

Smooth seas do not make skillful sailors.

-African Proverb

University of Alberta

**Ultrafast Nonlinear and Strong-Field Phenomena in Silicon-based
Nanoplasmonic Waveguides**

by

Matthew Shawn Bror Sederberg

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Photonics and Plasmas

Department of Electrical and Computer Engineering

©Matthew Shawn Bror Sederberg

Spring 2014

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

*This thesis is dedicated to my family,
for their endless love and support.*

Abstract

This thesis presents the realization and characterization of passive and active photonic and nanoplasmonic waveguides for applications in all-optical circuitry. The key results focus on generating visible light in nanoscale silicon waveguides through nonlinear interactions and demonstrating ultrafast all-optical modulation through nonlinear loss mechanisms.

Nanofabrication processes are developed to interface silicon photonic waveguides and silicon-based nanoplasmonic waveguides, along with a technique to integrate nanoplasmonic waveguides onto a macroscopic characterization beam. Passive propagation and nonlinear interactions are investigated in silicon-on-insulator photonic waveguides to provide a detailed understanding of nonlinear interactions present in silicon at $\lambda = 1550\text{nm}$ and the relevant timescales of the interactions. Extensive investigations into third-harmonic generation in silicon photonic waveguides are performed, and conversion efficiencies up to 2.8×10^{-5} are measured.

Measurements of the passive performance of silicon-based nanoplasmonic waveguides revealed a propagation length of $2.0\mu\text{m}$ at $\lambda = 1550\text{nm}$ and a coupling efficiency of 38% to silicon photonic waveguides. The concepts of nonlinear light generation and ultrafast modulation are then applied to sub-wavelength silicon-based nanoplasmonic waveguides. Third-harmonic generation with conversion efficiencies up to 2.3×10^{-5} is demonstrated in a nanoplasmonic waveguide with a footprint of $0.43\mu\text{m}^2$. Accurate investigations of ultrafast nonlinear interactions in silicon-based nanoplasmonic waveguides integrated onto a macroscopic characterization beam are performed using pump-probe time-domain measurements.

Ponderomotive acceleration of two-photon absorption-generated free-carriers in silicon-based nanoplasmonic waveguides is examined and it is demonstrated that electrons can be

accelerated to energies exceeding the threshold for impact ionization. Measurements reveal that the highly confined nanoplasmonic field drives an electron avalanche, and white light emission resulting from the avalanche is observed to scale exponentially with the input power. The electron avalanche effectively sweeps free-carriers from the nanoplasmonic waveguide on a timescale of ~ 2 ps, allowing for a reduction in the free-carrier recovery time by more than two orders of magnitude compared to silicon photonic waveguides.

Acknowledgements

I am deeply grateful for the excellent supervision of Dr. Abdul Elezzabi. Whether it was discussing the next step of the project or analyzing results, his insight, guidance, optimism, and spark of imagination were invaluable. None of this work would have been possible if it weren't for the time and funding that were invested in the equipment and my training.

The helpful administrative and technical support of all of the staff in the Department of Electrical and Computer Engineering was also a great asset. In particular, I wish to thank the staff in the NanoFab and the Machine Shop for the services, facilities, and training they provided.

My PhD experience would not have been the same without the friendship, camaraderie, and assistance of my lab-mates: Corey Baron, Pouya Maraghechi, