup & Measurements

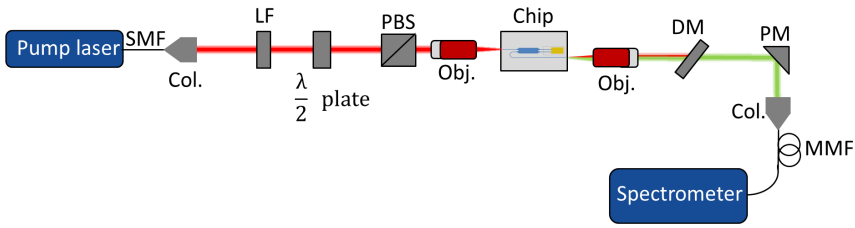


Figure 5.5: Schematic of the setup used to measure the SERS spectra. SMF: single-mode fiber, COL: collimator, LF: line filter, HWP: half-wave plate, PBS: polarizing beamsplitter, OBJ: objective , DM: dichroic mirror, PM: parabolic mirror, MMF: multimode fiber.

SERS spectra of the samples are acquired on the measurement setup depicted in Fig. 5.5 . The excitation CW laser light (785 nm) is provided by an XTRA II laser. A spectral filter is used in order to clean its spectrum. A polarizing beamsplitter ensures that only the TE-polarized light is transmitted towards the chip while the half-wave plate provides control over the amount of pump power. The light is coupled into and out of the chip by two identical Mitutoyo Plan Apo objectives (50×, NA=0.65). A dichroic mirror suppresses the residual excitation beam while transmitting the Raman scattered light. Finally, the light is captured in a collimator and guided in a multimode fiber towards the spectrometer (Shamrock 303i spectrometer with a Andor iDus 416 deep-cooled CCD camera).

Spectra are acquired for durations ranging from 60 to 180 seconds and excitation powers from 6 to 25 mW (measured before the objective). The excitation power is limited in order to prevent the photoreduction of NTP into a form of dimercaptoazobenzene (DMAB). A droplet of water is applied on top of the plasmonic slots when performing the acquisition. This allows a better mode matching at the interface of the cladded and uncladded parts of the circuit.

A typical spectrum acquired with our plasmonic sensor is presented in figure 5.6. It displays the characteristic peaks associated to NTP. The spectrum is presented with the dark counts subtracted. As the noise associated to the dark counts cannot be subtracted, we choose the acquisition parameters so that this noise is far lesser than other sources of noise. The spectrum displays the noticeable photon background we referred to in the introduction. To compare various results, we define the signal-to-background ratio (SBR) as the integrated number of counts (S) in the main peak (going from 1287 cm^{-1} to 1407 cm^{-1} and corresponding to 1341 cm^{-1} NTP-mode, shaded in green) by the integrated number of counts due to the photon background (BG) integrated over the same spectral span (shaded in red).

Defining a SBR rather than the signal-to-noise ratio is more relevant when parameters such as excitation power, incoupling loss, or acquisition time are varied. The SBR is independent of those factors. The SBR depends however on how well the signal and the background spectrum can be separated. That separation is perfect if the spectrum of the spurious photon background is completely known and if there would be no shot noise associated to both the background and the Raman scattered light. In the present case, we assume the background follows a purely linear trend over the integration span corresponding to the NTP peak at 1341 cm^{-1}. That approximation and the presence of the shot noise induces an uncertainty (relative standard deviation) of 12% on the SBR values we provide.

5.2 Results

5.2.1 Performance

Fig. 5.6 clearly shows the presence of a remaining photon background. We expect this background to be due to the access waveguide between the MMI and the plasmonic waveguide itself. We therefore measure the effect of the access waveguide length on the SBR. The results presented in Fig. 5.7 show that indeed the longest access (136 μm) waveguide is associated to a reduced SBR because the background increases. Furthermore, the background contribution due to the plasmonic slot itself is only limited since increasing the access waveguide length

has a significant effect on the SBR (which would not be the case if the background due to the plasmonic slot was dominant). We believe that the absence of a strong plasmonic slot contribution originates from the presence of a water top-cladding which pushes the plasmonic mode into the cladding. This results in less overlap with the dielectric core and a reduced background contribution (compared to an air-cladded measurement on the plasmonic slot [45]). In the remainder of this chapter, we will simply describe the background due to the access waveguide and plasmonic slot as the access waveguide contribution.

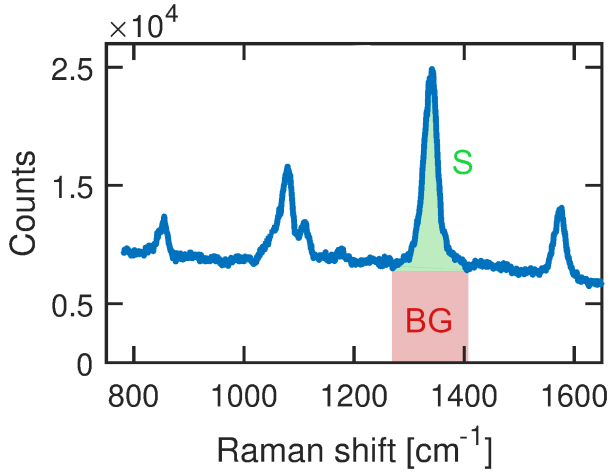


Figure 5.6: Measured Raman spectrum of NTP using our photonic-plasmonic sensor (fig. 5.2) with the instrumentation noise subtracted. The spectrum shows the characteristic NTP features (855 cm^{-1}, 1111 cm^{-1}, 1341 cm^{-1}, 1575 cm^{-1}) [152]. The green and red regions indicate the integration range (1287 cm^{-1} : 1407 cm^{-1}) used to calculate the signal (S) to background (BG) ratio.

In an ideal MMI-plasmonic slot hybrid, this contribution due to the access waveguide would be the only major background sources. However unwanted back-reflections at the interface between the dielectric and plasmonic slot, the interface between the MMI-taper and the mode converter, the interface between the mode converter and the slot waveguide, in the MMI from the input port (port 1) to the output port (port 2) or the interface between the silica cladded part of the dielectric slot waveguide and the water cladded part of that same waveguide can reflect a small portion of the excitation beam into the output waveguide. Through numerical simulations, the reflection at the plasmonic sensor interface is expected to be at least 10 dB larger than the other reflections. See Appendix A for a detailed discussion on the simulations.

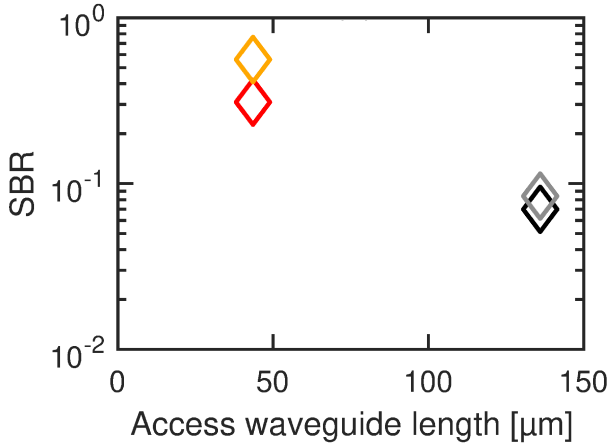


Figure 5.7: The signal-to-background ratio for our four photonic-plasmonic sensors.

5.2.2 Origin of extra background

While the length of the access waveguide could be reduced further, it does not necessarily lead to a reduced background if the reflectivity of the nanoplasmonic sensor becomes significant. Indeed, the reflected excitation beam will eventually reach the long output waveguide where it will generate a photon background. To quantify the importance of that undesired reflection on the SBR, we measured it for two output waveguide lengths on the same chip: 11 mm and 5 mm. Figure 5.8 presents the result of that measurement. We display the ratio (R) of the SBR when the output waveguide is the longest (11 mm) to the SBR when the output waveguide is only 5 mm. This ratio is therefore immune to fabrication variations occurring for chips sputtered with different conditions. The fact that this ratio is smaller than unity indicates that some spurious photon generation occurs in the output waveguide. The fact that it goes asymptotically to 1 when increasing the access waveguide to 100's of micrometers indicates that the background contribution due to the access waveguide length is strongly dominant in that case. At the opposite side, if the background contribution due to the output waveguide was strongly dominant, this ratio would be 46% (5 mm/11 mm).

For the structures with the short access waveguide the ratio R is 0.83, showing that the background due to the access waveguide and the output waveguide are of the same order. Other background sources due to the MMI and the input waveguide do not seem to play a major role under the device dimensions probed. If they did, R would be approaching unity for all the structures as changing the output waveguide length would not affect it.

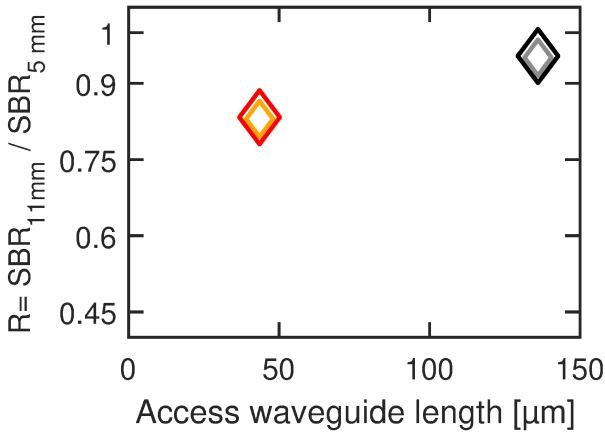


Figure 5.8: Decrease in SBR when increasing the output waveguide length from 5 mm to 11 mm for the four samples.

The fact that the background contribution due to the 11 mm output waveguide is nearly equal to that of the 43.5μm access waveguide (and plasmonic slot) is already an encouraging results. It shows we can integrate elements such as a notch filter [17] to remove the excitation beam by paying a low price in term of the SBR: a reduction by a factor of 2 (hence a reduction of the SNR by $\sqrt{2}$). As the length of the access waveguide can easily be reduced to the length of the mode converter (20 μm), we can conclude that the actual limitation of our architecture lies in the reflections arising at the interface between the dielectric and plasmonic slot waveguides. Before discussing how to further improve this, we are comparing this architecture with the simplest architecture possible.

5.2.3 Comparison with alternative architectures

In order to integrate the Raman sensor with other elements, the alternative to our architecture consists of placing the elements behind the sensor and collecting the forward-scattered Raman light. We refer to that configuration as the ”transmission configuration”(see Fig. 5.1).

To make a comparison, we assume a circuit for the transmission-configuration with a plasmonic slot waveguide of optimal length (1.8 μm [64]). The amount of Raman scattered power can then be taken equal in both configurations. In the case of the MMI-based circuit, we consider the photon-background contribution due to the access waveguide with a length of 43.5 μm (our best demonstrated sample), an equivalent background contribution due to the plasmonic slot and the contribution

due the output waveguide. As we know the amount of photon background generated per unit of length and per unit of optical power of the excitation beam, we can predict the amount of this photon background as a function of the output waveguide length (see Section 5.2.3.1 for the detailed calculations) . This requires knowledge of the fraction of the excitation laser beam power reflected to the output port. We experimentally estimated this reflection to be between -26.9 dB and -32 dB for our various samples (by removing the dichroic mirror of Fig. 5.5 and connecting the multimode fiber to a power meter instead of the spectrometer). From simulation, those numbers are between - 25 dB and - 30 dB where the uncertainty reflects the imprecise knowledge of the gold thickness. Because the excitation power is reduced, the photon background contribution of the output waveguide is 173 times weaker per unit of length than the contribution due to the access waveguide. For simplicity, we assumed the MMI to be perfectly balanced both at the excitation and Raman scattered wavelengths.

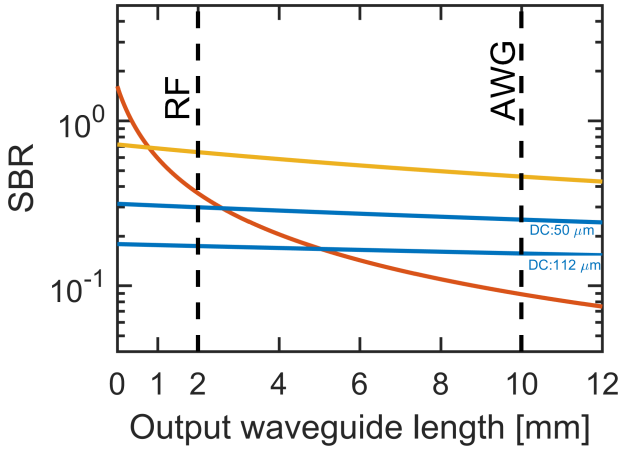


Figure 5.9: Evolution of the signal-to-background ratio with the length of the output waveguide for the case of the MMI-based configuration (yellow curve) compared to the simpler "transmission configuration" (red curve) and a directional coupler-based configuration (blue curves) with a DC length of 50 μm & 112 μm respectively. The dashed lines illustrate the typical length of elements that could be integrated to the plasmonic Raman sensor: rejection filter (RF) [17] and arrayed waveguide grating (AWG) [18].

In the case of the transmission configuration, we can proceed to the same reasoning but we have to make an assumption on the transmission of the pump and of the Raman scattered light through the plasmonic slot waveguide. We take this transmission to be -12.22 dB (coupling loss of -3.95 dB/facet and absorption loss of -2.4 dB/μm [16]). The result of the comparison is plotted in Fig. 5.9 via the expected SBR corresponding to both configurations. It can be seen that the SBR of the MMI-based configurations outperforms the transmission configuration for

an output waveguide length as short as 782 μm despite the current (not optimized) geometry of the access waveguide length. This MMI configuration thus proves to be a good option for further integration with spectral functionalities. It also eases the fabrication because collecting the back-scattered Raman light implies that the length of the nanoplasmonic waveguide does not need to be optimized while it requires to be precisely controlled for the forward-scattering case [64].

We can conclude that our MMI provides a significant advantage over the before mentioned transmission measurement and comes at nearly no cost. One could however argue that the same would apply to a balanced directional coupler (DC) with a 50/50 splitting ratio. This is not the case because a directional coupler comes with a significant photon background noise penalty. As a DC is simply made of two single-mode waveguides and as we know the quantity of background photons generated in such a waveguide per unit of length [64, 86], the effect on the SBR can be deduced for a DC with the same length as the MMI (112 μm) and a more compact DC with a length in the order of 50 μm [153]. Where we only consider the reflection at the plasmonic slot facet and no reflection in the DC itself (see Section 5.2.3.1). The result of that calculation can be seen in Fig. 5.9. It shows indeed that the SBR is between 1.75 and 4 times better for the MMI-case as compared to the DC-case, depending on output waveguide length and DC-length.

The fact that our MMI provides a better SBR is because the photon background originating from it is small enough to be negligible compared to background generated in the access and output waveguide. This can be understood when considering the multimode behaviour of the MMI. If only the modes with TE polarization are considered then the MMI supports 24 different guided modes. As a result the generated background photons, which are incoherent relative to each other, couple to all those modes and will upon propagation to the output of the MMI not be confined to the output waveguides. This is further confirmed by the fact that only 5 out of those 24 modes are responsible for supporting 90% of the transmission of in-coupled to out-coupled signal. Therefore, most of the background generated in the MMI will be radiated out rather than being collected at the output waveguide. The effect will be similar for the TM modes. This clearly demonstrates the benefit of using an MMI rather than a directional coupler.

5.2.3.1 Estimating the different signal-to-background ratios

Before starting the description of the different background contributions, let us firstly recall the general equations for the amount of background power P_{BG} generated in the backward- and forward-propagating direction along a dielectric waveg-

uide with length L (see Section 3.2.3.1):

$$P_{BG,back.} = \frac{1}{2} P_{In} \eta \overbrace{\rho_{SiN} \sigma_{SiN}}^{\beta_{SiN}} \left(\overbrace{\frac{1 - e^{-(\alpha_P + \alpha_S)L}}{(\alpha_P + \alpha_S)}}^{LF} \right) \tag{5.1}$$

$$P_{BG,forw.} = \frac{1}{2} P_{In} \eta \overbrace{\rho_{SiN} \sigma_{SiN}}^{\beta_{SiN}} \left(\overbrace{e^{-\alpha_{PL}} \frac{e^{(\alpha_P - \alpha_S)L} - 1}{(\alpha_P - \alpha_S)}}^{LF} \right) \tag{5.2}$$

Where P_{In} is the guided pump power in the input waveguide, η the Raman conversion efficiency, ρ_{SiN} and σ_{SiN} the molecular density and the scattering cross section of the Si_3N_4 waveguide core and LF the length factor with α_P, α_S the waveguide loss for the pump and stokes beam and L the waveguide length. From figure 5.10

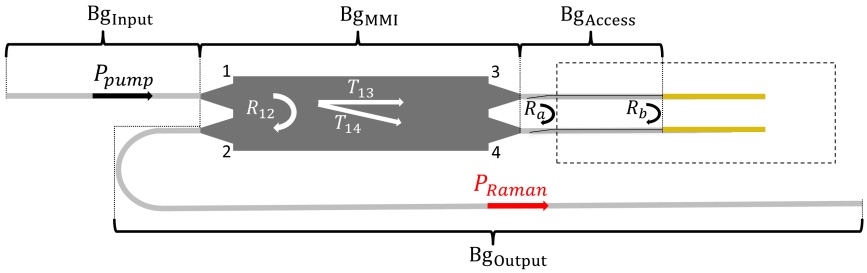


Figure 5.10: Different background (BG) sources in the MMI-plasmonic slot configuration. R_{12} *represent the reflection in the MMI from the input port (port 1) to the output port (port 2) when light is being transmitted towards the access waveguides (*T_{13} *&* T_{14}*).* R_a *and* R_b *represent the reflections at opening of the oxide cladding and the plasmonic slot facet respectively.* R_a, R_b *are assumed to be the same for both access waveguide arms.*

the different background contributions can be illustrated. Instead of expressing the different contributions as integrated detector counts (BG) it is better to express them in function of the guided background power on-chip (P):

$$P = \frac{\zeta \, h\nu_s}{(T_{Chip} T_{Set}) \, t_{int}} BG \tag{5.3}$$

Where ζ is conversion rate of the spectrometer [25], $h\nu_s$ the photon energy, T_{Chip} & T_{Set} the transmission loss when coupling from the chip to free-space and transmission loss in the setup after the chip respectively, t_{int} is the integration time used during experiment.

The total background power being present at the end of the output waveguide can now be expressed as the sum of the background power generated in the input waveguide ($P_{I,F}$), the MMI ($P_{M,B}$ & $P_{M,F}$), the access waveguide ($P_{A,F}$ & $P_{A,B}$)

and the output waveguide ($P_{O,F}$). Where F, B indicates if the background power is being generated by the forward or backward Raman scattering process respectively. Furthermore, we include the effect of the background generated in the plasmonic slot in P_A. Hence $P_{BG_{tot}} = P_{I,F} + P_{M,F} + P_{M,B} + P_{A,F} + P_{A,B} + P_{O,F}$. Now the background power due to the individual background sources can be written as:

$$P_{I,F} = \frac{1}{2}\,\eta_{Strip}\,\beta_{SiN}\,P_{In}\,L_{In} \times \overbrace{\Big(R_{12}(s) + \big(R_a(s) + R_b(s)\big)\,\big(T_{13}(s)\,T_{32}(s) + T_{14}(s)\,T_{42}(s)\big)\Big)}^{T_{I,F}} \tag{5.4}$$

$$P_{M,B} = \frac{1}{2}\,\eta_{MMI}\,\beta_{SiN}\,P_{In}\,L_{MMI} \times \overbrace{\Big(T_2(s)\Big)}^{T_{M,B}} \tag{5.5}$$

$$P_{M,F} = \frac{1}{2}\,\eta_{MMI}\,\beta_{SiN}\,P_{In}\,L_{MMI} \times \overbrace{\Big(\big(R_a(s) + R_b(s)\big)\,\big(T_3(s)T_{32}(s) + T_4(s)T_{42}(s)\big)\Big)}^{T_{M,F}} \tag{5.6}$$

$$P_{A,B} = \frac{1}{2}\,\eta_{Slot}\,\beta_{SiN}\,P_{In}\,L_A \times \overbrace{\Big(T_{13}(p)\,T_{32}(s) + T_{14}(p)\,T_{42}(s)\Big)}^{T_{A,B}} \tag{5.7}$$

$$P_{A,F} = \frac{1}{2}\,\eta_{Slot}\,\beta_{SiN}\,P_{In} \times \overbrace{\Big(\big(L_a\,R_a(p) + L_A\,R_b(p)\big)\,\big(T_{13}(p)\,T_{32}(s) + T_{14}(p)\,T_{42}(s)\big)\Big)}^{T_{A,F}} \tag{5.8}$$

$$P_{O,F} = \frac{1}{2}\,\eta_{Strip}\,\beta_{SiN}\,P_{In}\,L_O \times \overbrace{\Big(R_{12}(p) + \big(R_a(p) + R_b(p)\big)\,\big(T_{13}(p)\,T_{32}(p) + T_{14}(p)T_{42}(p)\big)\Big)}^{T_{O,F}} \tag{5.9}$$

$$P_{Sg,B} = \frac{1}{2}\,\eta_{Pl.}\,\beta_{NTP}\,P_{In}\,LF_{Pl.} \times \overbrace{\Big(T_{13}(p)\,T_b(p)\,T_b(s)\,T_{32}(s) + T_{14}(p)\,T_b(p)\,T_b(s)\,T_{42}(s)\Big)}^{T_{Sg}} \tag{5.10}$$

Where s, p indicate whether the reflection or transmission is occurring at the stokes or the pump wavelength. The assumption LF = L is made in order to simplify the equations, note that this holds for waveguide lengths up to 1 cm (with

$\alpha_P = \alpha_S = 1.5$ dB/cm for the cladded strip waveguide). From there, the LF and additional loss factors need to be introduced in the equations [64]. The length L_a refers to the small piece of access waveguide between the MMI and the opening of the oxide cladding while L_A refers to the complete access waveguide length (+ an extra length contribution whose background is equivalent to the background generated in the plasmonic slot). L_O corresponds to the output waveguide length. The reflections at the mode converter and the tapers are not considered in the equations, as from simulations (see Appendix A) they are expected to be much lower than the other reflection sources.

In order to better understand the terminology used in the equations above let us take the back-scattered background generated in the access waveguide ($P_{A,B}$) and the background power generated by the forward-scattering process in the output waveguide ($P_{O,F}$) as examples. Let us start with $P_{A,B}$, before the guided pump power (P_{In}) can reach the access waveguides it firstly needs to travel trough the MMI and hence the amount of pump power reaching the access waveguides will be $P_{in}(T_{13}(p) + T_{14}(p))$. The background generated in these access waveguides over an waveguide length L_A will therefore be proportional to $L_A\, P_{in}(T_{13}(p) + T_{14}(p))$. This background power is then transmitted towards the output waveguide before it contributes to the total background detected in our experiments. The amount of back-scattered background power due to the access waveguide is thus proportional to $L_A\, P_{in}(T_{13}(p)T_{32}(s) + T_{14}(p)T_{42}(s))$. In case of $P_{O,F}$, in order for the guided pump light to reach the output waveguide a small portion of it can be reflected in the MMI towards port 2 (R_{12}), at the opening of the oxide cladding (R_a) and at the interface of the plasmonic structure (R_b) after which it will again be transmitted through the MMI $R_{12}(p) + \left(R_a(p) + R_b(p)\right)\left(T_{13}(p)\, T_{32}(p) + T_{14}(p)T_{42}(p)\right)$. From there on these pump reflections can generate a background over the output waveguide length L_O.

Next, as mentioned in the Section 5.1.3, the SBR can be characterized as the ratio of the integrated signal counts (S) over the background counts (BG) (see figure 5.6). Using eq 5.3 an alternative way of characterizing the SBR is by using P_{Sg} and $P_{BG_{tot}}$:

$$\begin{aligned} SBR = \frac{S}{BG} &= \frac{\left(\frac{\zeta h\nu_s}{(T_{Chip}T_{Set})\, t_{int}}\right)}{\left(\frac{\zeta h\nu_s}{(T_{Chip}T_{Set})\, t_{int}}\right)} \times \frac{S}{BG} \\ &= \frac{P_{Sg}}{P_{BG,tot}} = \frac{\frac{P_{Sg}}{P_{In}}}{\frac{P_{BG,tot}}{P_{In}}} \end{aligned} \tag{5.11}$$

Besides P_{Sg} and $P_{BG_{tot}}$ being independent of the integration time, the ratio of P_{Sg} and $P_{BG_{tot}}$ also does not depend on the guided pump power used during experiments

as the amount of P_{in} is the same for both the signal and the background power. In Section 5.2.2 the ratio R was defined as:

$$R = \frac{SBR_{1.1cm}}{SBR_{0.5cm}} = \frac{\frac{P_{S,1.1cm}}{P_{BG,1,1cm}}}{\frac{P_{S,0.5cm}}{P_{BG,0.5cm}}} = \frac{\frac{P_{Bg,0.5cm}}{P_{in,0.5cm}}}{\frac{P_{Bg,1.1cm}}{P_{in,1.1cm}}} \tag{5.12}$$

Where we assume that $\frac{P_{Sg,0.5cm}}{P_{In,0.5cm}} = \frac{P_{Sg,1.1cm}}{P_{In,1.1cm}}$ since the variation of output waveguide from 1.1 cm to 0.5 cm does not have an effect on the NTP signal being generated by the plasmonic structure ($\eta_{Pl.}$ is unchanged). So instead of making a direct comparison of the background counts (BG) obtained for the different output waveguide lengths a better alternative is to use the normalized background variation R since it is independent of the integration time, plasmonic structure and the guided power in the experiment.

Next, the SBR of the MMI-slot configuration is compared to the transmission configuration and a directional coupler configuration. An identical plasmonic structure as the one experimentally demonstrated in this work is assumed in each collection scheme. This means that η_{pl} is the same in each configuration. Using eq.5.11 it is now possible to write the SBR_{MMI} as:

$$SBR_{MMI} = \frac{P_{Sg}}{P_{BG_{tot}}} = \frac{P_{Sg}}{P_{I,F} + P_{M,F} + P_{M,B} + P_{A,F} + P_{A,B} + P_{O,F}} \tag{5.13}$$

In order to come to a more workable expression of SBR_{MMI} some background contributions will be left out. Since the access waveguide and the MMI have a direct path towards the output waveguide (without the need of reflections), $P_{A,F}$, $P_{M,F}$ will be much smaller than $P_{A,B}$, $P_{M,B}$ respectively such that:

$$SBR_{MMI} = \frac{P_{Sg}}{P_{I,F} + P_{M,B} + P_{A,B} + P_{O,F}} \tag{5.14}$$

From Section 5.2.2 we know that both the background generated in the MMI and the input waveguide do not play a major role in the background. Therefore, those background contributions will not be considered in the theoretical estimate. The taper lengths of the MMI will be added to the access waveguide length ($L_{A_1} = 2 \times L_A + L_{taper}$) since the background generation in those tapers will be much higher than the MMI (as they guide a single mode). We will use $2 \times L_A = 2 \times 43.5$ µm i.e. doubling the access waveguide length, as we do not know the exact background contribution due to the plasmonic slot (only that it is smaller than the background contribution of the access waveguide). Therefore, we will be conservative and assume that it has the same background contribution as the

43.5 μm access waveguide. Including the effect of the plasmonic slot on the background is thus equivalent to doubling the access waveguide length.

$$\mathrm{SBR_{MMI}} = \frac{\eta_{\mathrm{Pl.}}\,\beta_{\mathrm{NTP}}\,\mathrm{LF_{Pl,B}}\,(\mathrm{T_{Sg}})}{\eta_{\mathrm{Strip}}\,\beta_{\mathrm{SiN}}\,(\frac{1}{3}\,\mathrm{L_{A_1}}\,\mathrm{T_{A,B}}\,+\,\mathrm{L_O T_{O,F}})} \tag{5.15}$$

Now for the reflections ($\mathrm{R_{12}}, \mathrm{R_a}, \mathrm{R_b}$), in Appendix A it is shown that the reflection at the facet of the plasmonic structure $\mathrm{R_b}$ is over 10 dB stronger than $\mathrm{R_{12}}$ and $\mathrm{R_a}$. In order to calculate $\mathrm{SBR_{MMI}}$, only $\mathrm{R_b}$ will be considered and $\mathrm{R_b(s)}$ will be assumed to be equal to $\mathrm{R_b(p)}$. Furthermore, the MMI transmissions (T) are all set to 0.5 (this will induce a tiny error $\sim 20\%$, as the transmission at the Stokes wavelengths is less efficient, which is sufficiently low for the type of calculation done here). So the $\mathrm{SBR_{MMI}}$ becomes:

$$\mathrm{SBR_{MMI}} = \frac{\eta_{\mathrm{Pl.}}\,\beta_{\mathrm{NTP}}\,\mathrm{LF_{Pl.,B}}\,(\mathrm{T_b})^2}{\eta_{\mathrm{Strip}}\,\beta_{\mathrm{SiN}}\,(\mathrm{R_b}\,(\mathrm{L_O})\,+\,\frac{1}{3}\,\mathrm{L_{A_1}})} \tag{5.16}$$

Where $\mathrm{T_b}$ is the transmission when transitioning from the access waveguide into the plasmonic slot and vice-versa.

For the transmission-configuration we have a background that is generated in the access waveguide, in the plasmonic slot and in the long output waveguide behind the plasmonic slot. For the access waveguide, we take a length of $\mathrm{L_A} = 43.5$ μm. Any background generated in the access waveguide will need to be transmitted through the plasmonic slot ($\mathrm{T_b^2 T_{pl}}$) before it can be collected in the output waveguide. Where $\mathrm{T_{pl}}$ is the transmission when passing through the plasmonic slot. In the same way the pump light that generates a background in the output waveguide also first needs to be transmitted through the plasmonic slot. The background contribution due to the plasmonic slot is represented as $\frac{1}{1.35}\,\mathrm{L_A}$, where the factor 1.35 includes the effect of the difference in the length factor between the forward and backward Raman generation [64]. Now, the SBR can be written as:

$$\mathrm{SBR_{Tran.}} = \frac{\eta_{\mathrm{Pl.}}\,\beta_{\mathrm{NTP}}\,\mathrm{LF_{Pl.,F}}\,(\mathrm{T_b})^2}{\eta_{\mathrm{Strip}}\,\beta_{\mathrm{SiN}}\,\frac{1}{3}(\mathrm{T_b^2 T_{pl}}\,(\mathrm{L_A}+\mathrm{L_O})+\frac{1}{1.35}\mathrm{L_A})} \tag{5.17}$$

Where we again use $\eta_{\mathrm{Slot}} = \frac{1}{3}\eta_{\mathrm{Strip}}$.

Finally, we can derive the SBR for the directional coupler-configuration. Since a directional coupler simply consists of two single-mode (strip) waveguides its effect on the background is similar as that of the access waveguide. Therefore, an

extra contribution (L_{DC}) due to the directional coupler is added such that the SBR becomes:

$$SBR_{DC} = \frac{\eta_{Pl.}\,\beta_{NTP}\,LF_{Pl.,B}\,(T_b)^2}{\eta_{Strip}\,\beta_{SiN}\,(R_b\,(L_O) + \frac{2}{3}\,L_A + L_{DC})} \tag{5.18}$$

Note that we use $\frac{2}{3}\,L_A$ instead of $\frac{1}{3}\,L_{A_1} = \frac{1}{3}\,(2 \times L_A + L_{taper})$ (as in Eq.5.16) since the DC does not require a taper section.

Now we can fill in the SBR formulas of the different configurations using the parameters as described in Table 5.1. The signal term $S = \eta_{Pl.}\,\beta_{NTP}\,LF_{Pl.,B}\,(T_b)^2$ of the different formulas is calculated by taking Eq. 5.16 and multiplying the denominator with $SBR_{MMI} = 0.56$ (for $L_{A_1} = 104\,\mu$m and $L_O = 5$ mm and using $LF_{Pl,F} = LF_{Pl,B}/1.35$ [64] for $SBR_{Tran.}$).

Parameter	Value	Unit
β_{SiN} [105]	3.3×10^{-8}	$[sr^{-1}cm^{-1}]$
η_{Strip} [64]	0.507	
L_{A_1}	0.0104	[cm]
L_A	0.00435	[cm]
L_{DC}	0.0112 0.0050	[cm] [cm]
R_b	-27	[dB]
T_b	-3.95	[dB]
T_{pl}	-4.32	[dB]

Table 5.1: Parameters used together with Eq. 5.16-5.18 to make the plots in Fig. 5.9.

Additionally, we can also make a comparison between the background generated in the output waveguide and the access waveguide (including the plasmonic slot contribution) since they are the most important background sources in our device. If we now use Eq. 5.7 and 5.9, we can describe the ratio of the output background to the access waveguide background as:

$$\frac{P_{O,F}}{P_{A,B}} = \frac{\eta_{Strip}\,\beta_{SiN}\,P_{In}\,L_O\,T_{O,F}}{\eta_{Slot}\,\beta_{SiN}\,P_{In}\,L_{A_1}\,T_{A,B}} \tag{5.19}$$

If we now again assume that the MMI transmissions (T) are all set to 0.5 and the reflection due to the plasmonic slot waveguide (R_b) is dominant and wavelength independent (i.e. $R_b(\lambda_s) = R_b(\lambda_p)$) then this ratio becomes:

$$\frac{P_{O,F}}{P_{A,B}} = \frac{\eta_{Strip} L_O R_b}{\eta_{Slot} L_{A_1}} \tag{5.20}$$

We can now also include the effect of a pump rejection filter on the output waveguide ($R_{filter,out}$) (ignoring the background contribution of the filter itself which is acceptable for $R_b \geq -27$dB, see Section 5.3.1) such that Eq. 5.20 becomes:

$$\frac{P_{O,F}}{P_{A,B}} = \frac{\eta_{Strip} L_O (R_b + R_{filter,out})}{\eta_{Slot} L_{A_1}} \tag{5.21}$$

Note that for long output waveguides ($>$ 1 cm) the propagation loss of the output waveguide also needs to be accounted for such that:

$$\frac{P_{O,F}}{P_{A,B}} = \frac{\eta_{Strip} LF_O (R_b + R_{filter,out})}{\eta_{Slot} L_{A_1} \exp(-\alpha_s L_O)} \tag{5.22}$$

where we now assume that $\alpha_s = \alpha_p$, such that (see Eq. 3.33):

$$\frac{P_{O,F}}{P_{A,B}} = \frac{\eta_{Strip} L_O \exp(-\alpha_p L_O) (R_b + R_{filter,out})}{\eta_{Slot} L_{A_1} \exp(-\alpha_p L_O)} \tag{5.23}$$

and we can see that we come to the same expression as Eq. 5.21.

Let us now apply the condition that the background due to the output waveguide needs to be an order of magnitude smaller than the access waveguide background (i.e. $P_{O,F}/P_{A,B} = BG_{out}/BG_{acc} = 0.1$). We can now express the maximum output waveguide length for this condition to hold as:

$$L_{out} = L_O = 0.1 \times \frac{L_{A_1}}{3 (R_b + R_{filter,out})} \tag{5.24}$$

where we used $\eta_{Slot} = \eta_{Strip}/3$.

In the same way we can now compare the background of the access waveguide to the background generated in a output fiber behind the output waveguide (with a coupling loss $\zeta_p = \zeta_s = \zeta$ between the waveguide and the fiber). The background ratio at the end of the fiber then corresponds to:

$$\frac{P_{fiber,out}}{P_{Access,fiber}} = \frac{G_{fiber} (P_{In} T_{O,F} \zeta_p) LF_{fiber,out}}{(G_{SiN,slot} P_{In} L_{A_1} T_{A,B}) \zeta_s \exp(-\alpha_s L_{fiber,out})} \tag{5.25}$$

We can now again assume that the reflection at the plasmonic slot interface is dominant (R_b), that the MMI transmissions (T) are all set to 0.5. Furthermore, we

assume that the propagation loss in the fiber (α_s, α_p), reflection at the plasmonic slot interface and the MMI transmission from port-to-port are the same at the Stokes and pump wavelength. Next, we also include the effect of a pump rejection filter $R_{filter,out}$ on the output waveguide before the light couples out to the optical fiber such that:

$$\frac{P_{fiber,out}}{P_{Access,fiber}} = \frac{G_{fiber} P_{In} (R_b + R_{filter,out}) \zeta_p L_{fiber,out} \exp(-\alpha_p L_{fiber,out})}{G_{SiN,slot} P_{In} L_{A_1} \zeta_s \exp(-\alpha_s L_{fiber,out})} \quad (5.26)$$

which can be simplified to:

$$\frac{P_{fiber,out}}{P_{Access,fiber}} = \frac{G_{fiber} (R_b + R_{filter,out}) L_{fiber,out}}{G_{SiN,slot} L_{A_1}} \quad (5.27)$$

Let us now again apply the condition that the background due to the output fiber needs to be an order of magnitude smaller than the access waveguide background (i.e. $P_{fiber,out}/P_{Access,fiber} = BG_{fiber,out}/BG_{acc} = 0.1$). We can then express the maximum output fiber length for this condition to hold as:

$$L_{fiber,out} = 0.1 \times \frac{G_{SiN,slot} L_{A_1}}{G_{fiber} (R_b + R_{filter,out})} \quad (5.28)$$

Now we can also derive an equivalent expression for an input fiber proceeding the chip (evaluated at the end of the output waveguide):

$$\frac{P_{fiber,in}}{P_{A,B}} = \frac{(G_{fiber} \frac{P_{In}}{\zeta_p \exp(-\alpha_p L_{fiber,in})} LF_{fiber,in}) \zeta_s T_{I,F}}{G_{SiN,slot} P_{In} L_{A_1} T_{A,B}} \quad (5.29)$$

which again can be reduced to (assuming a Stokes rejection filter on the input waveguide):

$$\frac{P_{fiber,in}}{P_{A,B}} = \frac{G_{fiber} (R_b + R_{filter,in}) L_{fiber,in}}{G_{SiN,slot} L_{A_1}} \quad (5.30)$$

For which the input fiber length can again be expressed as:

$$L_{fiber,in} = 0.1 \times \frac{G_{SiN,slot} L_{A_1}}{G_{fiber} (R_b + R_{filter,in})} \quad (5.31)$$

5.3 Improvements & Conclusions

5.3.1 Improvement strategies

The architecture presented here can be improved in several ways. One route for improvement consists in further reducing the length of the access waveguide. It is nearly trivial to reduce its length to the length of the mode converter (20 μm). Reducing the background contribution due to the plasmonic slot is not that straightforward since it also affects the signal generation. Furthermore, reducing the length of the mode converter is also not as trivial because there is a trade-off between a shorter length and proper mode matching. Since we know that for access waveguide lengths in the order of 43.5 μm the contribution due to the output waveguide is as strong as that due the access waveguide and plasmonic slot, further reducing the reflections of the excitation beam is thus the actual limitation. In the current study, we have no possibility to experimentally distinguish (1) the reflection due to the interface between the dielectric and plasmonic slot waveguides from (2) the reflection due to the mode converter/MMI-tapers, (3) from the reflection in the MMI or even (4) the reflections occurring at the interface between the silica cladded and water cladded slot waveguide. From numerical simulation, it is expected that the reflection due to the plasmonic waveguide (1) is at least 10 dB larger than the other three sources. More advanced mode matching at those four critical points will ultimately reduce the reflection and therefore improve the performance of our MMI-based circuit. For instance, the reflection due to the plasmonic waveguide could be better controlled by a more precise deposition technique such as the atomic layer deposition of gold that furthermore results in an increased Raman response [154] (see Section 6.1.1 for an example of such a plasmonic slot made with ALD gold). If the reflection in the MMI would be the limiting factor, the MMI can be designed with a central wavelength (closer) to 785 nm which would reduce the pump reflection to the output port by an extra -30 dB (from simulations in Lumerical). Another way to improve the architecture is to engineer the MMI so that it acts as a wavelength division multiplexer [155] transmitting the excitation beam from the input port (port 1) to the Raman sensors (port 3,4) with a high transmission and transmitting it inefficiently from the Raman sensors to the output port (port 2) while the scattered Raman light is efficiently collected in the output port. This would lead to a reduction of the reflected pump light in the output waveguide therefore improving the SBR for even longer/more complex analyzing circuits. Our first option is to re-design the MMI itself. In literature, angled-MMIs (A-MMI) have been proposed to separate the different frequency bands [156] (see Fig. 5.11). The frequency band captured by each output arm corresponds to a different part of the Raman spectrum. The pump frequency is efficiently transmitted towards the Raman sensor but is not captured by any of the output waveguides. The second

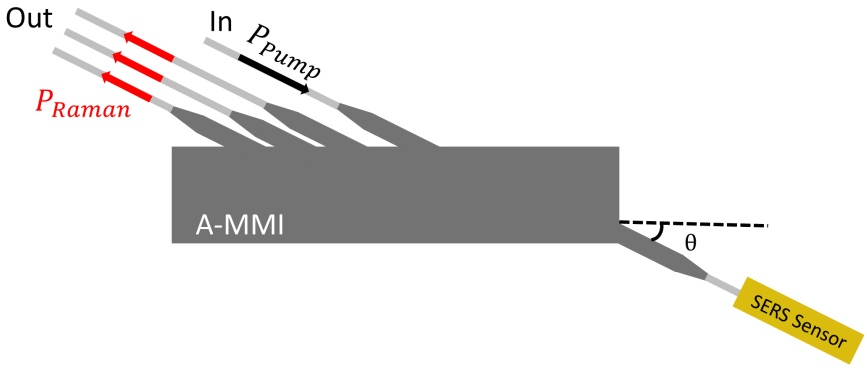


Figure 5.11: Proposed wavelength division multiplexer design using an angled-MMI. The axial positions of the input/output waveguides correspond to the self-imaging points of the different wavelengths propagating in the Raman arm [157].

option is to simply use the current MMI design and make a wavelength division multiplexer out of it by using interference. If we simply place the two SERS sensors $\sim \lambda/4$ apart, the reflected pump contributions from both SERS sensors will be in anti-phase compared to each other (see Fig. 5.12). The destructive interference of both contributions at the output port will then make sure that the reflected excitation beam is not coupled-back into the output waveguide. The Raman scattered light of the two SERS will still efficiently couple to the output port since Raman scattering is an incoherent process and therefore both contributions will maintain a random phase relationship compared to each other despite the extra path difference.

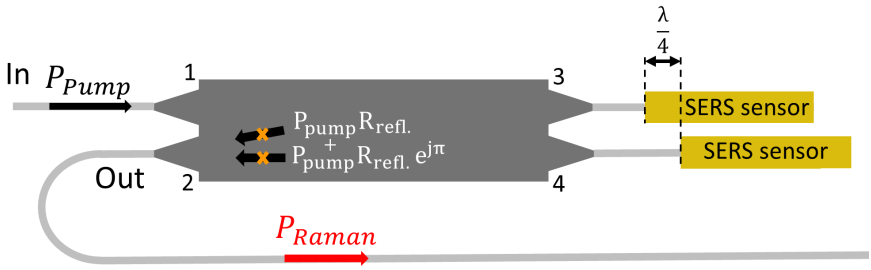


Figure 5.12: Proposed wavelength division multiplexer design using our current MMI design with the plasmonic sensors simply placed $\lambda/4$ apart.

Another improvement strategy consist in integrating a rejection filter [17] or lattice filter [158] on the output arm such that the back-reflected pump light can be further suppressed and does not generate a significant background contribution in

the output waveguide. The total pump rejection of our device (R_{tot}) can then be expressed as a combination of the pump rejection by the architecture itself (R_b) and the rejection by the filter on the output waveguide ($R_{filter,out}$). In the future we can design our device such that the background contribution of the output waveguide is no longer a relevant background source compared to the background of the SERS sensor itself (i.e. $BG_{out}/BG_{acc.} \sim 0.1$). In Section 5.2.3.1 (Eq. 5.24) we showed that the maximum allowed output waveguide length (L_{out}) for this condition to hold can be expressed as:

$$L_{out} = 0.1 \times \frac{L_{A_1}}{3\,(R_b + R_{filter,out})} \tag{5.32}$$

where L_{A_1} is the access waveguide length (which also includes the background contribution due to the plasmonic slot itself, see Table 5.1). We can see that we can control the allowed output waveguide length by tuning the total pump rejection of our device. The relationship between the pump rejection and the corresponding output waveguide length is plotted in Fig. 5.13. The required pump rejection can be achieved by improving the architecture (as described above) and/or increasing the filter length. Note that we showed in Fig. 5.9 that integrating a rejection filter of maximum length on the output waveguide (2 mm, with -65 dB rejection) would only lead to a minor deterioration of the SBR ($\sim$ –10%, for R_b = –27dB). In practice such a long filter is not needed as a filter of $\sim$ 100 μm (-30 dB rejection) would already suffice for most applications (under the condition that R_b = –27 dB as encountered in this work).

Now since since the full integration of the Raman spectroscopic system will be challenging in the short term, practical implementations will most likely make use of fiber-packaged NWERS (SERS) sensors linked to off-chip components that are difficult to integrate [146]. Since these input and output fibers will have a background contribution associated with them it is also interesting to look at the required Stokes/pump rejection of our device (R_{tot}) for these background contribution to be negligible (i.e. $BG_{fiber}/BG_{acc.} \sim 0.1$). Let us now first consider the case where a pump rejection filter (lattice filter) is integrated on the output waveguide before the light is coupled into a optical fiber. In Section 5.2.3.1 (Eq. 5.28) we showed that the maximum allowed length for the output fiber $L_{fiber,out}$ in that case corresponds to:

$$L_{fiber,out} = 0.1 \times \frac{G_{SiN}\, L_{A_1}}{G_{fiber}\,(R_b + R_{filter,out})} \tag{5.33}$$

where G_{SiN} is the spontaneous Raman gain factor for the dielectric slot waveg-

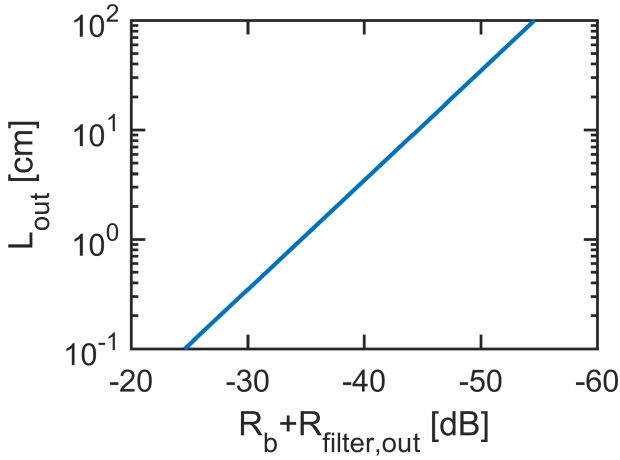


Figure 5.13: Relationship between the maximum allowed output waveguide length (L_{out}) (for a negligible output waveguide background contribution) and the total pump rejection of our device (achieved through a combination of the rejection of the MMI-architecture (R_b) and the rejection by the filter on the output waveguide ($R_{filter,out}$)).

uide and G_{fiber} the spontaneous Raman gain factor for the optical fiber, with $G_{SiN}/G_{fiber} = 42$. Analogously, the same filter can be integrated on the input waveguide in order to separate the incoming pump light from any background contribution accumulated in the fiber proceeding the Raman sensor (since they allow to collect both the pump and Raman light at different ports). In Section 5.2.3.1 (Eq. 5.31) we derived a similar expression as Eq. 5.33 for the input fiber length:

$$L_{fiber,in} = 0.1 \times \frac{G_{SiN}\, L_{A_1}}{G_{fiber}\,(R_b + R_{filter,in})} \tag{5.34}$$

under the assumption that $R_b(\lambda_s) = R_b(\lambda_p)$ we can see that the only difference between both expressions corresponds to $R_{filter,in}$ as the integrated filter on the input waveguide filters the Raman background (from the fiber) compared to the pump wavelength for the output waveguide. The relationship between the total Stokes/pump rejection and allowed input/output fiber length is however the same for both fibers. Therefore we can make a singular plot where we use R_{filter} (representing $R_{filter,in}$ for the input fiber and $R_{filter,out}$ for the output fiber) and L_{fiber} (representing $L_{filter,in}$ for the input fiber and $L_{filter,out}$ for the output fiber), see Fig. 5.14.

Note again that the effect of integrating the Stokes/pump rejection filter on the input/output waveguide will only have a modest impact on the SBR ratio. The conclusion of Fig. 5.9 (i.e. a 2 mm rejection filter only leads to a minor deterioration of the SBR $\sim$ –10%) also holds for a Stokes rejection filter of 2 mm integrated on

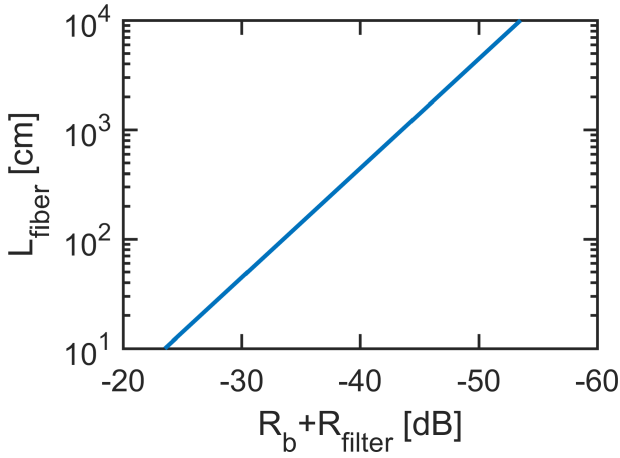


Figure 5.14: Relationship between the maximum allowed fiber length (L_{fiber}) (for a negligible input/output fiber background contribution) and the total Stokes/pump rejection of our device (achieved through a combination of the rejection of the MMI-architecture (R_b) and the rejection by the filter on the input/output waveguide (R_{filter})).

the input waveguide (where we assume that $R_b(\lambda_s) = R_b(\lambda_p)$, the port-to-port MMI transmission is $T_{MMI} = 0.5$ for both λ_s & λ_p and $R_b = -27$dB as encountered in this work).

5.3.2 Conclusion

In conclusion, our novel configuration based on a 2×2 MMI works effectively. It allows to extract the Raman scattered light out of a Raman sensor while enabling further integration with an analysing photonic circuit i.e. lasers, filters and spectrometers which are required for a complete Raman-on chip spectrometer. The penalty in terms of extra photon background generated is minimal. The background was shown to originate from the back-scattered Raman background in the short access waveguide (& Raman sensor) and from the output waveguide. The photon flux generated in that output waveguide is due to the same process occurring in the forward direction as the result of the presence of a remaining excitation beam. We found no contribution of the MMI to the photon background. This is an advantage of the MMI over a directional coupler that is able to perform the same functionality. We discussed the possible improvement to our design and showed that the focus should be put on reducing spurious reflections of the excitation beam on its path to the Raman sensor. In the future, an MMI having a spectral functionality may help to improve the capability of our architecture further.

6

SE-SRS on nanoplasmonic slot waveguides

SE-SRS on nanoplasmonic waveguides brings the advantage of boosting the strength of the generated Raman signal by three orders of magnitude compared to on-chip SERS while at the same time providing an improvement in the SBR by three orders of magnitude compared to on-chip SRS (see Section 4.5 for additional details). In this chapter, we will experimentally explore SE-SRS on nanoplasmonic waveguides and report our findings. Firstly, we discuss the sample fabrication, followed by the description of the measurement setup and the experimental verification of its correct functioning by performing SRS measurements on a optical fiber and a dielectric slot waveguide. Then, we present the result of our SE-SRS measurement on the nanoplasmonic slot waveguide and demonstrate that a combined photothermal and thermo-optic effect in the gold material induces a strong background signal that prevents the detection of our chosen analyte. Next, these experimental results are backed up by theoretical estimates on the strength of the photothermal effect and further confirmed by thermal simulations for which a 3D model of the nanoplasmonic slot is build. Finally, we propose several methods to reduce or counteract this background.

The text/figures of this chapter have been adapted from my published paper [90].

6.1 Sample fabrication

6.1.1 Nanoplasmonic slots

Photonic circuits of our Raman sensor are fabricated on 200 mm oxidized silicon wafers by depositing 300 nm of silicon nitride on top of a 3.3 μm silica layer through a plasma-enhanced chemical vapor deposition process (PECVD) (Imec's BioPIX platform [159]). Subsequent patterning is performed by deep UV-lithography and reactive ion etching. In this way, a 4 mm long (l) slot waveguide with width (w) × height (h) = 700 × 300 nm^2 and a 150 nm wide slot are fabricated (see light grey parts of Fig. 6.1 c). A silica top-cladding of 1 μm is deposited over the full dielectric slot length except for an open-cladded region of w × l = 50 × 10 μm^2 in the middle of the chip.

Nanoplasmonic slot waveguides are fabricated in the open-cladded region by firstly covering the dielectric slot with a 60 nm conformal Al_2O_3 layer via ALD (in order to reduce the slot gap) followed by the sputtering of a 2-nm Ti adhesion layer. Afterwards, an open window is defined by using AZ 5214 photo-resist and contact lithography corresponding to the region where the gold needs to be deposited. This window is 1.8 μm along the direction of the waveguide and is 12.7 μm wide.

We have two varieties of samples that differ by the gold deposition method: a gold sputtering process (detailed fabrication steps are given in Section 5.1.2) or atomic layer deposition (ALD). Gold deposition through ALD distinguishes itself by being more conformal and allowing a better controlled growth of the gold thicknesses than the sputtering alternative [154]. Furthermore, it leads to the formation of semi-continuous nanotrenches within the gold that further increases the plasmonic enhancement. In [160, 161] it has been shown that this leads to an increase in the generated Stokes signal by an order of magnitude ($P_{S,ALD}/P_{S,sputtered} \approx 10$) when probing a monolayer of NTP on the nanoplasmonic waveguides using SERS. A similar increase in the generated Stokes signal can be expected in our SE-SRS experiments for the ALD versus sputtered gold waveguides. The ALD gold layers are deposited at a substrate temperature of 100 °C using a home made ALD reactor with a base pressure of 2×10^{-6} mbar. Au is deposited using Me_3AuPMe_3 and H_2 plasma (20% H_2 in Ar) as the precursor and reactant with a 1 mbar and 6×10^{-3} mbar pressure respectively. The readers are suggested to consult [154] for the complete details on the ALD Au deposition process.

Fig. 6.1 a, b provide the SEM pictures of the fabricated plasmonic samples used during experiments which show that both slot gaps are open. Using the definitions of the thicknesses at various locations provided by Fig. 6.1 c, for the sputtered gold sample the inner side walls of the slot t_{slot} = 5.5 nm, the outer

side walls of the plasmonic waveguide t_{side} = 30 nm, and the horizontal surfaces next and on top of the waveguide $t_{top} \approx t_{next} \approx$ 100 nm. For the ALD gold $t_{top} = t_{next} = t_{side} = t_{slot}$ = 23 nm and the alumina thickness is set to 48 nm. Note that for the thermal simulation later in this chapter Fig. 6.1 c will be used which is modeled on the sputtered slot in Fig. 6.1 a. Both the ALD and sputtered gold nanoplasmonic waveguide have a length of 1.8 μm with 2 mm access waveguides on each side.

Two separate chips are fabricated containing six plasmonic slots each, with gold deposition by sputtering and ALD respectively. Besides the plasmonic slot waveguides, each chip also has two fully-cladded dielectric slot waveguides without any plasmonic structure. Those dielectric waveguides serve as references for our measurement apparatus.

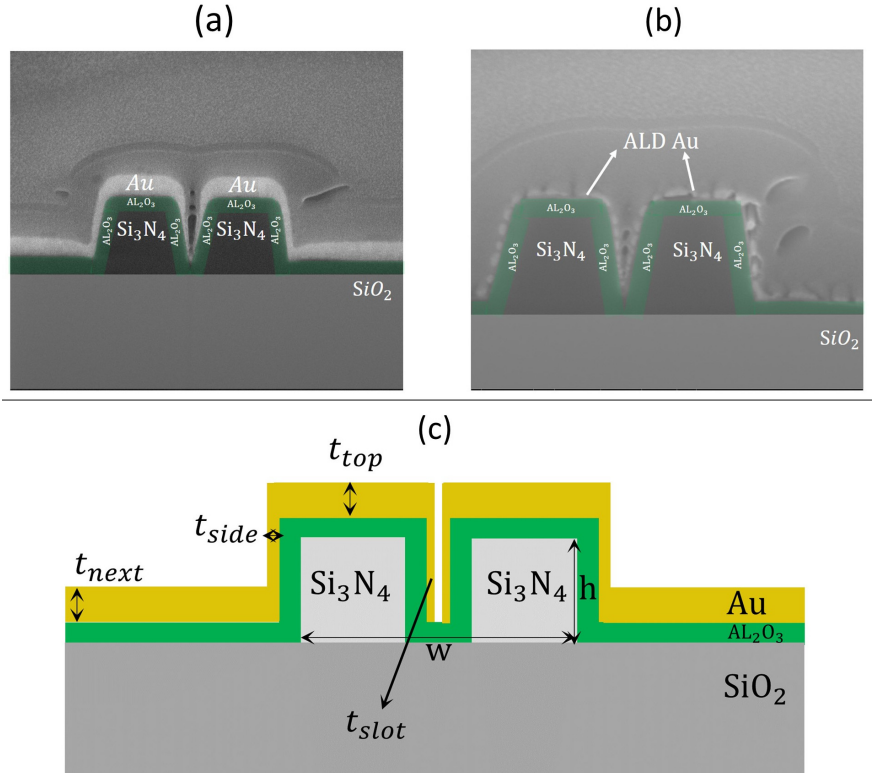


Figure 6.1: a) SEM image of the plasmonic slots fabricated using gold sputtering deposition b) and ALD gold deposition c) Schematic of the plasmonic slot modeled on the sputtered gold sample with the different gold thicknesses (t) and the width (w), height (h) of the silicon nitride slot waveguide indicated. The dimensions of a), b) are reported in the body of the text.

After fabrication, the Raman sensors are functionalized with a monolayer of 4-Nitrothiophenol (NTP) that selectively binds to gold but not to silicon nitride avoiding any extra contribution of SRS from the access waveguides themselves, see Section 5.1.2 for fabrication details (in the current measurements the NTP and plasmonic slot are exposed to air and no droplet of water is applied on top of them). The presence of the NTP monolayer on the plasmonic slot is verified by top-down measurements using a commercial confocal Raman microscope.

6.1.2 Suppressing cavity formation

Dicing and polishing are employed to terminate both chips with a 15 deg. angled facet in order to suppress back-reflections [35]. Those reflections result in the formation of a low quality Fabry-Perot cavity between dielectric slot end-facet and the plasmonic waveguide. Furthermore, chips are placed under a 30 deg. angle when measuring in order to suppress the formation of an unbalanced Mach-Zender interferometer. If not addressed both effects can swamp the (SE-)SRS signal, more details are provided in Section 6.2.3.1.

6.2 Setup & Verification

In this section we discuss the setup employed in our (SE-)SRS experiments followed by a experimental verification of its functioning on a optical fiber and dielectric reference waveguide.

6.2.1 Lock-in detection scheme

The (SE-)SRS setup used in this work is presented in Fig. 6.2 and validated in a previous study [19]. A lock-in detection scheme is employed in order to detect the NL-Raman gain of the Stokes beam (red). In our setup the amplitude of the pump beam (green) is modulated and sent through the plasmonic slot where it induces an intensity increase on the Stokes intensity via the (SE-)SRS process by an amount ΔI_S. This modulated increase of the Stokes beam is then picked up by the lock-in amplifier. By tuning the Stokes wavelength, the Raman spectra for a given analyte of interest is probed. In order to ease the separation of both beams prior to detection both beams are counter-propagating. The pump beam is originating from a laser diode and is sinusoidally modulated via its current resulting in a modulation depth of 30% on the intensity (peak-to-peak) at the wavelength of 785 nm. The intensity modulation on the pump is provided by the lock-in amplifier through the direct

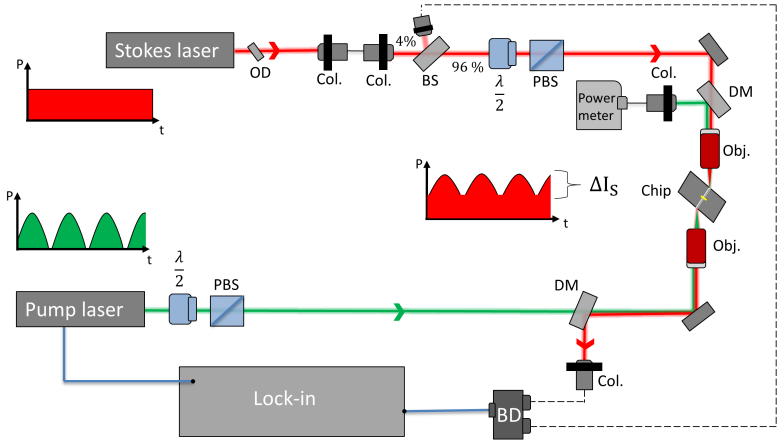


Figure 6.2: Schematic of the setup for the SE-SRS measurements. λ/2: half-wave plate, PBS: polarization beam splitter alligned to the TE-polarizion of the waveguide, DM: dichroic mirrors, Col: collimators, OD: Neutral density filter, Obj: Mitutoyo Plan Apo 50 × (NA=0.65), BD: balanced detector, BS: 4 %-reflective beam splitter.

modulation of the drive current. The Stokes beam is originating from a tunable CW Ti:sapphire laser. Half-wave plates and polarization beam splitters are used in order to control the power and ensure that only TE-polarized light is coupled into the chip. Dichroic mirrors are employed to separate the counter-propagating beams and collimators to collect light into optical fibers. The Stokes beam is eventually measured using a balanced detector to remove any fluctuation from the source itself. The Raman gain in the sample imprints a shallow modulation on the Stokes beam that is picked up by a lock-in amplifier. A 4 % power tap of the original Stokes beam (not affected by the Raman interaction) is provided as a second input to the balanced detector.

The (SE-)SRS spectra presented in this work (excepted for the optical fiber measurement) are measured with the modulation frequency set to 1 MHz, the build-in gain of the balanced detector set to 10^5, the lock-in amplifier to a 500 ms time constant and the optical power 25 mW and 20 mW before the objectives for the pump and Stokes beams respectively.

The presented spectra are obtained by averaging two successive measurements, applying a Savitzky–Golay filter [162] and normalizing by the pump and Stokes power. Measured data points have a uncertainty of $\sim$ 20% between successive measurements which originates from a combination of alignment drifts, the shot noise contribution due to the Stokes laser and small sample variations (a more

detailed discussion is presented in [35]).

6.2.2 Calibration: SRS on a optical fiber

The proper operation of our (SE-)SRS setup is first confirmed by the acquisition of reference SRS spectra on a commercially available optical fiber with a silica core. In this case the chip and objectives of Fig. 6.2 are replaced by two collimators to collect light from free-space and a L= 6 m long optical fiber. A pump and stokes power of 1.3 mw (modulated pump power) & 2.8 mW respectively are propagating in the fiber and a 1 MHz modulation frequency is set on the lock-in amplifier together with a 500 ms time constant and a gain of 10^4 . Fig. 6.3 shows the result of this measurement and clearly has a good overlap with the spectrum presented from [163].

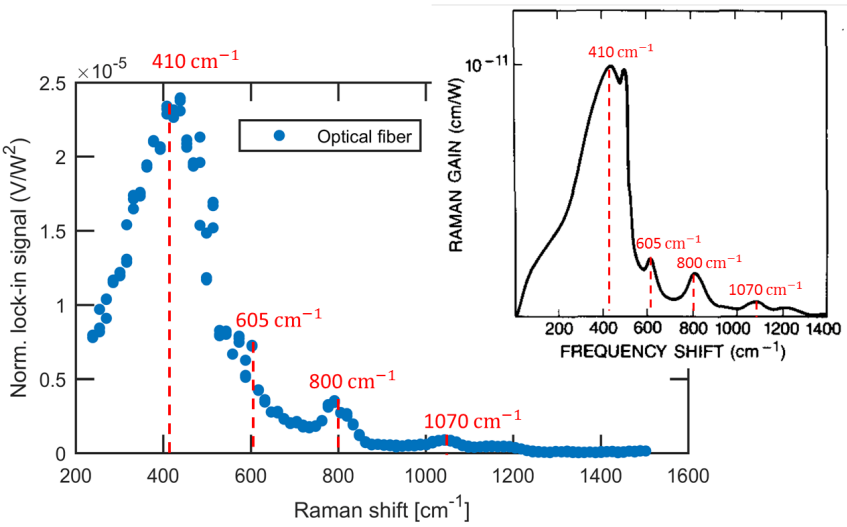


Figure 6.3: SRS-spectrum of a optical fiber (blue dots) using our experimental setup together with the Raman gain spectrum of silica glass as a comparison (from [163]).

Now let us calculate the stimulated Raman gain that we obtained from our measurement and comparing it to the peak Raman gain coefficient of [163]. Recall from Section 2.2.2 that we know that the stimulated Raman gain of the optical fiber can be calculated as:

$$g_{stim} = \frac{P_{SRS}}{P_P P_S L} \approx 0.0035 \pm 0.0007 \frac{1}{Wm} \tag{6.1}$$

where we neglect the loss in the optical fiber.

The peak Raman gain coefficient of [163] (see Fig. 6.3) corresponds to $G_{peak,lit.} = 10^{-11}$[cm/W]. Which can also be written as $G_{peak,lit.} = g_{stim,lit.}$[1/(cmW)]$\times$ $A_{lit.}$[cm^2] with $A_{lit.}$ the fiber core area. We can now calculate that $g_{stim,lit.} = 0.0021$/(Wm) (using $A_{lit.} = 50\mu m^2$).

Since this value was obtained at a pump wavelength of 1 μm and since $g_{stim,lit.}$ scales inversely with the wavelength ($1/\lambda$), its stimulated Raman gain at 785 nm will correspond to $g_{stim,lit.} = (1000/785) \times 0.002$ [1/(Wm)] ≈ 0.0025 [1/(Wm)]. This value corresponds reasonably well with the value we report in Eq. 6.1 ($\approx$ 0.0035 [1/(Wm)]).

6.2.3 SRS using dielectric slot waveguides

Next, the SRS spectrum of the fully-cladded reference waveguides (fabricated together with the plasmonic slots see Section 6.1.1) are acquired. This measurement is more challenging than that of the optical fiber as the formation of cavities and stray light paths on the chip need to be properly suppressed otherwise they can swamp the SRS signal. Furthermore, other focus also needs to be laid on preventing back-reflections on the chip facet to cause instability in the Stokes laser and to perform the measurement into the correct modulation frequency regime in order to not be hindered by unwanted signals that are picked up by the lock-in. First we will discuss these different challenges into more detail before presenting the successfully measured SRS spectrum of the dielectric reference waveguides.

6.2.3.1 Suppressing Kerr-induced interferometic signals

Since our SRS experiment makes use of a lock-in detection scheme, parasitic signals competing with the SRS signal can be induced by the modulated pump beam that leads to a refractive index change through the Kerr effect (dominant for the dielectric waveguides discussed here [35]) or a thermal effect (dominant for the plasmonic slot, discussed in Section 6.3.2). For the dielectric waveguides considered here, a modulated nonlinear phase change ($\Delta\phi \propto n_2 I_p$) will be induced on the Stokes beam by the modulated pump beam through the Kerr effect. This nonlinear phase change finally results in a parasitic signal being picked up by the lock-in. To better understand this phenomena, it is first important to note that a nonlinear phase change of the Stokes beam does not directly lead to a magnitude change. There are however two interferometric paths on the chip that can result in that change. Firstly, in our dielectric slot waveguide, a Fabry-Pérot cavity (FPC) can be formed between the two end facets of the waveguide such that a fraction

of the propagating Stokes beam is reflected at the facet and is coupled back to the waveguide. The presence of a modulated pump beam leads to varying nonlinear (NL) phase change of the Stokes beam through the Kerr effect. The FPC will convert this NL phase change to a transmission variation of the Stokes beam with the wavelength (with its own characteristic fringe pattern). In Section 6.3.2 we will give a detailed calculation on the effect of a FBC on the Stokes transmission. If unaddressed the presence of such a FPC causes the magnitude of the Stokes beam to change and results in a fringe pattern that is two orders of magnitude stronger than the SRS signal [35] (see Fig. 6.4 a). Therefore, the facets of our waveguide are polished at a 15 deg. angle so that the Fresnel reflection does not couple back into the waveguide (see Fig. 6.4 c). Secondly, a Mach-zender interformeter (MZI) can form because the stray light being guided in the optical cladding can find its way to the other side of the chip where it interferes with the light propagating in the waveguide. Since the light in the waveguide core is highly confined, the modulated pump will induce a nonlinear phase change to the Stokes beam unlike the cladding mode which experience a negligible Kerr-induced phase shift [35]. As only the waveguide mode experiences this extra phase change, the interference between the two paths is altered resulting in the output of this MZI to be a varying with the modulation frequency. Again this phenomena causes a fringe pattern in the recorded spectrum when varying the Stokes wavelength (see Fig. 6.4 b). To avoid this, the entire chip is set at angle of 30 deg, such that path of the stray light no longer overlaps with that of the light of the propagating beam (see Fig. 6.4 c). The parasitic signal is therefore suppressed and no longer appears in the recorded SRS spectrum.

With both measures of Fig. 6.4 c taken, no such fringe pattern is observed in our measured spectra as they are successful in sufficiently suppressing both phenomena. A more complete discussion on these parasitic signal due to FP and MZI formation for on-chip SRS can be found in [19, 35].

6.2.3.2 Solving Laser instabilities

Despite the absence of fringe patterns, the recorded spectra do not obviously show the broad Raman features of the PECVD Si_3N_4 [13]. While recording these spectra it was noticed that the captured lock-in signal due to the dielectric slot varied substantially over time ($\sim$ 80% per second) while the Stokes wavelength and power were kept fixed. A meticulous investigation by which the different parts of the setup (Fig. 6.2) were separately tested, revealed that this instability originates from a fraction of the Stokes laser reflecting on the chip facet and coupling back into the laser cavity of our Ti:Sapph laser. The unstable laser makes for an incorrect normalization and results in a distortion of our recorder SRS

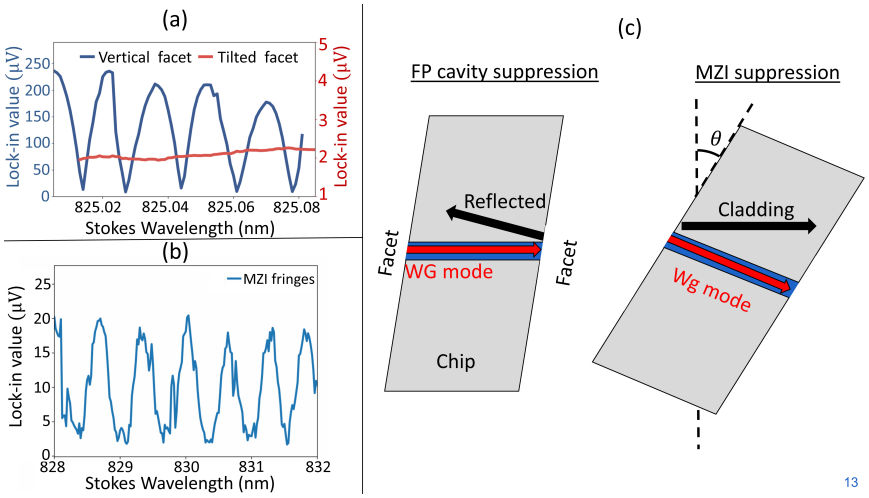


Figure 6.4: a) Illustration of the FP-cavity fringes that appear if output the waveguide facet is vertical and disappear if it is titled [from [35]] b) Illustration of the MZI fringes [from [35]] c) Both measures (facet titling and setting the chip on an angle) taken to suppress the FP and MZI parasitics.

spectrum. Such a problem is not encountered for our earlier SRS measurements on the optical fiber since the coupling to collimators is much more efficient than to Si_3N_4 waveguides. This issue was solved by placing a neutral density filter (rejecting Stokes wavelengths) under an angle in front of the laser. In this way it acts as an isolator such that any back-reflections do not couple into the laser cavity.

6.2.3.3 Modulation frequency operating regime

The modulation frequency (f1) of the pump beam in the (SE-)SRS experiments is set to 1 MHz. This frequency is chosen as a trade-off between on one hand suppressing the laser noise and any thermal sources (for which higher frequencies are better) and on the other hand avoiding the lock-in offset (see Fig. 6.5) that varies with the modulation frequency.

This lock-in offset (blue line in Fig. 6.5) is measured by using the detection scheme in Fig. 6.2 (with the dielectric slot included) and by simply blocking the optical path of the Stokes beam after chip right before it couples into the collimator towards the balanced detector. Next, this block is removed and the resulting variation of the SRS signal with the modulation frequency is shown in red. It can clearly be seen that the lock-in offset is also present in the spectral profile of the SRS signal. Since the offset can reach signal levels in the order of $\sim 10^{-5}$V it is

important to choose the modulation frequency outside this range since the estimated SE-SRS signal is in the order of $\sim 10^{-6}$V, hence a 1 MHz modulation frequency is chosen in our case.

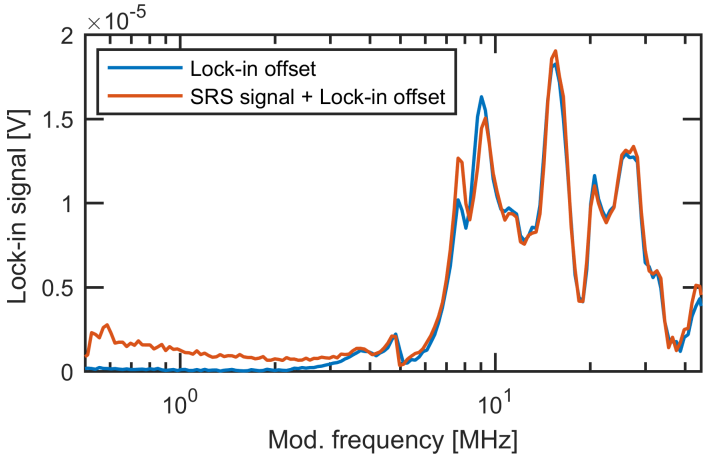


Figure 6.5: The lock-in signal when varying the modulation frequency in case the SRS signal of the dielectric slot is blocked (blue) and transmitted (red).

An effort was made to try to understand and suppress this lock-in offset by separating the lock-in detection scheme from the rest of the setup and: changing its physical location, wrapping aluminium foil around the different RF-cables, terminating all the RF outputs of the lock-in with 50 ohm impedance, replacing the balanced detector and laser diode, changing the RF-cables, etc. In this way the spectral behaviour of the lock-in offset could be altered but none off these efforts were fruitful in suppressing it. A few possible explanations for this offset include:

- The lock-in amplifier sends out a powerful modulation signal to the pump laser and part of the electromagnetic inference emitted by the RF-cable might be picked up by the lock-in
- The input offset of the detectors might be picked-up
- Frequencies emitted by radio stations might be picked up

Since the origin of this offset is not completely clear, it was concluded that it is sufficient to simply bypass it by setting the modulation frequency to 1 MHz (where the offset corresponds to $\sim$50 nV) such that its impact is minimized in the (SE-)SRS measurement.

6.2.3.4 Experimental result

Next, we present our successful SRS measurements on the dielectric reference slot. Fig. 6.6 displays the result of this measurement and shows good agreement with the corresponding spontaneous Raman spectrum. The spectra are taken over a wavelength range of 800-852 nm with a 2 nm resolution. The lock-in signal obtained for the dielectric slot waveguide is 2.5×10^{-5}V corresponding to a modulation depth $\Delta I_s/I_s = 4.7 \times 10^{-6}$.

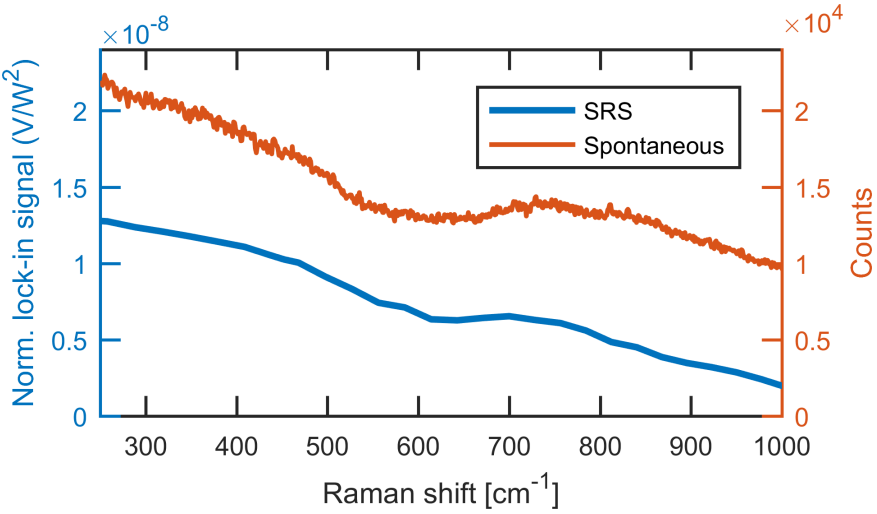


Figure 6.6: Raman background of a fully-cladded Si_3N_4 slot waveguide probed through SRS (blue & left-axis) and spontaneous Raman scattering (red & right-axis).

6.3 SE-SRS on a nanoplasmonic slot waveguide

After the experimental confirmation of the correct functioning of our setup, the next step is to move towards the SE-SRS measurement on a plasmonic slot waveguides. First, we present our experimental results which show that the SE-SRS spectra have a diverging response to what is expected, pointing towards a spurious SRS signal. Next, we investigate the frequency response of this spurious SRS signal and support the claim that it is a photothermal effect that generates this parasitic signal.

6.3.1 Experimental result

The experimentally captured SE-SRS spectra of the NTP-coated plasmonic slots are shown in Fig. 6.7. The spectra of the plasmonic slots made with sputtered gold (dotted blue) are probed over the 819 - 883 nm range with a 2 nm resolution, while the ALD gold slots (bold blue) are probed over the 804-948 nm range with a 4 nm resolution. A successful SE-SRS measurement should result in the appearance of the broadband Si_3N_4 Raman background with NTP features on top as presented by the light grey curve in Fig. 6.7 (obtained by a spontaneous Raman measurement of a plasmonic slot). The presented spectra obviously do not show such NTP peaks while they also do not match the photoemission signature that we did observe from the dielectric waveguides of Fig. 6.6. Moreover, the response obtained in Fig. 6.7 far exceeds the one we can observe using a dielectric waveguide. For the dielectric slot measurement of Fig. 6.6 an absolute lock-in signal of 2.5×10^{-5}V and a modulation depth ($\Delta I_s/I_s$) of 4.7×10^{-6} are captured. In comparison, for the plasmonic waveguides the absolute lock-in signal is $1.5 - 2 \times 10^{-4}$V corresponding to a modulation depth of $4.46 - 6.72 \times 10^{-4}$. This is much more than the modulation depth theoretically estimated from the NTP monolayer at 5.6×10^{-6}(see Section 4.3.2). Furthermore, such a strong signal can not originate from the dielectric access waveguides prior to the plasmonic slot as they are shorter and experience more loss than the reference dielectric slots (due to the presence of the plasmonic structure in the middle). This results in a reduction of the absolute SRS-signal with a factor 32 and a drop of the modulation depth with a factor 2 compared to the dielectric reference slots (because of the 12 dB extra absorption loss in the plasmonic slot and the reduction of the interaction length by half).

Detailed measurements with a 0.005 nm & 0.1 nm resolution in the 872-73 nm & 872-885 nm range respectively also did not reveal the appearance of Kerr-induced fringe patterns like those shown in Fig. 6.4 a,b. Furthermore, the free-spectral-range (FSR) of the MZI formed by the stray light in the top cladding and the propagating light in the access waveguide can be calculated as [165]:

$$\mathrm{FSR_{MZI}} = \Delta\lambda_{\mathrm{MZI}} = \frac{\lambda^2}{n_{gw}L_{w1} - n_{gc}L_c} = 7 \times 10^{-10}\mathrm{m} \tag{6.2}$$

where $n_{gw} = 1.98$ is the group index of the guided TE mode at the Stokes frequency (from modeling using Lumerical mode solutions), $L_{w1} = 2$mm the access waveguide length, $\lambda = 841$nm the Stokes wavelength and $n_{gc} = 1.48$, $L_c = 2$mm are the group index and the length of the cladding path (this is the length after which the access waveguide is no longer covered by a top cladding and the cladding mode no longer propagates [35]).

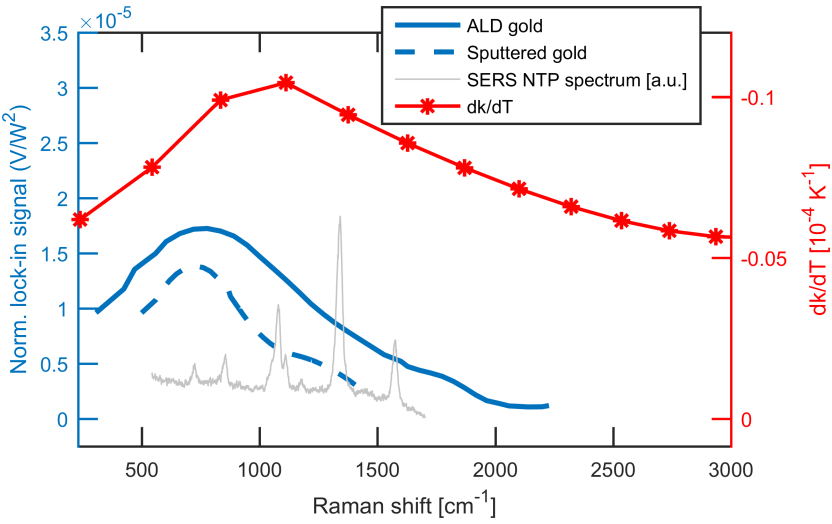


Figure 6.7: Recorded spurious SRS spectra for the plasmonic slot made with ALD gold (bold blue curve & left-axis) and sputtered gold (dotted blue curve & left-axis). Wavelength dependence of the extinction coefficient dk/dT of gold (red curve & right-axis) [164]. Indicative Raman spectra of NTP obtained using a regular spontaneous Raman setup and the sample under investigation are shown in grey (arbitrary units & left-axis).

Similarly the FSR of the FPC formed between the end facet of access waveguide and the plasmonic slot facet can be calculated as:

$$\mathrm{FSR_{FPC}} = \Delta\lambda_{\mathrm{FBC}} = \frac{\lambda^2}{n_{gw}L_{w2}} = 1.8 \times 10^{-10}\mathrm{m} \tag{6.3}$$

With L_{w2} = 2mm and n_{gw} = 1.98 again the group index of the plasmonic slot waveguide. This shows that both FSR's are too small to explain the response of Fig. 6.7 as being a part of their fringe pattern. In addition, the SRS spectrum shown in Fig. 6.6 is measured on the same chip (i.e. the angle of the tilted facet and the chip angle shown in Fig. 6.4 c are the same) such that it can be expected that those MZI & FPC fringes are sufficiently suppressed.

6.3.2 Photothermal effect

To explain the observed spectrum, we have to look elsewhere. It has been reported before that competing heterodyne optical processes such as photo-thermal effects, cross-phase modulation (XPM, as the Kerr-induced FPC & MZI discussed above), two-photon absorption (TPA) and transient absorption (TA) can result in a large

(thermal) background spectrally overlapping with the (SE-)SRS signal [61, 166]. Among those effects, some are fast (ps or faster) while others may be slow enough to be dependent on the frequency of the modulation we impart on the pump beam [167]. We have therefore investigated this response using the lock-in as a function of the modulation frequency (fixing the Stokes wavelength to 834 nm, the time constant to 75 ms and sweeping the modulation frequency from 0.15 MHz to 5 MHz). The result is presented in Fig. 6.8 and shows a clear decrease of the response for increasing modulation frequency. This points at a slow effect such as a thermal effect, with the other optical processes (XPM, TPA, TA) being considerably faster [168–175]. Note that in theory TA by induced fluorescence might be a slow enough process such that its contribution declines within our modulation frequency range. However, in practice it is very unlikely that this process would induce such a strong intensity modulation since earlier demonstrations of SERS on the plasmonic slot waveguide did not show any presence of fluorescence [16]. A closer look at the result presented in Fig. 6.8 reveals that the decay is best fitted by a bi-exponential. This may hint at the presence of two phenomena with different time scales.

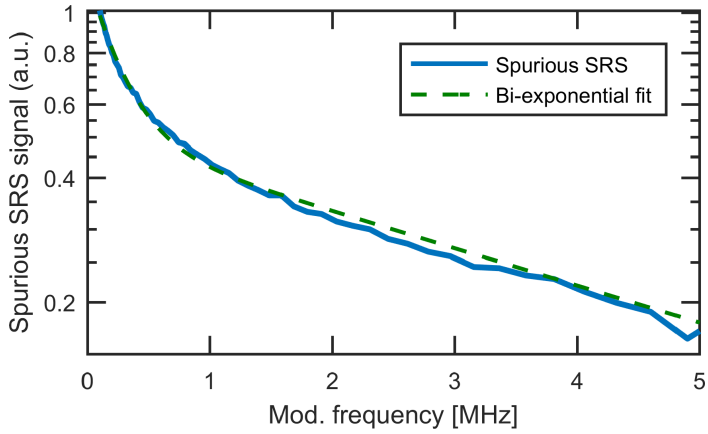


Figure 6.8: Effect of the modulation frequency on the experimentally captured spurious SRS signal (blue curve) and bi-exponential fit (dashed green).

The thermal effect that could account for our large response is a combined photothermal and thermo-optic effect explained as follows. A modulated heating of the gold occurs via absorption of the modulated pump beam. This periodic heating affects the absorption and refractive index of the gold [168, 176–178] that in turn modulates the Stokes beam. The change of the refractive index and absorption with temperature are quantified via the thermo-refractive coefficient (dn/dT) and thermo-extinction coefficient (dk/dT). These coefficients vary with the wavelength and we present the thermo-extinction coefficient measured by Wilson *et al.* [164] in Fig. 6.7 (red line). This spectrum is remarkably close

to the spectra we have measured, hinting at an effect dominated by the thermo-optic absorption. One should note that the contribution of the thermo-refractive coefficient was dominant in a recent SE-SRS experiment [61, 179] relying on a confocal microscope. In the present case, a contribution of the thermo-refractive effect to the signal collected in the lock-in amplifier is also possible. Indeed, the presence of a weak cavity in combination with a modulation of the refractive index would lead to a modulated Stokes signal (equivalent to the Kerr induced cavities of Section 6.2.3.1). Such cavity can be formed by the mode mismatch at the transition between the nanoplasmonic waveguide and the dielectric slot serving as access waveguides, such that the 1.8 μm long plasmonic slot itself act as the cavity. However, as already discussed in Section 6.3.1 the presence of such a cavity should be evidenced by the fringe pattern characteristic of a Fabry-Perot cavity. Given the geometry of the nanoplasmonic waveguide, the group index for the plasmonic mode is $n_g = 6.8$ (from modeling using Lumerical mode solutions), which gives a free-spectral range $FSR = \lambda_s^2/(n_g L) = 58$ nm for a cavity length $L = 1.8$ μm at a wavelength $\lambda = 841$ nm. A transmission measurement carried over 150 nm shows no evidence of a cavity. Therefore, the thermo-optic absorption seems a better candidate to explain our spectra.

Let us now make a calculation of the temperature increase needed in the plasmonic slot for the thermo-optic absorption to result in a modulation depth of 6.72×10^{-4} (corresponding to our recorded spurious SRS signal). Since we know that the modulated pump is converted into heat and it results in a temperature increase (ΔT_{Exp}) in the plasmonic slot, the change in the extinction coefficient (Δk) can be expressed as:

$$\Delta k = \frac{dk}{dT} \times \Delta T_{Exp} \tag{6.4}$$

From [164] we know that $dk/dT = -0.1045 \times 10^{-4} K^{-1}$, while ΔT_{Exp} is the parameter that we try to estimate. Δk can now be expressed as:

$$\Delta\alpha = \frac{4\pi\Delta k}{\lambda_s} \tag{6.5}$$

With $\Delta\alpha$ the change in the absorption coefficient and $\lambda_s = 841$nm the Stokes wavelength. Now let us express the power of the Stokes beam after being propagated through the plasmonic slot with $L = 1.8$ μm and the modulated pump being off:

$$P_{s,nM}(L) = P_s(0) \times e^{-\alpha L} \tag{6.6}$$

and in case the modulated pump is on:

$$P_{s,M}(L) = P_s(0) \times e^{-(\alpha+\Delta\alpha)L} \tag{6.7}$$

So that the increase in the Stokes power (due to the decrease in the absorption coefficient since dk/dT has a negative sign) corresponds to:

$$\Delta P_s = P_{s,M} - P_{s,nM} \tag{6.8}$$

such that the intensity modulation at the end of the plasmonic waveguide becomes:

$$\frac{\Delta I_s}{I_s} = \frac{\Delta P_s}{P_{s,nM}} = e^{-(\Delta\alpha)L} - 1 \tag{6.9}$$

Since we know that $\Delta I_s/I_s = 6.72 \times 10^{-4}$ we can calculate that only a modest temperature modulation depth (peak-to-peak) $\Delta T_{Exp} = 2.5$ K is sufficient to induce this experimentally captured spurious modulation depth on the Stokes intensity.

In the same way a calculation can be made using $dn/dT = 2 \times 10^{-4}$ K^{-1} and $\Delta T_{Exp} = 2.5$ K such that we can calculate the intensity modulation that the plasmonic slot cavity would induce for such a modest temperature increase.

The transmission of intensity of the weak Fabry-perot cavity formed by the plasmonic slot can be expressed as (adapted from [35]):

$$Tr = \frac{I_o}{I_i} = \frac{(1-R)^2 e^{-\alpha L}}{(1-Re^{-\alpha L})^2 + 4Re^{-\alpha L}\sin^2\phi_0} \tag{6.10}$$

where (I_i) is the intensity of the Stokes beam entering the cavity and (I_o) of it leaving and R=-26.9 dB is the reflectively at the plasmonic slot facet [128], L = 1.8 μm is the plasmonic slot length, α = 2.4 dB/μm the loss in the plasmonic slot [16] and ϕ_0 is the half-pass optical length:

$$\phi_0 = \frac{2\pi n_{eff} L}{\lambda} \tag{6.11}$$

with n_{eff} = 3.4741 the effective index of the Stokes beam in the plasmonic slot (from Lumerical). Now since we know that the effect of the modulated pump is to induce a refractive index change (Δn) through the thermo-optic effect we can include this effect in the half-pass optical length:

$$\phi = \phi_0 + \Delta\phi = \frac{2\pi}{\lambda_s}(n + \Delta n)L \tag{6.12}$$

where the refractive index change due to the thermo-optic effect is

$$\Delta n = \frac{dn}{dT} \times \Delta T_{Exp} = 5 \times 10^{-4} \tag{6.13}$$

Eq. 6.10 can now be expanded at the vicinity of $\Delta\phi = 0$ as:

$$Tr = \frac{(1-R)^2 e^{-\alpha L}}{(1-Re^{-\alpha L})^2 + 4Re^{-\alpha L}\sin^2\phi_0} + \overbrace{\frac{4R(1-R)^2 e^{-2\alpha L}\sin(2\phi_0)}{((1-Re^{-\alpha L})^2 + 4Re^{-\alpha L}\sin^2\phi_0)^2}}^{\Delta Tr}\Delta\phi + ... \tag{6.14}$$

The maximum modulation depth due to the Fabry-Pérot cavity can now be calculated for $\Delta T_{exp} = 2.5K$ and corresponds to $\Delta Tr_{max} = \Delta I_s/I_s = 1.4 \times 10^{-5}$ which is an order of magnitude smaller than that of the thermo-optic absorption. This agrees with our earlier statement that the spurious SRS signal generated by the thermo-extinction coefficient (dk/dT) is dominant over that of thermo-refractive coefficient (dn/dT).

6.4 3D Thermal simulations

In order to further investigate this thermo-optic effect we perform a time-dependent thermal simulation of our system (using COMSOL multiphysics). This way a value ΔT_{Sim} can be acquired to compare with the experimentally captured $\Delta T_{Exp} = 2.5K$ and check its validity. First, we introduce the assumptions made for the heat source in our model and then the custom-built 3D model of our device that includes the gold part of the nanoplasmonic waveguide, its silicon nitride core, the optical cladding and the substrate. Finally, we discuss our simulation settings and results obtained.

6.4.1 Heat source

Our thermal simulation models a time dependent heat flow. As a starting point, we have to make an estimate of the amount of heat that is generated at the nanoplasmonic slot waveguide as that will serve as a source term in our thermal simulation.

The source of this heat originates from the absorption of the (modulated) pump and Stokes laser propagating through the nanoplasmonic slot. The modulated part of the pump beam in the dielectric slot waveguide is set at $P_{P,Ac}$ = 1 mW (peak-to-peak, corresponding to the guided AC pump power used during experiments) and the coupling to the plasmonic waveguide is approximately –4 dB leading to 0.4 mW of optical modulated power propagating in it.

Then, we make the approximation that the loss within the plasmonic waveguide is solely due to absorption by gold and that the scattering loss is neglected. In that condition, all lost optical power is converted to heat. As the loss within the plasmonic waveguide is –4.3 dB, we have an absorbed power of $P_{heat,Ac} \approx 0.25$ mW. The same reasoning can be performed for the total DC optical power associated to the pump and Stokes beam giving $P_{heat,Dc} \approx 1.25$ mW. Thus, a total of 0.25 mW AC (sin. wave with a 1 MHz frequency corresponding to the modulation frequency used during experiments) and 1.25 mW DC heat source is assumed.

In our thermal simulation, we assume that heat is distributed uniformly in two regions corresponding to the inner walls of the plasmonic slot (see Fig. 6.9) as these regions overlap with the highest intensity regions of the simulated optical mode. This means that we neglect any temperature gradient that may arise along the propagation direction.

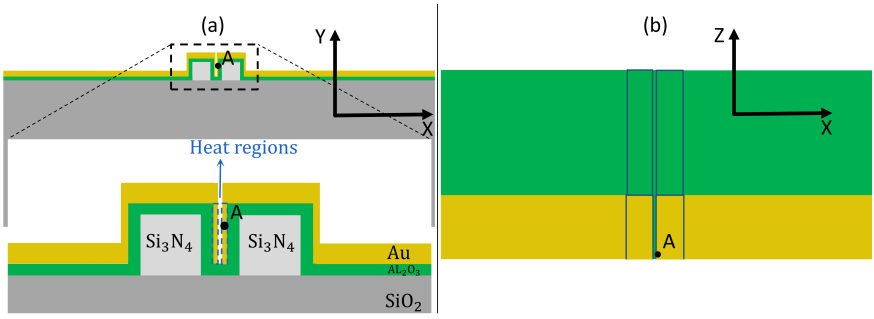


Figure 6.9: a) Cross section of the plasmonic slot build in the simulation environment with the thermal monitor in the slot gold (point A) and the two heat regions that originate from the absorption of the (modulated) pump and Stokes laser propagating through the plasmonic slot (dotted blue lines) b) Top-down view of the plasmonic slot (gold) showing the transition into the dielectric slot (green).

6.4.2 3D model of the nanoplasmonic slot

Thermal simulations are made using the Heat Transfer module in Solids and Fluids of COMSOL Multiphysics platform. A custom 3D model of the nanoplasmonic slot

is build (see Fig. 6.9) the dimensions of which are those associated to the sputtered gold samples (see Fig. 6.1 a, c). The ALD-gold samples are slightly different since their gold thickness are uniform and the gold layer consists of ALD gold patches instead of a uniform layer defined in the simulation environment. Note that SEM images of Fig. 6.1 reveal a discontinuous gold layer at the bottom of the slot for samples fabricated via sputtering (not clearly visible without extra magnification) but a continuous one for samples fabricated using ALD. Our simulations have been run for a scenario with and without a gold layer on the bottom but the result is practically insensitive to it owing to the fact that the fraction of gold constituting the bottom of the slot is insignificant at the scale relevant for the thermal dissipation: the entire nanoplasmonic structure and beyond.

In the 3D model, the gold and alumina sheets next to the plasmonic waveguide extended by 6 μm on each side transversely to the waveguide (along the x-axis), reflecting the geometry depicted in Fig. 6.9 a. The plasmonic slot has a length (l) of 1.8 μm along the propagation direction (z-direction in Fig. 6.9 b) after which the plasmonic slot transitions into the dielectric slot waveguides on both sides. The length of these dielectric waveguides and their surrounding (substrate, cladding) is assumed infinite by setting such a boundary condition. The silica cladding underneath the plasmonic slot has a thickness of 3.3 μm, the silicon substrate is approximated to be infinite and so is the air above the chip.The transverse dimension of the simulation window is 38 μm after which we also set infinite boundary conditions.

6.4.3 Simulation settings

The boundaries of the model are set as infinite element domains (which approximate infinitely large boundary layers with a constant temperature of 293.15 K at the end). Material properties are provided from COMSOL's built-in material library. Air is modeled using temperature dependent thermal parameters (e.g. thermal conductivity, etc.), while these parameters are constant for the other materials. The free convection flow of air is not considered in this model.

A 3D time-dependent study is performed on this simulation model over a 10 μs time range with a 8 ns time step. A thermal monitor is placed in the middle of the gold in the plasmonic slot (point A in Fig. 6.9) to monitor the evolution of the temperature in the slot over time.

6.4.4 Simulation result

The result of this time-dependent thermal simulation is presented in Fig. 6.10 and shows that the temperature in the plasmonic slot rises from 293.5 K to 352 K after 10 µs in quasi-steady-state with a remaining temperature modulation with an amplitude of ΔT_{Sim} = 5.8 K.

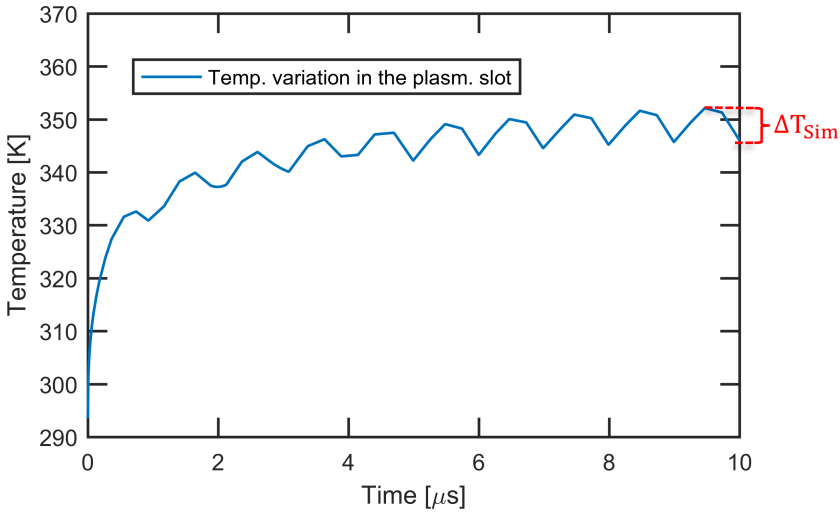


Figure 6.10: Variation of the temperature in the plasmonic slot (point A in Fig.6.9) using our custom build 3D simulation model and a 8 ns time step. The simulated temperature modulation (ΔT_{Sim}) is indicated in red.

The simulated ΔT_{Sim} = 5.8 K corresponds reasonably well with the value extracted from experiments ΔT_{Exp} = 2.5 K (see Section 6.3.2). The small discrepancy could be due to uncertainties in the very details of the geometry of the plasmonic waveguide (the size of the slot gap) as well as the coupling loss between plasmonic and dielectric waveguides (as it relates to the amount of light actually absorbed by the gold nanostructure). A full visualisation of the temperature distribution in the solved simulation model after 10 µs can be found in Appendix B.

Note that modelling the ALD gold slot could lead to even less discrepancy between simulation and experiment results. In this work we only modeled the sputtered gold slot as modeling the ALD gold slot is less straightforward due to the random distribution of the gold trenches which needs to be recreated in the simulation environment. In the future, image processing could be used on SEM images of the ALD gold in order to establish the size distributions of the gold patches which could then be modeled in the simulation environment. Since ALD

provides a more reliable layer thickness it would become even more straightforward to predict the exact coupling/absorption loss of the plasmonic slot in the simulation environment based on previous experimental loss-characterizations with the same layer thickness (under the assumption that random nature of the trenches does not provide too much variation between samples). This could further reduce the discrepancy between ΔT_{Sim} and ΔT_{Exp}.

Now since the "steady-state" ΔT_{Sim} in Fig. 6.10 is calculated for a 1 MHz modulation frequency it is interesting to see the effect of the modulation frequency on it. We make a comparison to the experimental captured variation of the spurious SRS signal with the modulation frequency (presented by Fig. 6.8 in Section 6.3.2). Remarkably, the simulation reproduces the bi-exponential behaviour we deduced from our measurement as is shown in Fig. 6.11.

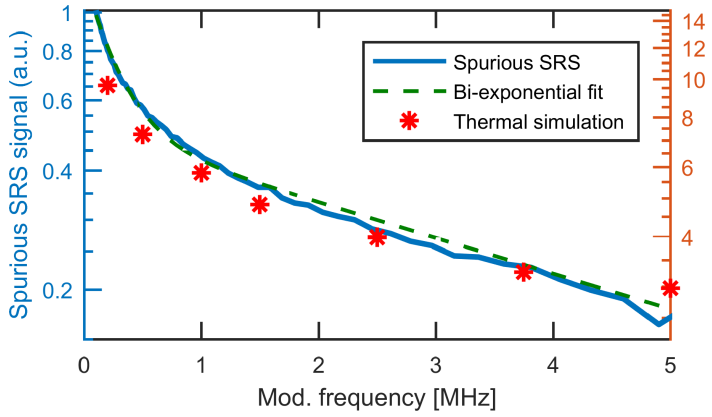


Figure 6.11: Effect of the modulation frequency on the simulated thermal modulation (ΔT_{sim}) in the plasmonic slot (red stars & right-axis). The experimentally captured spurious SRS signal (blue curve & left-axis) and bi-exponential fit (dashed green) are added in comparison.

The good match between the theoretical model and the experimental captured results further strengthens the argument that the spurious SRS signal has a thermal origin. Furthermore, the bi-exponential behaviour is recreated in the simulation environment which indicates that it is not sufficient to characterise the heat flow from the nanoplasmonic slot as a single phenomena. By simulating the amplitude of the thermal modulation at various positions in the vicinity of the plasmonic waveguide, we can understand why two heat transport phenomena are characteristic of the heat flow out of the waveguide.

Fig. 6.12 (a) and (b) illustrate the different positions where a thermal monitor is placed in the geometry of our structure in the transverse plane (cross section) and from the top respectively. The "steady state" thermal modulation amplitude

ΔT_{sim} is extracted (in the same manner as indicated in Fig. 6.10 using a 1 MHz modulation frequency) at the position denoted by letters A-J in Fig. 6.12 (a,b) and are summarized in Fig. 6.12 (c). The blue curve clearly drops slower then the other two and this demonstrates that the heat flows better along the X-axis than it does over the Y and Z-axis. This is actually meaningful as the gold layer making the plasmonic waveguide extends in a narrow strip extending 6 μm on each side along the X-direction. The thermal resistivity and capacitance of that gold strip are lower than the one of the silica cladding and silicon substrate associated to the heat flow along the Y-axis and than that for for the silicon nitride along the Z-axis. Hence higher modulation frequencies are supported in the gold compared to the surrounding materials. This results in a greater amplitude of ΔT_{sim} when moving away over the same distance from the plasmonic slot.

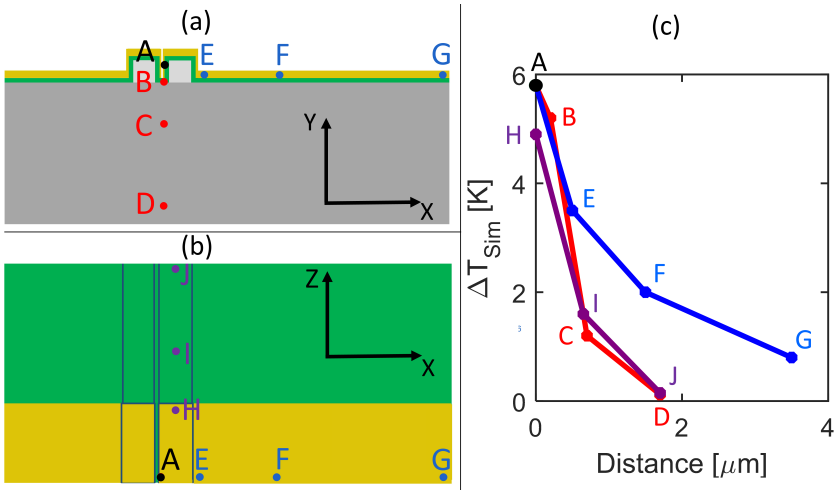


Figure 6.12: a) Cross section of the plasmonic slot with thermal monitors placed in the plasmonic slot (A,B), the silica (C,D) and along the gold sheet (E,F,G) b) Top-down view of the slot, monitors are placed along the propagating direction in the Si_3N_4 core (H,I,J) c) Variation of the thermal modulation (ΔT_{Sim}) with distance.

Finally, we can probe the impulse response of our thermal model in order to extract the time constants associated to our thermal system. Fig. 6.13 shows the response of our simulation model. By making a bi-exponential fit we can extract two time constant: τ_1 = 0.023 µs & τ_2 = 46.9 µs. The clear difference between both time constants further indicates that there is a faster thermal response in the gold compared to the surrounding materials. It is also not surprising that the response of the gold can be faster since gold is a good heat conductor and in our current plasmonic slot there is only a limited amount of gold (mass) compared to the silicon/silica layers.

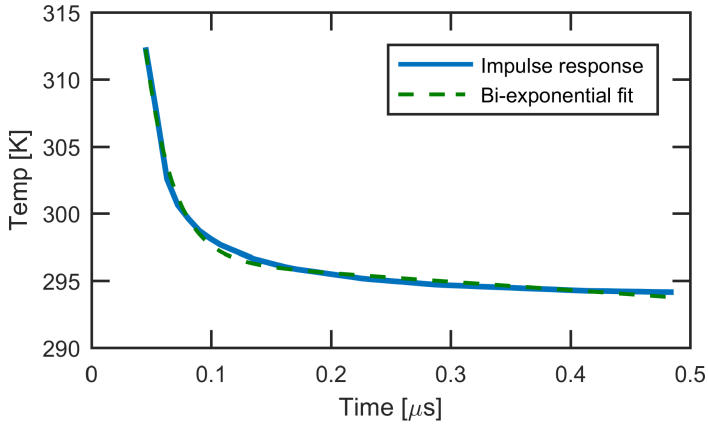


Figure 6.13: Impulse response of the nanoplasmonic slot (blue) for a 50 ns long pulse (with a 9 ns time step resolution). Fitted bi-exponential curve (green) to extract the two time constants τ_1 = 0.023 µs & τ_2 = 46.9 µs *(with amplitudes* A_1 = 111.4 K & A_2 = 296.8 K*).*

6.5 Mitigation & Conclusions

We believe all the evidence presented in this chapter is clearly pointing at the thermo-optic absorption mechanism to explain our spectra in Fig. 6.7. A remaining question can be raised concerning the wavelength shift of our spectra as compared the one of the thermo-extinction coefficient. A possible explanation could be that the thermo-extinction coefficient does not scale perfectly linearly with temperature. As the reported curve originating from [164] was obtained at temperatures from 298.15 K to 430.15 K, this may have some impact. As reported in Section 4.1, the field enhancement associated to our plasmonic waveguide is mostly nonresonant and thus spectrally far broader than the spectra depicted in Fig. 6.7. As compared to sputtered gold samples [16], the ALD-gold ones present a slightly more resonant field enhancement due to the presence of nanostructured gaps within the material that allow localized plasmons resonances [154]. We cannot reliably infer any difference between ALD and sputtered gold sample in the present investigation. It is also interesting to note for studies on SE-SRS that use gold nanoparticles as the plasmonic medium and a free-space excitation and collection (see Section 4.3), the presence of such a non-resonant thermal background was not reported or it could be sufficiently suppressed such that it could be subtracted [61, 179]. This is different for a waveguide platform as the generated signal has to pass through the same plasmonic structure in order to be collected while for the earlier SE-SRS demonstrations the generated signal is scattered into free-space. Therefore, the spurious signal can build up over a greater distance and become too strong for the SE-SRS signal to be distinguishable from it.

6.5.1 Mitigation strategies

Both in the present case or in the free-space demonstration of SE-SRS reported in [179], the photothermal/thermo-optic effect can in principle be mitigated. Below we provide several strategies that can be combined to achieve this goal:

- Looking at Fig. 6.11 one obvious mitigation strategy consists of increasing the modulation frequency as lock-in systems operating at 100's of MHz exist. Ultrafast SRS techniques making use of pulsed lasers and a reduced average power may also suffer less from the thermal effects [59].

- Another strategy is to increase the thermal capacitance of the gold sheet (i.e. the amount of heat to be supplied to a given mass of a material to produce a unit change in its temperature [180], [J/K]). As demonstrated in Fig. 6.12 higher modulation frequencies are better supported by the

gold sheets, such that lowering the thermal response at these frequencies can be achieved by increasing their thermal capacitance. In order to better understand the reasoning behind this strategy let us consider the generated heat in the plasmonic slot to only being able to flow into the gold sheets next to the waveguide such that the transfer function of this thermal system can be described as a simple first-order system with a thermal resistance (R_T) and thermal capacity (C_T):

$$TF = \frac{R_T}{1 + j\omega R_T C_T} \tag{6.15}$$

From Eq. 6.15 we can derive that if we work far below the cutoff frequency then we need to decrease R_T in order to reduce the thermal response while C_T doesn‘t really matter. Secondly if we work far above the cutoff frequency then we need to decrease C_T in order to lower the thermal response while R_T doesn‘t really matter. So increasing C_T for higher modulation frequencies will lower the thermal response at these frequencies.Therefore, a good strategy would be to connect the gold sheet to a bigger gold mass such that C_T would increase. Another strategy could be to reduce R_T of the nanoplasmonic waveguide by for example applying a good thermal conductor as top cladding, however decreasing the thermal resistance will weaken the thermal effect but at the same time make it more broadband thereby countering the effect of increasing C_T. Therefore, special care will need to be laid in optimizing the new design of the nanoplasmonic slot in order to sufficiently increase the gold mass to make the thermal-effect less broadband while still maintaining the thermal resistance high enough such that it does not counteract the increase of C_T. One way to achieve this could be by keeping the gold sheets sufficiently narrow before it connects to a greater gold mass. If the modulation frequency increase discussed in the first strategy is sufficiently above the thermal cutoff frequency R_T will not have too much influence on the thermal response and optimization efforts can mostly be focused on increasing C_T.

- Another mitigation procedure consists in removing the thermal modulation all together without affecting the Raman signal itself. This could be achieved by using a second pump beam with anti-phase modulation compared to the first one so that the heating in the plasmonic slot is constant at all time (see Fig. 6.14). This way the refractive index and extinction coefficient experienced by the Stokes beam remain constant and the spurious signal can be suppressed. The auxilliary pump beam could be set at a different wavelength either sufficiently remote from the first pump (< 600 nm or $>$ 1000 nm) so as to avoid any Stokes response in the region of interest or close to the first pump (e.g. 790 nm) in which case two identical superimposed but

shifted spectra will be measured and signal processing allows to disentangle them. An important aspect of this setup consists in making sure that the (SE-)SRS response induced by the first pump laser is not the same or comparable to that of the second pump beam. In that case the Stokes beam will experience the same generated (SE-)SRS amplification for both pump beams and will no longer be picked-up by the lock-in since both contributions are in anti-phase and cancel out the modulation (see Fig. 6.15 b). For the SE-SRS measurement, placing the second pump beam off the Raman resonance of the analyte should suffice in achieving this goal.

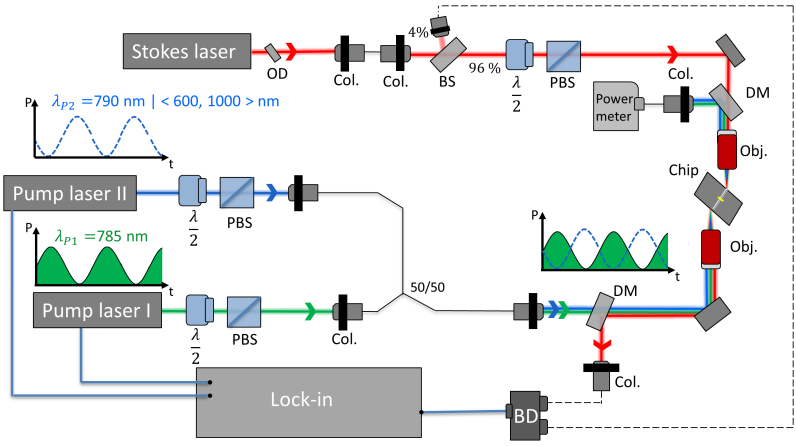


Figure 6.14: Schematic of the adjusted setup for the SE-SRS measurements with a second pump added (blue line).The two pump beams are combined by the 3-dB splitter (50/50). λ/2: half-wave plate, PBS: polarization beam splitter aligned to the TE-polarizion of the waveguide, DM: dichroic mirrors, Col: collimators, Obj: Mitutoyo Plan Apo 50 × (NA=0.65), BD: balanced detector, BS: 4 %-reflective beam splitter.

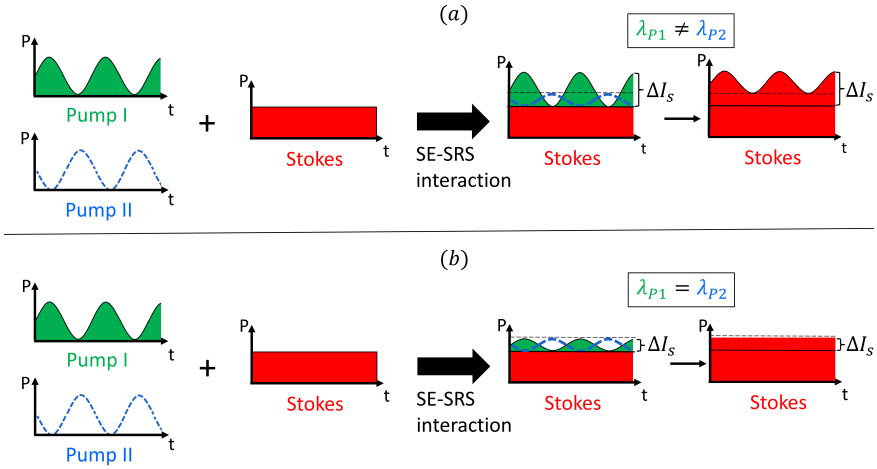


Figure 6.15: Effect of pump wavelength on the generated SE-SRS signal (ΔI_s) a) in case the SE-SRS response due to both pump beams sufficiently differs in strength such that the resulting SE-SRS signal will be modulated and picked-up by the lock-in. b) in case they have a comparable response and both contributions cancel out the modulation.

6.5.2 Conclusion

In conclusion, we have shown that while SE-SRS using nanoplasmonic waveguides is appealing, it does not come without challenges. The heating intrinsic to linear absorption in the gold nanostructure induces a thermo-optic effect that affects the Stokes beam overshadowing the Raman signature of the analyte (NTP in the present case). Not only is this thermo-optic effect two orders of magnitude stronger than the Raman response, but it also presents its own non-trivial spectrum overlapping with that of the thermo-extinction coefficient presented in literature. The thermal origin of this spurious SRS signal is confirmed by demonstrating its clear decline with modulation frequency which points at a slow process like the thermal effect while other common heterodyne processes competing with the (SE-)SRS signal like XPM, TPA and TA are considerably faster. Additionally supportive calculations show that only a mild (2.5 K) thermal modulation in the plasmonic slot is sufficient for such a strong spurious background to be present. This calculation is backed up by thermal simulations using a 3D model of the plasmonic slot which is probed under the same measurement conditions as we use during our experiments and shows a comparable thermal modulation (5.8 K) in the plasmonic slot. Furthermore, these thermal simulations recreate the bi-exponential frequency behaviour of the spurious SRS signal we deduced in our measurements which can be attributed to the faster heat flow in the gold sheets compared to the surrounding materials. All this evidence clearly points towards the fact that our captured spurious SRS signal has a thermal origin. In order to suppress this thermal background in the future we have introduced mitigation strategies that may be combined in future investigations in order to realize the actual potential of nanoplasmonic-based SE-SRS.

7

Conclusion and Perspectives

In the first part of this PhD project, we explored an alternative circuit layout based on a multi-mode interferometer (MMI). We demonstrated that:

- Our novel configuration based on a 2×2 MMI works effectively and allows to extract the Raman scattered light out of a SERS sensor while enabling further integration with an analysing photonic circuit i.e. lasers, filters and spectrometers, which are required for a complete Raman-on chip spectrometer. The penalty in terms of extra photon background generation compared to a common SERS measurement on the plasmonic slot is minimal.

- The remaining photon background originates from the Raman-like back-scattered light in the access waveguide (& Raman sensor) and from the output waveguide. The photon flux generated in the output waveguide is due to the same process occurring in the forward direction as the result of the presence of a remaining excitation beam which is reflected at the (plasmonic slot) interfaces. We found no contribution of the MMI on the photon background which is to be expected since most of the background photons generated in the MMI will be radiated out rather than being collected in the output waveguide.

- For typical output waveguide lengths (mm to cm), our MMI-based configuration performs significantly better than an alternative configuration relying on

forward-Raman scattering and a directional coupler-configuration which are both able to perform the same functionality.

From this work, several future directions can be followed:

- **Improve the presented MMI-plasmonic slot architecture (1)**. The first easy win to further reduce the photon background generation in our architecture comes by reducing the access waveguide length to the length of the mode converter (which was not yet optimized in our experimental demonstration). In that optimized regime, the true limit lies in reducing the reflections of the excitation beam which ultimately leads to the background contribution of the output waveguide. From numerical simulations we know that the reflection at the plasmonic slot interface will be the dominant reflection source and it will therefore be key to suppress it. This reflection could be better controlled by a more precise deposition technique such as the atomic layer deposition of gold that furthermore results in an increased Raman response [154].

- **Improve the presented MMI-plasmonic slot architecture (2)**. Another way to improve the architecture is to engineer the MMI so that it acts as a wavelength division multiplexer [155, 156], transmitting the excitation beam from the input port to the Raman sensor with a high transmission and transmitting it inefficiently from the Raman sensor to the output port while the scattered Raman light is efficiently collected in the output port. This would lead to a reduction of the reflected pump light in the output waveguide therefore improving the SBR for even longer/more complex analyzing circuits.

- **Towards a fully-integrated Raman spectroscopic system.** An interesting route to follow will be to proceed to the integration of the MMI-plasmonic slot configuration with other components of the Raman spectroscopic system that are readily available on the Si_3N_4-platform (i.e. lasers [67], detectors [143], spectrometers [66, 181] and pump rejection filters [17]). Since we know that the MMI-plasmonic slot configuration (combined with a pump rejection filter) allows for the integration of the analyzing circuit with only a limited background penalty let us now focus on other important components such as the laser, detector and Raman sensor. In this work we showed that SERS (SESRS) on the nanoplasmonic slot has a conversion ratio ($\zeta = P_{Raman}/P_{pump}$) of $\zeta_{SERS} = 10^{-9}$ ($\zeta_{SESRS} = 10^{-6}$). Furthermore, integrated lasers sources in the visible can reach power levels in the order of ~ 0.15 mW [unpublished result] (with $\sim$ 1mW power levels being the practical power limit for the nanoplasmonic slot in order to prevent thermal damage to the probed analyte and the sensor itself). Let us now further assume that future improvements

on the laser (or using a off-chip laser source) allow for 1 mW power levels, the generated Raman signals in that case corresponds to $P_{SERS} = 10^{-12}$W ($P_{SESRS} = 10^{-9}$W). Note that the signal power for SERS correspond to the integrated power over the whole signal peak. Therefore, in practice the power incident on a single pixel will be $P_{SERS,pix} \sim 10^{-13} - 10^{-14}$W. Commercially available uncooled detectors show detection sensitivities in the order of a few (tens of) $\sim$ fW (for a 1s integration time) [182–184] and can be integrated using hybrid or heterogeneous integration techniques. Cooled detectors are to be avoided since they require volume, add extra cost, have additional power consumption and require challenging thermal management on-chip. In practical applications often low concentration or weak-scattering molecules are to be detected (unlike the NTP monolayer considered in the calculation here). In that scenario it seems further advances are required for both the laser, detector and the Raman sensor. Further progress on the Raman sensor itself might still be possible: experimentally demonstrating SE-SRS (SE-CARS), using ALD plasmonic slots and integrated pulsed lasers at the visible can further aid to sufficiently increase the Raman response to signal levels that can be picked up by the (uncooled) integrated photodetectors. Note that for applications such as in-vitro rapid diagnostics with use-once chips the full-integration of the Raman spectroscopic system is not required where the laser and detector are off-chip and not disposed after a single use. On one hand on-chip Raman boosts the performance relative to free-space excitation and collection and on the other the analyte needs to be in contact with the waveguide and therefore can affect the sensitivity of the sensor after-use (jeopardizing the next measurement). In that case, integrating the laser and detector on-chip will make sense for analytes that will not have much effect on the sensor (e.g. gas) or if the effect of the analyte on the sensor can be mitigated. Short-term practical implementations can make use of fiber-packed NWERS (SERS) sensors linked to off-chip components that are difficult/expensive to integrate [146].

- **Practical applications of on-chip Raman spectroscopy** Once a (partially) integrated Raman spectroscopic system is demonstrated it will be interesting to look towards potential practical applications of such an integrated Raman sensor. Traditional (confocal) Raman microscopic systems have found applications in many fields (for non-destructive material analysis) e.g. to determine the distribution of active pharmaceutical ingredients [185], to study meteorites [186], to characterize graphene in device fabrication [187], to determine pigment concentrations in paint [188] and to detect biomolecules in living cells [189]. However, these desktop devices are costly (a few hundred thousand euros) and bulky which limits their widespread use, especially in outside-of-the-lab environments [62]. Therefore, there is an increased

demand for miniaturized Raman devices that can be used in the field [11]. Portable handheld devices have recently emerged as a rapid in-situ tool that can be used for applications as pharmaceutical material screening [190], identifying different types of bacterial growth [191] and quality control in the food industry [192]. These devices are however still rather costly (tens of thousands of euros) and the miniaturization of the Raman spectroscopic system results in a reduced performance compared to their desktop equivalents [193, 194]. Integrated Raman sensors boost the performance relative to free-space excitation and collection [14, 76] (allowing for the detection of low-concentration analytes in complex samples e.g. blood). Further, they are also cheap, compact, reliable and mass-producible (allowing for widespread use on-site). They therefore can be used for i.a. point-of-care applications allowing for high-performance sensing outside of the medical laboratory (e.g. bedside (drug) monitoring in/of patients blood [81], in-line measurements during endoscopy), drug development (e.g. detecting protease activity [84, 195]), thread detection (e.g. in-situ detection of trace amounts of chemical warfare agents [196]) and environmental-quality monitoring (e.g. indoor monitoring of volatile organic compounds (VOCs) [78]). A partial attempt has been made towards a cheap, high-performing handheld Raman spectroscopic device in the form of a handheld spectrometer (the size of a smartphone). This device uses photonic waveguide interferometers that are monolitically integrated on top of a CMOS image sensor showing high optical throughput and spectral resolution (like the desktop versions) [193, 197]. However, it only involves the integration of the spectrometer itself and does not take advantage of the enhanced signal collecting capacity of waveguide-enhanced Raman spectroscopy.

In the second part of this PhD project, we explored on-chip surface-enhanced coherent Raman spectroscopy. We showed that:

- On-chip surface-enhanced coherent Raman spectroscopy provides a (theoretical) signal enhancement of more than three orders of magnitude compared to SERS on the nanoplasmonic waveguide (allowing for room temperature detection), while proving an improvement in the signal-to-background ratio (SBR) of more than three orders of magnitude compared to on-chip SRS on a dielectric strip waveguide (thereby drastically reducing the background contribution of the waveguide core).
- Both on-chip SE-SRS and SE-CARS provide a comparable Raman response. SE-SRS is preferred over SE-CARS in this work with the main argument that it has a linear dependence on the analyte concentration (compared to quadratic for SE-CARS) and will therefore be more sensitive to relevant low concentrations analytes such as biomolecules in aqueous solution.
- The experimental demonstration of SE-SRS on the nanoplasmonic slot is challenging. The heating intrinsic to linear absorption in the gold nanostructure induces a thermo-optic effect that affects the Stokes beam overshadowing the Raman signature of the analyte (NTP in the present case). Not only is this thermo-optic effect two order of magnitude stronger than the Raman response, but it also presents its own non-trivial spectrum overlapping with that of the thermo-extinction coefficient presented in literature.

From this work the following future directions can be followed:

- **Thermal background mitigation.** The obvious next step for SE-SRS is to mitigate the thermal background encountered in practical SE-SRS measurements. Firstly, the modulation frequency of the lock-in system can be increased to 100‘s of MHz such that it will lead to a significant drop in strength of the parasitic thermal background. Secondly, we can increase the thermal capacitance of the gold sheets next to plasmonic slot such that it lowers the thermal response of the plasmonic slot at higher modulation frequencies. Thirdly, we can try to remove the thermal modulation (and hence the thermal background) all together by adding a second pump laser to our detection scheme that is in anti-phase to the first pump laser so that the heating in the plasmonic slot is constant over time. (A detailed discussion on the different mitigation strategies can be found in Section 6.5).
- **Performing SE-CARS over SE-SRS.** In this work, we chose SE-SRS over SE-CARS without knowing about the presence of the photothermal background. Since SE-CARS does not suffer from such a background it would

be tempting to experimentally explore SE-CARS instead of focusing on the thermal background mitigation in SE-SRS. However, our original argument still stands in that SE-SRS will be much more sensitive to low concentration analytes such as biomolecules in aqueous solution. Furthermore, the signal enhancement for SE-CARS will only be limited (for low concentrations) compared to on-chip SERS such that it will be difficult to justify its additional setup complexity. Therefore, as a first step, further investigating on-chip SE-SRS would be worth the effort since it provides a better performance compared to SE-CARS. The proposed mitigation strategies for SE-SRS are relatively straightforward and it would therefore only take a reasonable amount of time to verify if they are successful or not. In case the thermal background is persistent or another parasitic effect is present, the shift to SE-CARS can be made. SE-CARS has a more straightforward detection scheme but also carries some unknown risk since it has not yet been demonstrated on a waveguide platform. When performing SE-CARS on the plasmonic slot, special care should be taken to avoid thermal damage to the sample and the analyte as the first preliminary results on the plasmonic slot show that this might be an issue. At this moment, the risk of (further) investigation SE-SRS and SE-CARS seems more or less equal, while the return in performance for SE-SRS is higher.

- **Determining the detection limit of on-chip SE-SRS.** Once the thermal background is (sufficiently) suppressed, the next step will be to perform a successful measurement of SE-SRS on the plasmonic slot. Firstly, one performs a proof-of-concept measurement on the nanoplasmonic slot using an analyte with a strong Raman response such as a monolayer of NTP [16]. Afterwards, the experimental detection limit of our SE-SRS sensor can be determined by probing the Raman response of small fractions of DMSO dissolved into a solvent with a low surface tension like ethanol or IPA (in order to ensure the wetting of the plasmonic slot [103]).

- **Integration of the SE-SRS sensor.** Further along the road, the SE-SRS sensor can be integrated with the rest of the Raman spectroscopic system to form an all-on-chip SE-SRS system. In our SE-SRS measurements we use CW lasers [67, 68], (balanced) detectors [143, 144] and lock-in amplifiers [145, 198] which all have been demonstrated on integrated platforms with decent performance. Furthermore, by integrating an array of lasers onto the chip (with each laser dedicated to a specific Raman mode) we can probe the Raman spectrum without the need for tuning of the laser wavelength.

A

Appendix A: Simulating reflections in the MMI-plasmonic slot configuration

In our MMI-plasmonic slot configuration, unwanted back-reflections (at the interface between the dielectric and plasmonic slot, the interface between the MMI-taper and the mode converter, the interface between the mode converter and the slot waveguide, in the MMI from the input port to the output port, the interface between the silica cladded part of the dielectric slot waveguide and the water cladded part of that same waveguide) can reflect a small portion of the excitation beam into the output waveguide. In order to get an idea about the reflections at the different interfaces we make use of numerical simulations (Lumerical's EME and FDTD solver). In the models, a refractive index of 2.02 is taken for the PECVD Si_3N_4, 1.76 for Al_2O_3, 1.45 for SiO_2, 1 for air, 1.33 for water and the build in refractive index model for gold (Johnson and Christy [199]). The different models follow the device dimensions as described in Section 5.1.1 and the reflections are evaluated at 785 nm which corresponds to the pump wavelength used in our experiments.

Fig. A.1 shows the FDTD simulation model (width $\times$ length = 1.5 μm $\times$ 2 μm) build for the interface between the dielectric and plasmonic slot (image taken directly from the simulation). A fundamental TE-mode of power P_{in} is injected in the dielectric slot waveguide, where part of it reflects ($P_{refl.}$) at the interface and is recorder by the power monitor placed behind the excitation source. A mesh resolution of 1 nm is used in the simulation. For practical plasmonic slots it was

discovered (using SEM) that the gold lift-off at the edge of the plasmonic slot is not completely perfect resulting in a reduced/increased gold gap. Therefore, in our simulations the gold thickness in the gap is varied from 1 nm to 10 nm. This results in a recorder reflection of -25 dB to -30 dB, where the reflection is calculated as:

$$\text{Reflection [dB]} = 10 \times \log(\frac{P_{\text{refl.}}}{P_{\text{in}}}) \tag{A.1}$$

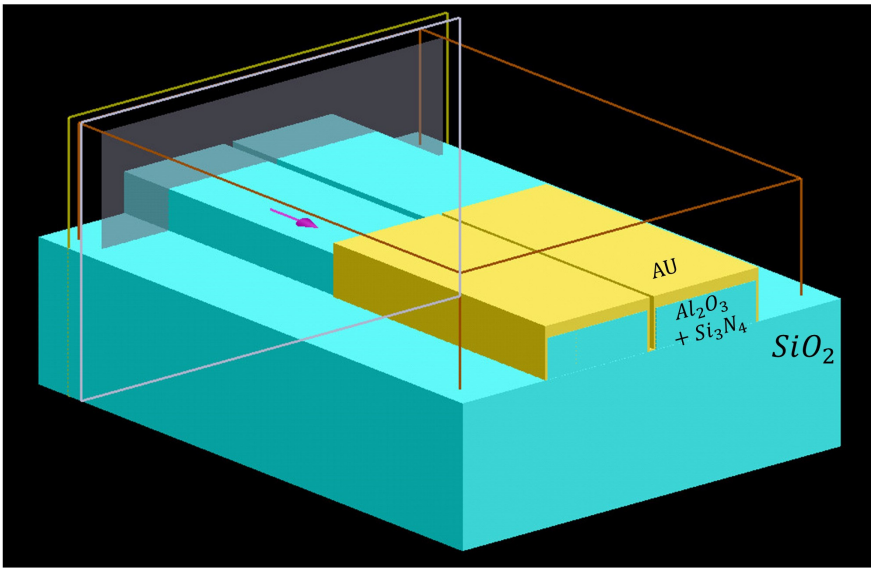


Figure A.1: 3D view of the plasmonic slot connected to the Al_2O_3 *coated* Si_3N_4 *slot waveguide. The FDTD simulation environment is defined by the red box, while the reflection power monitor (yellow rectangle) is placed behind the excitation source (purple arrow).*

In Fig.A.2, the build model is shown for the interface between the silica cladded part of the dielectric slot waveguide and the water cladded part of that same waveguide. Where we again define a width $\times$ length = 1.5 µm $\times$ 2 µm simulation environment and follow an analogous procedure as for the plasmonic slot interface. A reflection of -36 dB is found for the water cladded dielectric slot.

Next, we consider the interface between the MMI-taper and the mode converter and the interface between the mode converter and the slot waveguide. Fig. A.3 displays the build model for both (following the dimensions as described in Section 5.1.1). In the EME solver, the fundamental TE-mode is excited in port 1, transmitted to port 2 and the reflection coefficient (s_{11}^2) is extracted from the calculated s-matrix in the model. A total of 30 modes are calculated over the taper length, 60 modes

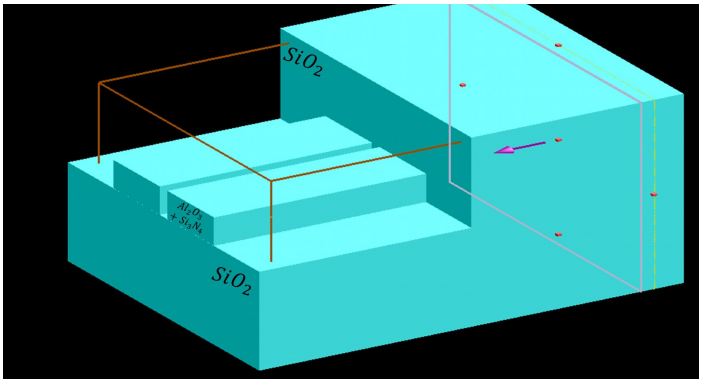


Figure A.2: 3D view of the transition between the silica cladded dielectric slot waveguide to the Al_2O_3 coated Si_3N_4 slot waveguide with a water cladding (not shown). The FDTD simulation environment is defined by the red box (not shown in silica), while the reflection power monitor (yellow rectangle) is placed behind the excitation source (purple arrow).

over the strip-to-slot converter and a 10 nm mesh resolution is used. The calculated reflection for the taper corresponds to -120 dB, while it is -300 dB for the mode converter. Finally, we simulate the amount of pump power reflected in the MMI towards the output waveguide (port 2) when injecting a fundamental TE-mode at the input waveguide (port 1). Fig. A.4 shows the model build in Lumerical using the EME solver. A total of 50 modes are calculated in the MMI cell along the propagation direction and a mesh resolution of 30 nm is used. A reflection of - 43.5 dB is calculated for the MMI towards the output waveguide (port 2) when injecting a pump laser in the input waveguide (port 1).

Numerical Reflections	**dB**
R_{Plas}	-25 to -30
R_{Clad}	-36
R_{MMI}	-43.5
R_{Taper}	-120
R_{Mode}	-300

Table A.1: Simulated reflection values for the interface between the dielectric and plasmonic slot (Plas), the silica cladded part of the dielectric slot waveguide and the water cladded part (Clad), the MMI-taper and the mode converter (Taper), the mode converter and the slot waveguide (Mode) and the internal reflection in the MMI from the input port to the output port (MMI).

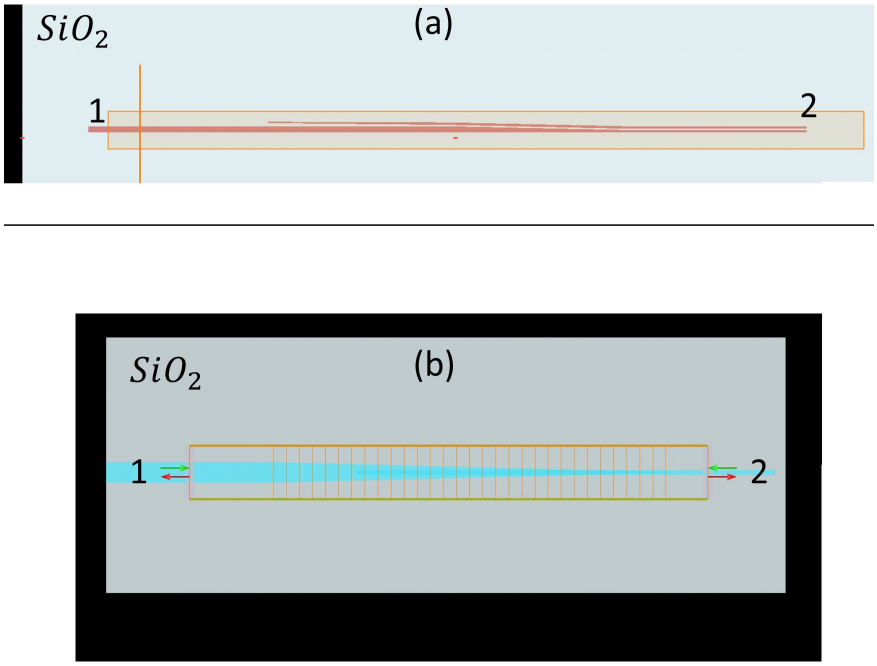


Figure A.3: Top-down view of (a) the mode converter and (b) MMI-taper build in the simulation model. The silica top cladding is not shown in the figures. For the MMI-taper the different cells are shown for which the EME-solver calculates the fundamental mode at each cell.

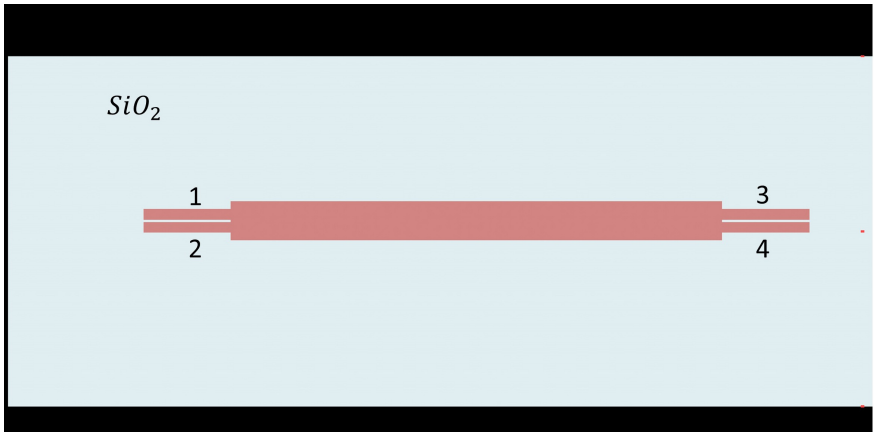


Figure A.4: Top-down view of the MMI simulation model. The silica top cladding is not shown in the figure. The fundamental TE-mode is injected in port 1 and the amount of reflection at port 2 is extracted from the s-matrix.

B

Appendix B: Additional illustrations of the 3D thermal simulation

In Section 6.4 we give a detailed description of the custom-build 3D simulation model of the nanoplasmonic slot (using COMSOL Multiphysics). This model has been used to make an estimate on the thermal modulation (ΔT_{Sim}) occurring in the plasmonic slot when a (modulated) pump and Stokes beam are propagating through it. A part of these pump and Stokes beams is absorbed and acts as a heat source for the nanoplasmonic slot. Here we present three supportive figures on the solved model taken directly from our COMSOL simulation. The three figures below show the (absolute) temperature in the simulation model (from three different perspective) after 10 μs when a quasi-steady-state is reached in the plasmonic slot (see Fig. 6.10). From Fig. B.1 we can see that the temperature is the highest in and close to the plasmonic slot and that the gold sheets next to the waveguide have a higher temperature than alumina and silica layers which is expected due to the lower thermal resistance of gold compared to the surrounding materials (Section 6.4.4 provides a more detailed discussion on this). Fig. B.2 & B.3 show the complete and a top-down view of the simulation model respectively where there is only a mild temperature increase in the silicon and silica layer since they act as heat sinks. Furthermore, these figures indicate the infinite element domains (these regions apply a semi-infinite coordinate stretch along the axis of interest such that they approximate the different material layers as infinitely large) that terminate the

simulation model on each side with a constant temperature of 293.15 K at the end.

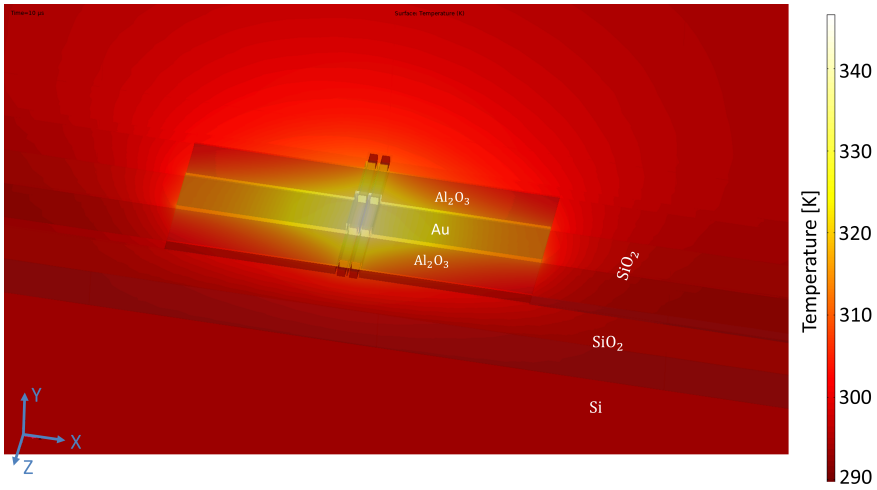


Figure B.1: 3D view of the temperature distribution in the plasmonic slot and its surrounding (from COMSOL). The different material layers are indicated in the figure. The heat source originates from the absorption of the (modulated) pump and Stokes laser propagating through the nanoplasmonic slot.

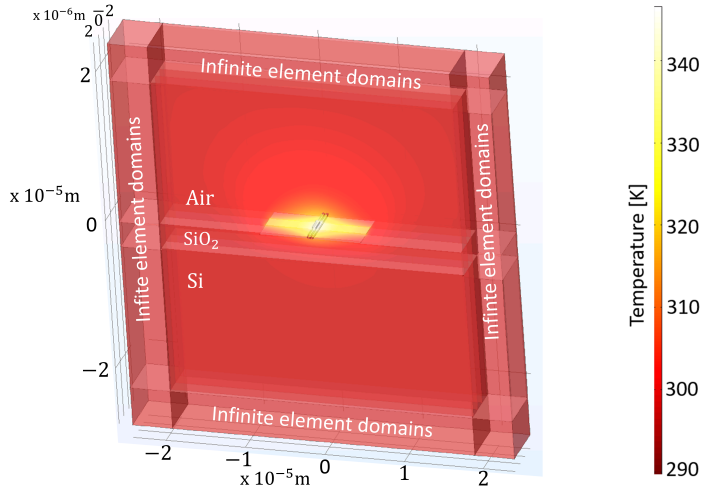


Figure B.2: Full view of the complete simulation environment, the main material layers and the infinite element domains set in the model. The heat source originates from the absorption of the (modulated) pump and Stokes laser propagating through the nanoplasmonic slot.

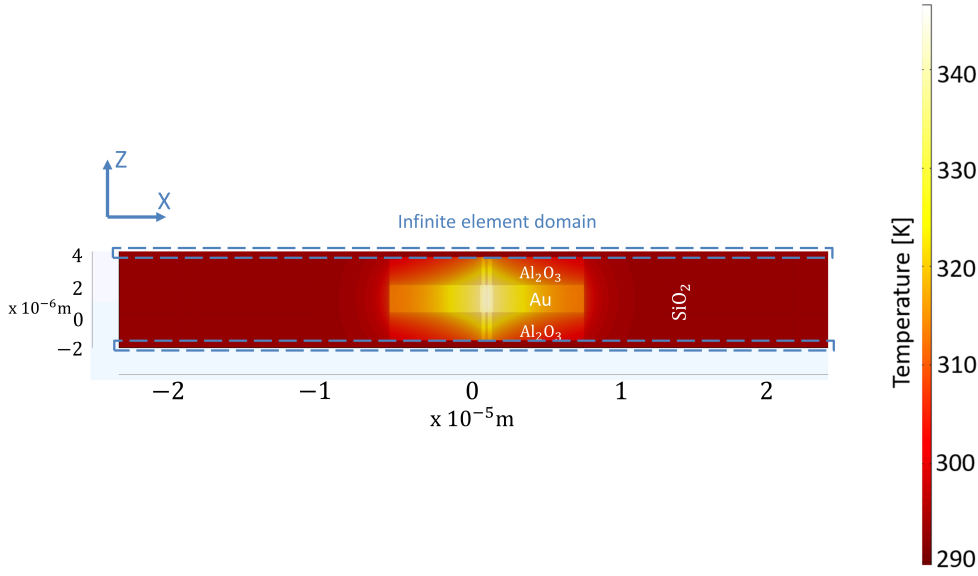


Figure B.3: Top-down view of the simulation model with the two infinite element domains that terminate the model in the Z-direction (blue dotted lines). The heat source originates from the absorption of the (modulated) pump and Stokes laser propagating through the nanoplasmonic slot.

C

Appendix C: Chip-scale investigation on the acoustic vibrations of dielectric Nanoparticles

C.1 Introduction

In the first two years of my PhD the focus was on performing a CARS-like (Coherent anti-Stokes Raman scattering) measurement on 22-nm polystyrene beads. These beads were dropcasted in aqueous solution on top of a silicon nitride slot waveguide (see Fig. C.1). In literature it is shown [200] that these polystyrene beads have an observable acoustic vibrational spectrum (GHz vibrations vs THz for molecular vibrations) using a similar four-wave mixing experiment in a free-space configuration (see Fig. C.2). By using photonic integrated circuits (PICs) the interaction length with the suspension of nanoparticles is further increased while the propagating light beam remains tightly confined. Furthermore, PICs allow for a much superior integration with the other spectroscopic components on a single chip. This demonstration on the polystyrene beads would serve as a proof-of-concept such that in a later phase the acoustic vibrations of viral particles (organic nanoparticles) could be probed. These viral particles possess a rigid outer shell (the capsid) which shows a distinct acoustic spectrum depending on their size,

shape, and stiffness [201–204] (just like plastic and metal nanoparticles [205, 206]) which can be used as a fingerprint to distinguish them from each other. In this way, the sensing capabilities of the waveguide-based Raman sensor can be expanded from the detection of molecules to the detection of (viral) particles paving the way for fast, cheap and mass-producible virus sensors.

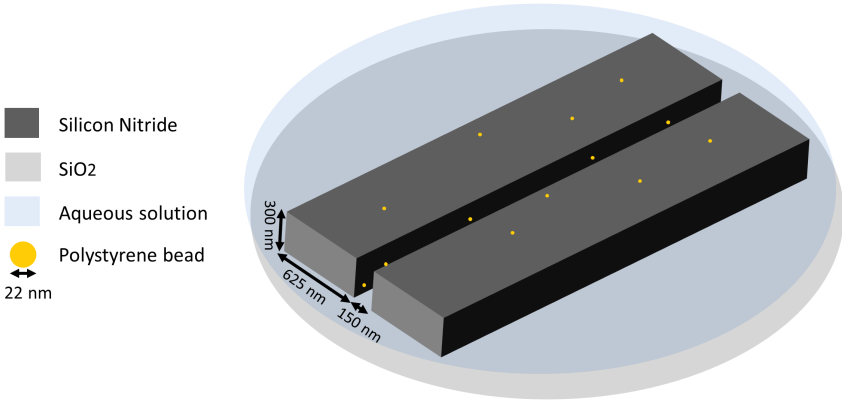


Figure C.1: Illustration of the nanoparticle solution dropcasted on top of a dielectric slot waveguide.

C.2 Methods

C.2.1 Fabrication

The photonic circuits used for experiments are fabricated on 200 mm silicon wafers. 300 nm Si_3N_4 is deposited through a low temperature plasma enhanced chemical vapor deposition process [74] on a 3.3 μm SiO_2 layer (imec's BioPIX platform [149]). Subsequent patterning is performed by deep UV-lithography and reactive ion etching. In this way two 5.5 mm long (l) strip waveguide with width (w) $\times$ height (h) = 1400 $\times$ 300 nm^2 are defined on both sides, that taper over 50 µm into the slot waveguide with width (w) $\times$ height (h) = 1400 $\times$ 300 nm^2 and a 150 nm wide slot. Five separate slot waveguides are defined with a varying length going from 0.4 mm to 6.7 mm. A silica top-cladding of 1 µm is deposited over the full waveguide expect over the length of the dielectric slot where the top-cladding is opened up. A well cut from a standard 96 wells-plate is glued to the chip using UV-curable glue. Having a long term confinement of the nanobead solution on the chip is essential since it prevents evaporation of the solution and assures the constant concentration of the particles during experiments. A small hole is made in

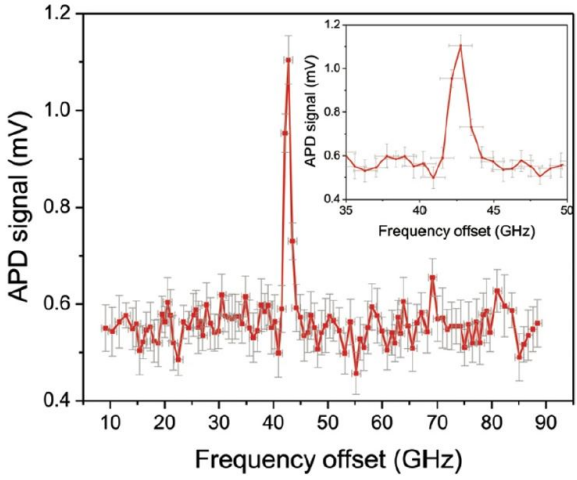


Figure C.2: Resonance observed (43 GHz) when probing the 22-nm polystyrene beads using a four-wave mixing experiment in a free-space configuration (Image taken directly from [200]). APD: Avalanche Photodetector.

the wells such that a droplet of the nanobead solution can be dropcasted on top of our waveguide structures. Afterwards, the well is closed using scotch tape.

C.2.2 Setup

Measurements are performed on the setup depicted in Fig. C.3. A pump laser (Santec TSL510, 1550.3 nm) is first amplified (EDFA booster) before being guided towards a 50/50-splitter where it is combined with the tunable Stokes laser (Santec TSL510). A Fiber Bragg grating (FBG) before chip filters out the (laser) noise in the spectral range where the anti-Stokes beam is expected to be generated (1549.7-1550.1 nm). Afterwards, the noise floor in that region corresponds to -87 dBm. Both the pump and Stokes beam are coupled onto the chip using vertical fiber coupling. A semi-automatic setup is employed to ensure the stable power input onto the chip (reducing the effect of fiber drifts) which is critical because CARS has a non-linear scaling with pump power. After the chip another FBG is employed to filter out both the pump and Stokes contribution before finally the spectra is detected by an OSA (Advantest 8381A). The exact beating frequency of the pump and Stokes laser was determined and its variations are in the MHz range while the linewidth of the vibration that is probed is in the $\sim$ 2 GHZ range such that this variation will not limit the amount of power effectively used in exciting the vibrations. A spectral range of 25 GHz-75 GHz is probed during the experiment

with a 1 pm (124 MHz) resolution, while having a pump and Stokes power of 24.3 dBm & 6.31 dBm respectively before chip. The correct function of the setup was verified by performing a degenerate four-wave mixing measurement on a 4 cm-long strip waveguide showing the generation of a similar peak as the one we expect in our CARS experiment.

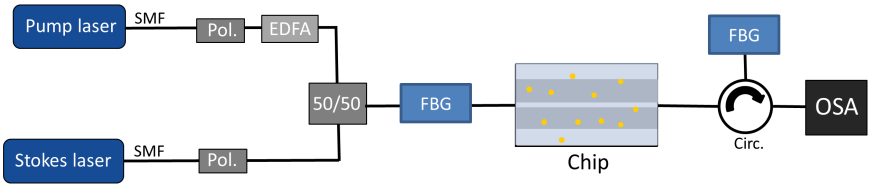


Figure C.3: Schematic of the setup used to probe the acoustic vibrations of the polysytrene beads. SMF: Single mode fiber, Pol: Polarization controller, EDFA: Erbium-Doped Fiber Amplifier, 50/50-splitter, FBG: Fiber Bragg grating, OSA: Optical spectrum analyzer.

C.2.3 Estimated signal

Since we know the nonlinear refractive index n_2 of the polystyrene beads in aqueous solution from [200], the generated anti-Stokes (idler) signal at the end of the waveguide for these beads using our dielectric slot and measurement setup can be estimated by filling in [38]:

$$P_{as}(L_{slot}) = P_s(0)(1 + \kappa^2/4g^2)\sinh^2(gL_{slot}) \tag{C.1}$$

Assuming no pump depleting and where:

$$\kappa = \Delta k + 2\gamma P_p \tag{C.2}$$

$$\gamma = \frac{n_2\omega}{cA_{eff}} \tag{C.3}$$

$$\Delta k = (n_s\omega_s + n_{as}\omega_{as} - 2n_p\omega_p)/c \tag{C.4}$$

$$g = \sqrt{(2\gamma P_p)^2 - (\kappa/2)^2} \tag{C.5}$$

Now using P_p = 24.3 dbm$-T_{coupl.}-T_{slot}$ = 12.14 dBm, P_s = 6.31 dbm$-T_{coupl.}-T_{slot}$ = –5.85 dBm, where $T_{coupl.}$ = 7.76 dB is the coupling loss from free-space to the waveguide and T_{slot} = 4.4 dB/mm $\times$ L_{slot} is the scattering/absorption loss in the dielectric slot with the particle solutions as the top cladding. L_{slot} = 1 mm, n2 = 2×10^{-15} m^2/W, A_{eff} = 0.695 μm^2 and n_s = 1.80157, n_p = 1.80174, n_{as} = 1.80191 (using Lumerical mode solutions) with Δk = 12.46 and $2\gamma P_p$ = 381.

The generated signal at the end of the dielectric slot corresponds to $P_{as}(L_{slot})$ = –14.05 dBm. Now including the coupling loss ($T_{coupl.}$) and setup loss after chip (-12 dB) the power to be detected by the OSA corresponds to $P_{as,OSA}$ = –33.8 dBm while the noise floor is at -87 dBm.

C.3 Results

The acoustic resonances that were observed with a free-space approach [200] (see Fig. C.2) were not observed in the acoustic spectra of our on-chip based measurement. Our theoretical calculation involving the effective nonlinearity of the nanoparticles and the wave propagation do confirm the CARS conversion efficiency should be in the range of -33.8 dBm while we observe no signal down to -87 dBm.

This calculation did however exclude a non-ideal behavior of the beads (i.e. damped vibrations, absence of particles in the slot). Therefore it is possible that a problem could be found in this direction. Several hypothesis were refuted during the course of two years as the cause for the absence of a measurable signal. These included:

- Laser heating effects pushing the particles out of the slot.
  (I verified that this effect is not relevant by performing transmission experiments as a function of power)
- Trapping of particles by the optical forces
  (Calculations show that this force is not relevant for our 22 nm beads [207])
- Geometry and size dispersion of the nanobeads
  (SEM images show that the beads are monodisperse and of expected size)
- Particles stick to waveguide walls
  (This only leads to limited damping due to acoustic impedance mismatch between the particle and the waveguide wall [208, 209])
- No-wetting of the waveguide slot
  (Raman Glucose experiments showed wetting of the slot [210])

- Issue with the polystyrene beads
  (Alternative nanobeads (silica, CdSe/Cds) also did not show observable resonances)
- Exclusion of the beads from the waveguide slot by entropic forces
  (Measurements on strip waveguides and on the dried-up particle solution also did not show any resonances)

Despite these good efforts, experimental results did not follow. Therefore it was concluded that it was no longer fruitful to continue this line of experimental work. The possibility still remains that there is a unexpected particle behaviour that is not accounted for in our experiment/theory models which explains the absence of observable resonances. We contacted Professor R. Gordon [200] and his team in order to learn from their experience in probing the vibrations of these polystyrene nanoparticles with a free-space approach (possibly through a visit). Unfortunately, we were told that their setup was dismantled and that they did not yet succeeded in reproducing the experimental results. In the future it might be best to first recreate the results of [200] with a free-space approach before advancing towards photonic integrated circuits.

References

[1] R. S. Das and Y. Agrawal, "Raman spectroscopy: recent advancements, techniques and applications," *Vibrational spectroscopy*, vol. 57, no. 2, pp. 163–176, 2011.

[2] C. Krafft, G. Steiner, C. Beleites, and R. Salzer, "Disease recognition by infrared and raman spectroscopy," *Journal of biophotonics*, vol. 2, no. 1-2, pp. 13–28, 2009.

[3] M.-L. Xu, Y. Gao, X. X. Han, and B. Zhao, "Detection of pesticide residues in food using surface-enhanced raman spectroscopy: A review," *Journal of agricultural and food chemistry*, vol. 65, no. 32, pp. 6719–6726, 2017.

[4] P. Ashok, G. Singh, K. Tan, and K. Dholakia, "Fiber probe based microfluidic raman spectroscopy," *Optics express*, vol. 18, no. 8, pp. 7642–7649, 2010.

[5] K. Kong, C. Kendall, N. Stone, and I. Notingher, "Raman spectroscopy for medical diagnostics—from in-vitro biofluid assays to in-vivo cancer detection," *Advanced drug delivery reviews*, vol. 89, pp. 121–134, 2015.

[6] A. D. Tormo, D. Khalenkow, A. G. Skirtach, and N. Le Thomas, "Raman and quantitative-phase microscope with counter-propagating beams demonstrated on hela cells," *OSA Continuum*, vol. 2, no. 3, pp. 797–804, 2019.

[7] P. C. Ashok, B. B. Praveen, and K. Dholakia, "Near infrared spectroscopic analysis of single malt scotch whisky on an optofluidic chip," *Optics Express*, vol. 19, no. 23, pp. 22982–22992, 2011.

[8] B. Bao, J. Riordon, F. Mostowfi, and D. Sinton, "Microfluidic and nanofluidic phase behaviour characterization for industrial co 2, oil and gas," *Lab on a Chip*, vol. 17, no. 16, pp. 2740–2759, 2017.

[9] M. Nieuwoudt, S. Holroyd, C. McGoverin, M. Simpson, and D. Williams, "Raman spectroscopy as an effective screening method for detecting adulteration of milk with small nitrogen-rich molecules and sucrose," *Journal of dairy science*, vol. 99, no. 4, pp. 2520–2536, 2016.

[10] P. Graves and D. Gardiner, "Practical raman spectroscopy," *Springer*, 1989.

[11] B. H. Ray and K. T. Carron, "From portable raman to mobile raman: the progression of raman spectroscopy," in *Next-Generation Spectroscopic Technologies XI*, vol. 10657, p. 1065704, International Society for Optics and Photonics, 2018.

[12] R. Baets, A. Z. Subramanian, S. Clemmen, B. Kuyken, P. Bienstman, N. Le Thomas, G. Roelkens, D. Van Thourhout, P. Helin, and S. Severi, "Silicon photonics: silicon nitride versus silicon-on-insulator," in *Optical Fiber Communication Conference*, pp. Th3J–1, Optical Society of America, 2016.

[13] A. Dhakal, A. Z. Subramanian, P. Wuytens, F. Peyskens, N. Le Thomas, and R. Baets, "Evanescent excitation and collection of spontaneous raman spectra using silicon nitride nanophotonic waveguides," *Optics letters*, vol. 39, no. 13, pp. 4025–4028, 2014.

[14] A. Dhakal, P. C. Wuytens, F. Peyskens, K. Jans, N. L. Thomas, and R. Baets, "Nanophotonic waveguide enhanced raman spectroscopy of biological submonolayers," *Acs Photonics*, vol. 3, no. 11, pp. 2141–2149, 2016.

[15] N. L. Thomas, A. Dhakal, A. Raza, F. Peyskens, and R. Baets, "Impact of fundamental thermodynamic fluctuations on light propagating in photonic waveguides made of amorphous materials," *Optica*, vol. 5, pp. 328–336, Apr 2018.

[16] A. Raza, S. Clemmen, P. Wuytens, M. Muneeb, M. Van Daele, J. Dendooven, C. Detavernier, A. Skirtach, and R. Baets, "Ald assisted nanoplasmonic slot waveguide for on-chip enhanced raman spectroscopy," *APL Photonics*, vol. 3, no. 11, p. 116105, 2018.

[17] X. Nie, N. Turk, Y. Li, Z. Liu, and R. Baets, "High extinction ratio on-chip pump-rejection filter based on cascaded grating-assisted contra-directional couplers in silicon nitride rib waveguides," *Optics letters*, vol. 44, no. 9, pp. 2310–2313, 2019.

[18] D. Martens, A. Z. Subramanian, S. Pathak, M. Vanslembrouck, P. Bienstman, W. Bogaerts, and R. G. Baets, "Compact silicon nitride arrayed waveguide gratings for very near-infrared wavelengths," *IEEE Photonics Technology Letters*, vol. 27, no. 2, pp. 137–140, 2014.

[19] H. Zhao, S. Clemmen, A. Raza, and R. Baets, "Stimulated raman spectroscopy of analytes evanescently probed by a silicon nitride photonic integrated waveguide," *Optics letters*, vol. 43, no. 6, pp. 1403–1406, 2018.

[20] D. A. McQuarrie, *Quantum chemistry*. University Science Books, 2008.

[21] E. Smith and G. Dent, "Modern raman spectroscopy: a practical approach," 2005.

[22] R. Roque-Malherbe, *The physical chemistry of materials: energy and environmental applications*. CRC Press, 2016.

[23] J. R. Ferraro, *Introductory raman spectroscopy*. Elsevier, 2003.

[24] J.-X. Cheng and X. S. Xie, *Coherent Raman scattering microscopy*. CRC press, 2016.

[25] P. Wuytens, *Surface-enhanced Raman spectroscopy for intracellular sensing and protease activity detection: from chip technology to applications*. PhD thesis, Ghent University, 2017.

[26] "https://commons.wikimedia.org/wiki/file:morse-potential.png," Accessed: 2021-04-20.

[27] P. Vandenabeele, *Practical Raman spectroscopy: an introduction*. Wiley Online Library, 2013.

[28] M. Donahue, E. Botonjic-Sehic, D. Wells, and C. Brown, "Understanding infrared and raman spectra of pharmaceutical polymorphs," *American Pharmaceutical Review*, 2011.

[29] S. Dochow, C. Krafft, U. Neugebauer, T. Bocklitz, T. Henkel, G. Mayer, J. Albert, and J. Popp, "Tumour cell identification by means of raman spectroscopy in combination with optical traps and microfluidic environments," *Lab on a Chip*, vol. 11, no. 8, pp. 1484–1490, 2011.

[30] E. Le Ru and P. Etchegoin, *Principles of Surface-Enhanced Raman Spectroscopy: and related plasmonic effects*. Elsevier, 2008.

[31] J. R. Lakowicz, *Principles of fluorescence spectroscopy*. Springer science & business media, 2013.

[32] A. Fast and E. O. Potma, "Coherent raman scattering with plasmonic antennas," *Nanophotonics*, vol. 8, no. 6, pp. 991–1021, 2019.

[33] C. Zhang and J.-X. Cheng, "Perspective: Coherent raman scattering microscopy, the future is bright," *APL Photonics*, vol. 3, no. 9, p. 090901, 2018.

[34] R. W. Boyd, *Nonlinear optics*. Academic press, 2020.

[35] H. Zhao, *Coherent raman spectroscopy enabled by photonic integrated circuits*. PhD thesis, Ghent University, 2018.

[36] O. Aso, M. Tadakuma, and S. Namiki, "Four-wave mixing in optical fibers and its applications," *dEp*, vol. 1, p. 2, 1999.

[37] A. Zheltikov, "Coherent anti-stokes raman scattering: from proof-of-the-principle experiments to femtosecond cars and higher order wave-mixing generalizations," *Journal of Raman Spectroscopy*, vol. 31, no. 8-9, pp. 653–667, 2000.

[38] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*, pp. 195–211, Springer, 2000.

[39] G. I. Petrov, R. Arora, V. V. Yakovlev, X. Wang, A. V. Sokolov, and M. O. Scully, "Comparison of coherent and spontaneous raman microspectroscopies for noninvasive detection of single bacterial endospores," *Proceedings of the National Academy of Sciences*, vol. 104, no. 19, pp. 7776–7779, 2007.

[40] C. Valensise, A. Giuseppi, F. Vernuccio, A. De la Cadena, G. Cerullo, and D. Polli, "Removing non-resonant background from cars spectra via deep learning," *APL Photonics*, vol. 5, no. 6, p. 061305, 2020.

[41] E. Capitaine, N. O. Moussa, C. Louot, C. Lefort, D. Pagnoux, J.-R. Duclère, J. F. Kaneyasu, H. Kano, L. Duponchel, V. Couderc, *et al.*, "Coherent anti-stokes raman scattering under electric field stimulation," *Physical Review B*, vol. 94, no. 24, p. 245136, 2016.

[42] R. R. Jones, D. C. Hooper, L. Zhang, D. Wolverson, and V. K. Valev, "Raman techniques: Fundamentals and frontiers," *Nanoscale research letters*, vol. 14, no. 1, pp. 1–34, 2019.

[43] A. Virga, C. Ferrante, G. Batignani, D. De Fazio, A. Nunn, A. Ferrari, G. Cerullo, and T. Scopigno, "Coherent anti-stokes raman spectroscopy of single and multi-layer graphene," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.

[44] J. Pitarke, V. Silkin, E. Chulkov, and P. Echenique, "Theory of surface plasmons and surface-plasmon polaritons," *Reports on progress in physics*, vol. 70, no. 1, p. 1, 2006.

[45] A. Raza, *Raman spectroscopy enhanced by on-chip dielectric and metal waveguides*. PhD thesis, Ghent University, 2019.

[46] R. Dragila, B. Luther-Davies, and S. Vukovic, "High transparency of classically opaque metallic films," *Physical Review Letters*, vol. 55, no. 10, p. 1117, 1985.

[47] I. Pockrand, "Coupling of surface plasma oscillations in thin periodically corrugated silver films," *Optics Communications*, vol. 13, no. 3, pp. 311–313, 1975.

[48] J. D. Jackson, "Classical electrodynamics," 1999.

[49] P. Zijlstra, M. Orrit, and A. F. Koenderink, "Metal nanoparticles for microscopy and spectroscopy," in *Nanoparticles*, pp. 53–98, Springer, 2014.

[50] G. Sun, J. B. Khurgin, and A. Bratkovsky, "Coupled-mode theory of field enhancement in complex metal nanostructures," *Physical Review B*, vol. 84, no. 4, p. 045415, 2011.

[51] J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, "Biosensing with plasmonic nanosensors," *Nanoscience and Technology: A Collection of Reviews from Nature Journals*, pp. 308–319, 2010.

[52] E. Le Ru and P. Etchegoin, "Rigorous justification of the— e— 4 enhancement factor in surface enhanced raman spectroscopy," *Chemical Physics Letters*, vol. 423, no. 1-3, pp. 63–66, 2006.

[53] S. A. Maier, *Plasmonics: fundamentals and applications*. Springer Science & Business Media, 2007.

[54] M. Moskovits, "Surface-enhanced raman spectroscopy: a brief retrospective," *Journal of Raman Spectroscopy: An International Journal for Original Work in all Aspects of Raman Spectroscopy, Including Higher Order Processes, and also Brillouin and Rayleigh Scattering*, vol. 36, no. 6-7, pp. 485–496, 2005.

[55] T. J. Seok, A. Jamshidi, M. Eggleston, and M. C. Wu, "Mass-producible and efficient optical antennas with CMOS-fabricated nanometer-scale gap," *Optics express*, vol. 21, no. 14, pp. 16561–16569, 2013.

[56] Y. Zhang, Y.-R. Zhen, O. Neumann, J. K. Day, P. Nordlander, and N. J. Halas, "Coherent anti-stokes raman scattering with single-molecule sensitivity using a plasmonic fano resonance," *Nature communications*, vol. 5, no. 1, pp. 1–7, 2014.

[57] C. Steuwe, C. F. Kaminski, J. J. Baumberg, and S. Mahajan, "Surface enhanced coherent anti-stokes raman scattering on nanostructured gold surfaces," *Nano letters*, vol. 11, no. 12, pp. 5339–5343, 2011.

[58] D. V. Voronine, Z. Zhang, A. V. Sokolov, and M. O. Scully, "Surface-enhanced fast cars: en route to quantum nano-biophotonics," *Nanophotonics*, vol. 7, no. 3, pp. 523–548, 2018.

[59] R. R. Frontiera, A.-I. Henry, N. L. Gruenke, and R. P. Van Duyne, "Surface-enhanced femtosecond stimulated raman spectroscopy," *The Journal of Physical Chemistry Letters*, vol. 2, no. 10, pp. 1199–1203, 2011.

[60] C. L. Lee and K. Hewitt, "First demonstration of surface enhanced-stimulated raman spectroscopy (se-srs) using low-power cw sources," *Faraday Discussions*, vol. 205, pp. 227–232, 2017.

[61] C. Zong, R. Premasiri, H. Lin, Y. Huang, C. Zhang, C. Yang, B. Ren, L. D. Ziegler, and J.-X. Cheng, "Plasmon-enhanced stimulated raman scattering microscopy with single-molecule detection sensitivity," *Nature communications*, vol. 10, no. 1, pp. 1–11, 2019.

[62] A. Z. Subramanian, E. Ryckeboer, A. Dhakal, F. Peyskens, A. Malik, B. Kuyken, H. Zhao, S. Pathak, A. Ruocco, A. De Groote, *et al.*, "Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip," *Photonics Research*, vol. 3, no. 5, pp. B47–B59, 2015.

[63] D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Virot, J.-M. Fédéli, *et al.*, "Roadmap on silicon photonics," *Journal of Optics*, vol. 18, no. 7, p. 073003, 2016.

[64] A. Raza, S. Clemmen, P. Wuytens, M. De Goede, A. S. Tong, N. Le Thomas, C. Liu, J. Suntivich, A. G. Skirtach, S. M. Garcia-Blanco, *et al.*, "High index contrast photonic platforms for on-chip raman spectroscopy," *Optics express*, vol. 27, no. 16, pp. 23067–23079, 2019.

[65] X. Nie, E. Ryckeboer, G. Roelkens, and R. Baets, "CMOS-compatible broadband co-propagative stationary fourier transform spectrometer integrated on a silicon nitride photonics platform," *Optics express*, vol. 25, no. 8, pp. A409–A418, 2017.

[66] E. Ryckeboer, X. Nie, A. Z. Subramanian, D. Martens, P. Bienstman, S. Clemmen, S. Severi, R. Jansen, G. Roelkens, and R. Baets, "CMOS-compatible silicon nitride spectrometers for lab-on-a-chip spectral sensing," *Proc. SPIE*, vol. 9891, p. 98911K, 2016.

[67] S. Kumari, E. P. Haglund, J. S. Gustavsson, A. Larsson, G. Roelkens, and R. G. Baets, "Vertical-cavity silicon-integrated laser with in-plane waveguide emission at 850 nm," *Laser & Photonics Reviews*, vol. 12, no. 2, p. 1700206, 2018.

[68] E. Haglund, M. Jahed, J. S. Gustavsson, A. Larsson, J. Goyvaerts, R. Baets, G. Roelkens, M. Rensing, and P. O'Brien, "High-power single transverse and polarization mode vcsel for silicon photonics integration," *Optics express*, vol. 27, no. 13, pp. 18892–18899, 2019.

[69] A. Dhakal, F. Peyskens, S. Clemmen, A. Raza, P. Wuytens, H. Zhao, N. Le Thomas, and R. Baets, "Single mode waveguide platform for spontaneous and surface-enhanced on-chip raman spectroscopy," *Interface focus*, vol. 6, no. 4, p. 20160015, 2016.

[70] D. J. Griffiths and D. F. Schroeter, *Introduction to quantum mechanics*. Cambridge University Press, 2018.

[71] M. O. Scully and M. S. Zubairy, "Quantum optics," 1999.

[72] M. Eichhorn and M. Pollnau, "Spectroscopic foundations of lasers: Spontaneous emission into a resonator mode," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 21, no. 1, pp. 486–501, 2014.

[73] F. Peyskens, A. Dhakal, P. Van Dorpe, N. Le Thomas, and R. Baets, "Surface enhanced raman spectroscopy using a single mode nanophotonic-plasmonic platform," *ACS photonics*, vol. 3, no. 1, pp. 102–108, 2016.

[74] A. Z. Subramanian, P. Neutens, A. Dhakal, R. Jansen, T. Claes, X. Rottenberg, F. Peyskens, S. Selvaraja, P. Helin, B. Du Bois, K. Leyssens, S. Severi, P. Deshpande, R. Baets, and P. Van Dorpe, "Low-loss singlemode pecvd silicon nitride photonic wire waveguides for 532–900 nm wavelength window fabricated within a CMOS pilot line," *IEEE Photonics Journal*, vol. 5, no. 6, pp. 2202809–2202809, 2013.

[75] A. Dhakal, A. Raza, F. Peyskens, A. Z. Subramanian, S. Clemmen, N. Le Thomas, and R. Baets, "Efficiency of evanescent excitation and collection of spontaneous raman scattering near high index contrast channel waveguides," *Optics express*, vol. 23, no. 21, pp. 27391–27404, 2015.

[76] A. Dhakal, P. Wuytens, F. Peyskens, A. Subramanian, A. Skirtach, N. Le Thomas, and R. Baets, "Nanophotonic lab-on-a-chip raman sensors: a sensitivity comparison with confocal raman microscope," in *2015 International Conference on BioPhotonics (BioPhotonics)*, pp. 1–4, IEEE, 2015.

[77] C. M. Dissanayake, I. D. Rukhlenko, M. Premaratne, and G. P. Agrawal, "Raman-mediated nonlinear interactions in silicon waveguides: Copropagating and counterpropagating pulses," *IEEE Photonics Technology Letters*, vol. 21, no. 19, pp. 1372–1374, 2009.

[78] H. Zhao, B. Baumgartner, A. Raza, A. Skirtach, B. Lendl, and R. Baets, "Multiplex volatile organic compound raman sensing with nanophotonic slot waveguides functionalized with a mesoporous enrichment layer," *Optics Letters*, vol. 45, no. 2, pp. 447–450, 2020.

[79] Z. Liu, H. Zhao, B. Baumgartner, B. Lendl, A. Stassen, A. Skirtach, N. Le Thomas, and R. Baets, "Ultra-sensitive slot-waveguide-enhanced raman spectroscopy for aqueous solutions of non-polar compounds using a functionalized silicon nitride photonic integrated circuit," *Optics Letters*, vol. 46, no. 5, pp. 1153–1156, 2021.

[80] S. A. Holmstrom, T. H. Stievater, D. A. Kozak, M. W. Pruessner, N. Tyndall, W. S. Rabinovich, R. A. McGill, and J. B. Khurgin, "Trace gas raman spectroscopy using functionalized waveguides," *Optica*, vol. 3, no. 8, pp. 891–896, 2016.

[81] D. A. Coucheron, D. N. Wadduwage, G. S. Murugan, P. T. So, and B. S. Ahluwalia, "Chip-based resonance raman spectroscopy using tantalum pentoxide waveguides," *IEEE Photonics Technology Letters*, vol. 31, no. 14, pp. 1127–1130, 2019.

[82] Z. Wang, S. J. Pearce, Y.-C. Lin, M. N. Zervas, P. N. Bartlett, and J. S. Wilkinson, "Power budget analysis for waveguide-enhanced raman spectroscopy," *Applied spectroscopy*, vol. 70, no. 8, pp. 1384–1391, 2016.

[83] C. C. Evans, C. Liu, and J. Suntivich, "Tio2 nanophotonic sensors for efficient integrated evanescent raman spectroscopy," *Acs Photonics*, vol. 3, no. 9, pp. 1662–1669, 2016.

[84] N. Turk, *Waveguide-based surface-enhanced Raman spectroscopy for protease activity detection*. PhD thesis, Ghent University, 2020.

[85] N. Turk, *Waveguide-Based Surface-Enhanced Raman Spectroscopy for Protease Activity Detection*. PhD thesis, Ghent University, 2020.

[86] N. Le Thomas, A. Dhakal, A. Raza, F. Peyskens, and R. Baets, "Impact of fundamental thermodynamic fluctuations on light propagating in photonic waveguides made of amorphous materials," *Optica*, vol. 5, no. 4, pp. 328–336, 2018.

[87] Z. Liu, H. Zhao, A. Raza, N. Le Thomas, and R. Baets, "On the performance of tantalum pentoxide and silicon nitride slot waveguides for on-chip raman spectroscopy," in *21st European Conference on Integratred Optics (ECIO 2019)*, pp. 1–3, 2019.

[88] N. Turk, A. Raza, P. Wuytens, H. Demol, M. Van Daele, C. Detavernier, A. Skirtach, K. Gevaert, and R. Baets, "Comparison of free-space and waveguide-based sers platforms," *Nanomaterials*, vol. 9, no. 10, p. 1401, 2019.

[89] N. Le Thomas, Z. Liu, C. Lin, H. Zhao, and R. Baets, "Raman on-chip: Current status and future tracks," in *Integrated Optics: Devices, Materials, and Technologies XXV*, vol. 11689, p. 1168908, International Society for Optics and Photonics, 2021.

[90] K. Reynkens, S. Clemmen, H. Zhao, A. Raza, T. Vanackere, A. Stassen, M. Van Daele, J. Dendooven, and R. Baets, "Gold-induced photothermal background in on-chip surface enhanced stimulated raman spectroscopy," *Optics Letters*, vol. 46, no. 5, pp. 953–956, 2021.

[91] P. C. Wuytens, A. Z. Subramanian, W. H. De Vos, A. G. Skirtach, and R. Baets, "Gold nanodome-patterned microchips for intracellular surface-enhanced raman spectroscopy," *Analyst*, vol. 140, no. 24, pp. 8080–8087, 2015.

[92] A. X. Wang and X. Kong, "Review of recent progress of plasmonic materials and nano-structures for surface-enhanced raman scattering," *Materials*, vol. 8, no. 6, pp. 3024–3052, 2015.

[93] B.-S. Lee, D.-Z. Lin, and T.-J. Yen, "A low-cost, highly-stable surface enhanced raman scattering substrate by si nanowire arrays decorated with au nanoparticles and au backplate," *Scientific reports*, vol. 7, no. 1, pp. 1–7, 2017.

[94] F. Peyskens, P. Wuytens, A. Raza, P. Van Dorpe, and R. Baets, "Waveguide excitation and collection of surface-enhanced raman scattering from a single plasmonic antenna," *Nanophotonics*, vol. 7, no. 7, pp. 1299–1306, 2018.

[95] A. Raza, F. Peyskens, S. Clemmen, and R. Baets, "Towards single antenna on-chip surface enhanced raman spectroscopy: Arch dipole antenna," in *7th International Conference on Metamaterials, Photonic Cyrstals and Plasmonics (META'16)*, 2016.

[96] G. Singh, R. Bi, U. Dinish, and M. Olivo, "Generating localized plasmonic fields on an integrated photonic platform using tapered couplers for biosensing applications," *Scientific reports*, vol. 7, no. 1, pp. 1–8, 2017.

[97] F. Tang, P.-M. Adam, and S. Boutami, "Theoretical investigation of sers nanosensors based on hybrid waveguides made of metallic slots and dielectric strips," *Optics express*, vol. 24, no. 19, pp. 21244–21255, 2016.

[98] P. C. Wuytens, A. G. Skirtach, and R. Baets, "On-chip surface-enhanced raman spectroscopy using nanosphere-lithography patterned antennas on silicon nitride waveguides," *Optics express*, vol. 25, no. 11, pp. 12926–12934, 2017.

[99] L. Kong, C. Lee, C. M. Earhart, B. Cordovez, and J. W. Chan, "A nanotweezer system for evanescent wave excited surface enhanced raman spectroscopy (sers) of single nanoparticles," *Optics express*, vol. 23, no. 5, pp. 6793–6802, 2015.

[100] S. Lin, W. Zhu, Y. Jin, and K. B. Crozier, "Surface-enhanced raman scattering with ag nanoparticles optically trapped by a photonic crystal cavity," *Nano letters*, vol. 13, no. 2, pp. 559–563, 2013.

[101] Y. C. Jun, R. Kekatpure, J. White, and M. Brongersma, "Nonresonant enhancement of spontaneous emission in metal-dielectric-metal plasmon waveguide structures," *Physical Review B*, vol. 78, no. 15, p. 153111, 2008.

[102] H. M. Wong, M. K. Dezfouli, L. Sun, S. Hughes, and A. S. Helmy, "Nanoscale plasmonic slot waveguides for enhanced raman spectroscopy," *Physical Review B*, vol. 98, no. 8, p. 085124, 2018.

[103] C. Chen, X. Xu, Y. Li, H. Jans, P. Neutens, S. Kerman, G. Vereecke, F. Holsteyns, G. Maes, L. Lagae, *et al.*, "Full wetting of plasmonic nanopores through two-component droplets," *Chemical science*, vol. 6, no. 11, pp. 6564–6571, 2015.

[104] D. O'kane and K. Mittal, "Plasma cleaning of metal surfaces," *Journal of Vacuum Science and Technology*, vol. 11, no. 3, pp. 567–569, 1974.

[105] A. Dhakal, P. Wuytens, A. Raza, N. Le Thomas, and R. Baets, "Silicon nitride background in nanophotonic waveguide enhanced raman spectroscopy," *Materials*, vol. 10, no. 2, p. 140, 2017.

[106] J.-X. Cheng and X. S. Xie, "Coherent anti-stokes raman scattering microscopy: Instrumentation, theory, and applications," *The Journal of Physical Chemistry B*, vol. 108, no. 3, pp. 827–840, 2004.

[107] C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. He, J. C. Tsai, J. X. Kang, and X. S. Xie, "Label-free biomedical imaging with high sensitivity by stimulated raman scattering microscopy," *Science*, vol. 322, no. 5909, pp. 1857–1861, 2008.

[108] Y. Ozeki, W. Umemura, Y. Otsuka, S. Satoh, H. Hashimoto, K. Sumimura, N. Nishizawa, K. Fukui, and K. Itoh, "High-speed molecular spectral imaging

of tissue with stimulated raman scattering," *Nature photonics*, vol. 6, no. 12, pp. 845–851, 2012.

[109] L. Wei, Z. Chen, L. Shi, R. Long, A. V. Anzalone, L. Zhang, F. Hu, R. Yuste, V. W. Cornish, and W. Min, "Super-multiplex vibrational imaging," *Nature*, vol. 544, no. 7651, pp. 465–470, 2017.

[110] L. Wei and W. Min, "Electronic preresonance stimulated raman scattering microscopy," *The journal of physical chemistry letters*, vol. 9, no. 15, pp. 4294–4301, 2018.

[111] C.-S. Liao, P. Wang, P. Wang, J. Li, H. J. Lee, G. Eakins, and J.-X. Cheng, "Spectrometer-free vibrational imaging by retrieving stimulated raman signal from highly scattered photons," *Science advances*, vol. 1, no. 9, p. e1500738, 2015.

[112] C.-R. Hu, M. N. Slipchenko, P. Wang, P. Wang, J. D. Lin, G. Simpson, B. Hu, and J.-X. Cheng, "Stimulated raman scattering imaging by continuous-wave laser excitation," *Optics letters*, vol. 38, no. 9, pp. 1479–1481, 2013.

[113] J. L. Doménech and M. Cueto, "Sensitivity enhancement in high resolution stimulated raman spectroscopy of gases with hollow-core photonic crystal fibers," *Optics letters*, vol. 38, no. 20, pp. 4074–4077, 2013.

[114] K. Krupa, K. Baudin, A. Parriaux, G. Fanjoux, and G. Millot, "Intense stimulated raman scattering in co 2-filled hollow-core fibers," *Optics letters*, vol. 44, no. 21, pp. 5318–5321, 2019.

[115] H. Bao, W. Jin, Y. Miao, and H. L. Ho, "Laser-induced dispersion with stimulated raman scattering in gas-filled optical fiber," *Journal of Lightwave Technology*, vol. 37, no. 9, pp. 2120–2125, 2019.

[116] P. G. Westergaard, M. Lassen, and J. C. Petersen, "Differential high-resolution stimulated cw raman spectroscopy of hydrogen in a hollow-core fiber," *Optics express*, vol. 23, no. 12, pp. 16320–16328, 2015.

[117] X. Fan and I. M. White, "Optofluidic microsystems for chemical and biological analysis," *Nature photonics*, vol. 5, no. 10, p. 591, 2011.

[118] J.-X. Cheng and X. S. Xie, "Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine," *Science*, vol. 350, no. 6264, 2015.

[119] R. Claps, D. Dimitropoulos, V. Raghunathan, Y. Han, and B. Jalali, "Observation of stimulated raman amplification in silicon waveguides," *Optics express*, vol. 11, no. 15, pp. 1731–1739, 2003.

[120] A. Liu, H. Rong, M. Paniccia, O. Cohen, and D. Hak, "Net optical gain in a low loss silicon-on-insulator waveguide by stimulated raman scattering," *Optics Express*, vol. 12, no. 18, pp. 4261–4268, 2004.

[121] A. Liu, H. Rong, R. Jones, O. Cohen, D. Hak, and M. Paniccia, "Optical amplification and lasing by stimulated raman scattering in silicon waveguides," *Journal of Lightwave Technology*, vol. 24, no. 3, p. 1440, 2006.

[122] R. Jones, A. Liu, H. Rong, M. Paniccia, O. Cohen, and D. Hak, "Lossless optical modulation in a silicon waveguide using stimulated raman scattering," *Optics Express*, vol. 13, no. 5, pp. 1716–1723, 2005.

[123] H. Rong, A. Liu, R. Jones, O. Cohen, D. Hak, R. Nicolaescu, A. Fang, and M. Paniccia, "An all-silicon raman laser," *Nature*, vol. 433, no. 7023, pp. 292–294, 2005.

[124] L. Sirleto, A. Vergara, and M. A. Ferrara, "Advances in stimulated raman scattering in nanostructures," *Advances in Optics and photonics*, vol. 9, no. 1, pp. 169–217, 2017.

[125] L. E. Buchanan, N. L. Gruenke, M. O. McAnally, B. Negru, H. E. Mayhew, V. A. Apkarian, G. C. Schatz, and R. P. Van Duyne, "Surface-enhanced femtosecond stimulated raman spectroscopy at 1 mhz repetition rates," *The journal of physical chemistry letters*, vol. 7, no. 22, pp. 4629–4634, 2016.

[126] Y. E. Monfared, T. M. Shaffer, S. S. Gambhir, and K. C. Hewitt, "continuous-wave coherent raman spectroscopy via plasmonic enhancement," *Scientific reports*, vol. 9, no. 1, pp. 1–7, 2019.

[127] P. Kumar, H. Kuramochi, S. Takeuchi, and T. Tahara, "Time-domain observation of surface-enhanced coherent raman scattering with 105–106 enhancement," *The Journal of Physical Chemistry Letters*, vol. 11, no. 15, pp. 6305–6311, 2020.

[128] K. Reynkens, S. Clemmen, A. Raza, H. Zhao, J. S.-D. Peñaranda, C. Detavernier, and R. Baets, "Mitigation of photon background in nanoplasmonic all-on-chip raman sensors," *Optics express*, vol. 28, no. 22, pp. 33564–33572, 2020.

[129] E. Liang, A. Weippert, J.-M. Funk, A. Materny, and W. Kiefer, "Experimental observation of surface-enhanced coherent anti-stokes raman scattering," *Chemical physics letters*, vol. 227, no. 1-2, pp. 115–120, 1994.

[130] D. V. Voronine, A. M. Sinyukov, X. Hua, K. Wang, P. K. Jha, E. Munusamy, S. E. Wheeler, G. Welch, A. V. Sokolov, and M. O. Scully, "Time-resolved

surface-enhanced coherent sensing of nanoscale molecular complexes," *Scientific reports*, vol. 2, no. 1, pp. 1–5, 2012.

[131] T. Ichimura, N. Hayazawa, M. Hashimoto, Y. Inouye, and S. Kawata, "Tip-enhanced coherent anti-stokes raman scattering for vibrational nanoimaging," *Physical review letters*, vol. 92, no. 22, p. 220801, 2004.

[132] Z. Yan, Z. Liu, M. Xia, A. Efimov, and Y.-H. Xie, "Broadband surface-enhanced coherent anti-stokes raman spectroscopy with high spectral resolution," *Journal of Raman Spectroscopy*, vol. 48, no. 7, pp. 935–942, 2017.

[133] L. Ouyang, T. Meyer-Zedler, K.-M. See, W.-L. Chen, F.-C. Lin, D. Akimov, S. Ehtesabi, M. Richter, M. Schmitt, Y.-M. Chang, *et al.*, "Spatially resolving the enhancement effect in surface-enhanced coherent anti-stokes raman scattering by plasmonic doppler gratings," *ACS nano*, 2020.

[134] M. Yan, L. Zhang, Q. Hao, X. Shen, X. Qian, H. Chen, X. Ren, and H. Zeng, "Surface-enhanced dual-comb coherent raman spectroscopy with nanoporous gold films," *Laser & Photonics Reviews*, vol. 12, no. 9, p. 1800096, 2018.

[135] C. Zong, Y. Xie, M. Zhang, Y. Huang, C. Yang, and J.-X. Cheng, "Plasmon-enhanced coherent anti-stokes raman scattering vs plasmon-enhanced stimulated raman scattering: Comparison of line shape and enhancement factor," *The Journal of Chemical Physics*, vol. 154, no. 3, p. 034201, 2021.

[136] T.-W. Koo, S. Chan, and A. A. Berlin, "Single-molecule detection of biomolecules by surface-enhanced coherent anti-stokes raman scattering," *Optics letters*, vol. 30, no. 9, pp. 1024–1026, 2005.

[137] S. Yampolsky, D. A. Fishman, S. Dey, E. Hulkko, M. Banik, E. O. Potma, and V. A. Apkarian, "Seeing a single molecule vibrate through time-resolved coherent anti-stokes raman scattering," *Nature Photonics*, vol. 8, no. 8, pp. 650–656, 2014.

[138] J. Levy, "Integrated nonlinear optics in silicon nitride waveguides and resonators," 2011.

[139] T. Bocklitz, T. Meyer, M. Schmitt, I. Rimke, F. Hoffmann, F. von Eggeling, G. Ernst, O. Guntinas-Lichius, and J. Popp, "Invited article: Comparison of hyperspectral coherent raman scattering microscopies for biomedical applications," *APL Photonics*, vol. 3, no. 9, p. 092404, 2018.

[140] M. Krakowski, P. Resneau, M. Garcia, E. Vinet, Y. Robert, O. Parillaud, B. Gérard, S. Kundermann, N. Torcheboeuf, and D. L. Boiko, "Stabilized high pulse energy passively mode-locked monolithic and external cavity

tapered lasers for space applications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 6, pp. 1–15, 2019.

[141] K. Van Gasse, S. Uvin, V. Moskalenko, S. Latkowski, G. Roelkens, E. Bente, *et al.*, "Recent advances in the photonic integration of mode-locked laser diodes," *IEEE Photonics Technology Letters*, vol. 31, no. 23, pp. 1870–1873, 2019.

[142] S. Cuyvers, S. Poelman, K. Van Gasse, and B. Kuyken, "Hybrid modeling approach for mode-locked laser diodes with cavity dispersion and nonlinearity," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.

[143] J. Goyvaerts, S. Kumari, S. Uvin, J. Zhang, R. Baets, A. Gocalinska, E. Pelucchi, B. Corbett, and G. Roelkens, "Transfer-print integration of gaas pin photodiodes onto silicon nitride waveguides for near-infrared applications," *Optics Express*, vol. 28, no. 14, pp. 21275–21285, 2020.

[144] P. Runge, G. Zhou, T. Beckerwerth, F. Ganzer, S. Keyvaninia, S. Seifert, W. Ebert, S. Mutschall, A. Seeger, and M. Schell, "Waveguide integrated balanced photodetectors for coherent receivers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1–7, 2017.

[145] A. Hu and V. P. Chodavarapu, "General-purpose high-speed integrated lock-in amplifier with 30 db dynamic reserve at 20 mhz," *Analog Integrated Circuits and Signal Processing*, vol. 75, no. 3, pp. 369–382, 2013.

[146] D. M. Kita, J. Michon, and J. Hu, "A packaged, fiber-coupled waveguide-enhanced raman spectroscopic sensor," *Optics Express*, vol. 28, no. 10, pp. 14963–14972, 2020.

[147] Y. Zhang, Q. Du, C. Wang, T. Fakhrul, S. Liu, L. Deng, D. Huang, P. Pintus, J. Bowers, C. A. Ross, J. Hu, and L. Bi, "Monolithic integration of broadband optical isolators for polarization-diverse silicon photonics," *Optica*, vol. 6, no. 4, pp. 473–478, 2019.

[148] L. B. Soldano and E. C. Pennings, "Optical multi-mode interference devices based on self-imaging: principles and applications," *Journal of lightwave technology*, vol. 13, no. 4, pp. 615–627, 1995.

[149] A. Rahim, J. Goyvaerts, B. Szelag, J. Fedeli, P. Absil, T. Aalto, M. Harjanne, C. Littlejohns, G. Reed, G. Winzer, S. Lischke, L. Zimmermann, D. Knoll, D. Geuzebroek, A. Leinse, M. Geiselmann, M. Zervas, H. Jans, A. Stassen, C. Domínguez, P. Muñoz, D. Domenech, A. L. Giesecke, M. C. Lemme, and R. Baets, "Open-access silicon photonics platforms in europe," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–18, 2019.

[150] E. Levrau, K. Van de Kerckhove, K. Devloo-Casier, S. Pulinthanathu Sree, J. A. Martens, C. Detavernier, and J. Dendooven, "In situ ir spectroscopic investigation of alumina ald on porous silica films: Thermal versus plasma-enhanced ald," *The Journal of Physical Chemistry C*, vol. 118, no. 51, pp. 29854–29859, 2014.

[151] J. C. Love, L. A. Estroff, J. K. Kriebel, R. G. Nuzzo, and G. M. Whitesides, "Self-assembled monolayers of thiolates on metals as a form of nanotechnology," *Chemical reviews*, vol. 105, no. 4, pp. 1103–1170, 2005.

[152] P. Du, X. Zhang, H. Yin, Y. Zhao, L. Liu, Z. Wu, and H. Xu, "In situ surface-enhanced raman scattering monitoring of reduction of 4-nitrothiophenol on bifunctional metallic nanostructure," *Japanese Journal of Applied Physics*, vol. 57, no. 3, p. 030308, 2018.

[153] S. Nevlacsil, M. Eggeling, P. Muellner, G. Koppitsch, M. Sagmeister, J. Kraft, and R. Hainberger, "Broadband sin asymmetric directional coupler for 840 nm operation," *OSA Continuum*, vol. 1, no. 4, pp. 1324–1331, 2018.

[154] M. Van Daele, M. B. E. Griffiths, A. Raza, M. M. Minjauw, E. Solano, J.-Y. Feng, R. K. Ramachandran, S. Clemmen, R. Baets, S. T. Barry, C. Detavernier, and J. Dendooven, "Plasma-enhanced atomic layer deposition of nanostructured gold near room temperature," *ACS Applied Materials & Interfaces*, vol. 11, no. 40, pp. 37229–37238, 2019.

[155] S. Triki, M. Najjar, and H. Rezig, "Optimization of mmi-wdm demultiplexer by using bpm method," in *ICTON Mediterranean Winter Conference*, pp. 1–6, IEEE, 2007.

[156] T. D. Bucio, A. Z. Khokhar, G. Z. Mashanovich, and F. Y. Gardes, "N-rich silicon nitride angled mmi for coarse wavelength division (de) multiplexing in the o-band," *Optics Letters*, vol. 43, no. 6, pp. 1251–1254, 2018.

[157] Y. Hu, R. Jenkins, F. Gardes, E. Finlayson, G. Mashanovich, and G. Reed, "Wavelength division (de) multiplexing based on dispersive self-imaging," *Optics letters*, vol. 36, no. 23, pp. 4488–4490, 2011.

[158] N. F. Tyndall, T. H. Stievater, D. A. Kozak, M. W. Pruessner, and W. S. Rabinovich, "Passive photonic integration of lattice filters for waveguide-enhanced raman spectroscopy," *Optics Express*, vol. 28, no. 23, pp. 34927–34934, 2020.

[159] A. Rahim, J. Goyvaerts, B. Szelag, J. Fedeli, P. Absil, T. Aalto, M. Harjanne, C. Littlejohns, G. Reed, G. Winzer, S. Lischke, L. Zimmermann, D. Knoll, D. Geuzebroek, A. Leinse, M. Geiselmann, M. Zervas, H. Jans, A. Stassen,

C. Domínguez, P. Muñoz, D. Domenech, A. L. Giesecke, M. C. Lemme, and R. Baets, "Open-access silicon photonics platforms in europe," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–18, 2019.

[160] A. Raza, S. Clemmen, M. Van Daele, J. Dendooven, M. B. Griffiths, S. T. Barry, A. Skirtach, C. Detavernier, and R. Baets, "On-chip surface enhanced raman spectroscopy using ald grown plasmonic nanotrenches integrated with a silicon nitride slot waveguide," in *Integratred Optics, 21st European conference, Papers*, p. 3, 2019.

[161] A. Raza, *Raman spectroscopy enhanced by on-chip dielectric and metal waveguides*. PhD thesis, Ghent University, 2019.

[162] R. W. Schafer, "What is a savitzky-golay filter?[lecture notes]," *IEEE Signal processing magazine*, vol. 28, no. 4, pp. 111–117, 2011.

[163] A. R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *Journal of Lightwave Technology*, vol. 8, no. 10, pp. 1548–1557, 1990.

[164] R. Wilson, B. A. Apgar, L. W. Martin, and D. G. Cahill, "Thermoreflectance of metal transducers for optical pump-probe studies of thermal properties," *Optics express*, vol. 20, no. 27, pp. 28829–28838, 2012.

[165] B. Ouyang, W. Bogaerts, and J. Caro, "Design of silicon mach-zehnder interferometer and ring resonator with a free spectral range tolerant against waveguide-width variations," in *IEEE Photonics Benelux Chapter/Annual Symposium 2018*, pp. 1–4, 2018.

[166] D. Zhang, M. N. Slipchenko, D. E. Leaird, A. M. Weiner, and J.-X. Cheng, "Spectrally modulated stimulated raman scattering imaging with an angle-to-wavelength pulse shaper," *Optics express*, vol. 21, no. 11, pp. 13864–13874, 2013.

[167] H. Qian, Y. Xiao, and Z. Liu, "Giant kerr response of ultrathin gold films from quantum size effect," *Nature communications*, vol. 7, no. 1, pp. 1–6, 2016.

[168] N. L. Gruenke, M. F. Cardinal, M. O. McAnally, R. R. Frontiera, G. C. Schatz, and R. P. Van Duyne, "Ultrafast and nonlinear surface-enhanced raman spectroscopy," *Chemical Society Reviews*, vol. 45, no. 8, pp. 2263–2290, 2016.

[169] C. Clavero, "Plasmon-induced hot-electron generation at nanoparticle/metal-oxide interfaces for photovoltaic and photocatalytic devices," *Nature Photonics*, vol. 8, no. 2, pp. 95–103, 2014.

[170] S. Hashimoto, D. Werner, and T. Uwada, "Studies on the interaction of pulsed lasers with plasmonic gold nanoparticles toward light manipulation, heat management, and nanofabrication," *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, vol. 13, no. 1, pp. 28–54, 2012.

[171] J. Olson, S. Dominguez-Medina, A. Hoggard, L.-Y. Wang, W.-S. Chang, and S. Link, "Optical characterization of single plasmonic nanoparticles," *Chemical Society Reviews*, vol. 44, no. 1, pp. 40–57, 2015.

[172] S. S. Lo, M. S. Devadas, T. A. Major, and G. V. Hartland, "Optical detection of single nano-objects by transient absorption microscopy," *Analyst*, vol. 138, no. 1, pp. 25–31, 2013.

[173] C. Voisin, N. Del Fatti, D. Christofilos, and F. Vallée, "Ultrafast electron dynamics and optical nonlinearities in metal nanoparticles," *The Journal of Physical Chemistry B*, vol. 105, no. 12, pp. 2264–2280, 2001.

[174] H. Harutyunyan, A. B. Martinson, D. Rosenmann, L. K. Khorashad, L. V. Besteiro, A. O. Govorov, and G. P. Wiederrecht, "Anomalous ultrafast dynamics of hot plasmonic electrons in nanostructures with hot spots," *Nature nanotechnology*, vol. 10, no. 9, pp. 770–774, 2015.

[175] R. Berera, R. van Grondelle, and J. T. Kennis, "Ultrafast transient absorption spectroscopy: principles and application to photosynthetic systems," *Photosynthesis research*, vol. 101, no. 2, pp. 105–118, 2009.

[176] A. Gaiduk, P. V. Ruijgrok, M. Yorulmaz, and M. Orrit, "Detection limits in photothermal microscopy," *Chemical Science*, vol. 1, no. 3, pp. 343–350, 2010.

[177] V. K. Pustovalov, "Light-to-heat conversion and heating of single nanoparticles, their assemblies, and the surrounding medium under laser pulses," *RSC Adv.*, vol. 6, pp. 81266–81289, 2016.

[178] S. Hashimoto, D. Werner, and T. Uwada, "Studies on the interaction of pulsed lasers with plasmonic gold nanoparticles toward light manipulation, heat management, and nanofabrication," *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, vol. 13, no. 1, pp. 28 – 54, 2012.

[179] C. Zong and J.-X. Cheng, "Origin of dispersive line shapes in plasmon-enhanced stimulated raman scattering microscopy," *Nanophotonics*, vol. 10, no. 1, pp. 617–625, 2020.

[180] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of physics*. John Wiley & Sons, 2013.

[181] P. Ramirez-Priego, D. Martens, A. A. Elamin, P. Soetaert, W. Van Roy, R. Vos, B. Anton, R. Bockstaele, H. Becker, M. Singh, *et al.*, "Label-free and real-time detection of tuberculosis in human urine samples using a nanophotonic point-of-care platform," *ACS sensors*, vol. 3, no. 10, pp. 2079–2086, 2018.

[182] Hamamatsu, "CMOS linear image sensors." https://www.hamamatsu.com/resources/pdf/ssd/s13014_kmpd1170e.pdf. Accessed: 2021-09-21.

[183] Hamamatsu, "Multi-pixel photon counter." https://www.hamamatsu.com/resources/pdf/ssd/s13720_series_kapd1060e.pdf. Accessed: 2021-09-21.

[184] Newport, "Femtowatt photoreceiver, silicon, 320-1050 nm, 1 mm diameter, 8-32 / m4." https://www.newport.com/p/2151. Accessed: 2021-09-21.

[185] H. Hisada, A. Okayama, T. Hoshino, J. Carriere, T. Koide, Y. Yamamoto, and T. Fukami, "Determining the distribution of active pharmaceutical ingredients in combination tablets using near ir and low-frequency raman spectroscopy imaging," *Chemical and Pharmaceutical Bulletin*, vol. 68, no. 2, pp. 155–160, 2020.

[186] S. E. Kichanov, D. P. Kozlenko, A. K. Kirillov, E. V. Lukin, B. Abdurakhimov, N. M. Belozerova, A. V. Rutkauskas, T. I. Ivankina, and B. N. Savenko, "A structural insight into the chelyabinsk meteorite: neutron diffraction, tomography and raman spectroscopy study," *SN Applied Sciences*, vol. 1, no. 12, pp. 1–9, 2019.

[187] J.-B. Wu, M.-L. Lin, X. Cong, H.-N. Liu, and P.-H. Tan, "Raman spectroscopy of graphene-based materials and its applications in related devices," *Chemical Society Reviews*, vol. 47, no. 5, pp. 1822–1873, 2018.

[188] R. Wiesinger, L. Pagnin, M. Anghelone, L. M. Moretto, E. F. Orsega, and M. Schreiner, "Pigment and binder concentrations in modern paint samples determined by ir and raman spectroscopy," *Angewandte Chemie International Edition*, vol. 57, no. 25, pp. 7401–7407, 2018.

[189] A. Yashchenok, A. Masic, D. Gorin, B. S. Shim, N. A. Kotov, P. Fratzl, H. Möhwald, and A. Skirtach, "Nanoengineered colloidal probes for raman-based detection of biomolecules inside living cells," *Small*, vol. 9, no. 3, pp. 351–356, 2013.

[190] C. V. Navin, L. S. Lawson, and J. D. Rodriguez, "Librarybased screening of pharmaceutical materials by handheld raman and nearinfrared spectrometers," 2016.

[191] J. A. Randall and M. G. Lyman, "Discrimination of bacterial growth media using portable raman spectroscopy with background fluorescence subtraction," 2013.

[192] A. Martín-Gómez, N. Arroyo-Manzanares, M. García-Nicolás, Á. I. López-Lorente, S. Cárdenas, I. López-García, P. Viñas, M. Hernández-Córdoba, and L. Arce, "Portable raman spectrometer as a screening tool for characterization of iberian dry-cured ham," *Foods*, vol. 10, no. 6, p. 1177, 2021.

[193] Imec, "How to shrink a raman spectroscope and analyze complex samples on-the-go?." https://fdocuments.in/reader/full/how-to-shrink-a-raman-spectroscope-and-analyze-complex-and-to-learn-more-about. Accessed: 2021-09-07.

[194] C. Rathmell, D. Bingemann, M. Zieg, and D. Creasey, "Portable raman spectroscopy: Instrumentation and technology," *Portable Spectroscopy and Spectrometry 1: Technologies and Instrumentation*, vol. 1, pp. 115–145, 2021.

[195] P. C. Wuytens, H. Demol, N. Turk, K. Gevaert, A. G. Skirtach, M. Lamkanfi, and R. Baets, "Gold nanodome sers platform for label-free detection of protease activity," *Faraday discussions*, vol. 205, pp. 345–361, 2017.

[196] N. F. Tyndall, T. H. Stievater, D. A. Kozak, K. Koo, R. A. McGill, M. W. Pruessner, W. S. Rabinovich, and S. A. Holmstrom, "Waveguide-enhanced raman spectroscopy of trace chemical warfare agent simulants," *Optics letters*, vol. 43, no. 19, pp. 4803–4806, 2018.

[197] K. Vunckx, B. Geelen, V. G. Munoz, W. Lee, H. Chang, P. Van Dorpe, H. A. Tilmans, S. H. Nam, and A. Lambrechts, "Towards a miniaturized application-specific raman spectrometer," in *Sensing for Agriculture and Food Quality and Safety XII*, vol. 11421, p. 1142108, International Society for Optics and Photonics, 2020.

[198] P. M. Maya-Hernández, L. C. Álvarez-Simón, M. T. Sanz-Pascual, and B. Calvo-López, "An integrated low-power lock-in amplifier and its application to gas detection," *Sensors*, vol. 14, no. 9, pp. 15880–15899, 2014.

[199] P. B. Johnson and R.-W. Christy, "Optical constants of the noble metals," *Physical review B*, vol. 6, no. 12, p. 4370, 1972.

[200] D. Xiang and R. Gordon, "Nanoparticle acoustic resonance enhanced nearly degenerate four-wave mixing," *ACS Photonics*, vol. 3, no. 8, pp. 1421–1425, 2016.

[201] S.-C. Yang, H.-C. Lin, T.-M. Liu, J.-T. Lu, W.-T. Hung, Y.-R. Huang, Y.-C. Tsai, C.-L. Kao, S.-Y. Chen, and C.-K. Sun, "Efficient structure resonance energy transfer from microwaves to confined acoustic vibrations in viruses," *Scientific reports*, vol. 5, no. 1, pp. 1–10, 2015.

[202] J. Burkhartsmeyer, Y. Wang, K. S. Wong, and R. Gordon, "Optical trapping, sizing, and probing acoustic modes of a small virus," *Applied Sciences*, vol. 10, no. 1, p. 394, 2020.

[203] N. Tcherniega, S. Pershin, A. Bunkin, E. Donchenko, O. Karpova, A. Kudryavtseva, V. Lednev, T. Mironova, M. Shevchenko, M. Strokov, *et al.*, "Laser excitation of gigahertz vibrations in cauliflower mosaic viruses' suspension," *Laser Physics Letters*, vol. 15, no. 9, p. 095603, 2018.

[204] E. R. May and C. L. Brooks III, "Determination of viral capsid elastic properties from equilibrium thermal fluctuations," *Physical review letters*, vol. 106, no. 18, p. 188101, 2011.

[205] S. Wheaton, R. M. Gelfand, and R. Gordon, "Probing the raman-active acoustic vibrations of nanoparticles with extraordinary spectral resolution," *Nature Photonics*, vol. 9, no. 1, pp. 68–72, 2015.

[206] M. Fujii, T. Nagareda, S. Hayashi, and K. Yamamoto, "Low-frequency raman scattering from small silver particles embedded in sio 2 thin films," *Physical Review B*, vol. 44, no. 12, p. 6243, 1991.

[207] A. H. Yang, S. D. Moore, B. S. Schmidt, M. Klug, M. Lipson, and D. Erickson, "Optical manipulation of nanoparticles and biomolecules in sub-wavelength slot waveguides," *Nature*, vol. 457, no. 7225, pp. 71–75, 2009.

[208] L. Saviot and D. B. Murray, "Long lived acoustic vibrational modes of an embedded nanoparticle," *Physical review letters*, vol. 93, no. 5, p. 055506, 2004.

[209] D. B. Murray and L. Saviot, "Phonons in an inhomogeneous continuum: Vibrations of an embedded nanoparticle," *Physical Review B*, vol. 69, no. 9, p. 094305, 2004.

[210] A. Raza, "Glucose sensing using photonics waveguide based evanescent raman spectroscopy," *Photonics Research*, 2015.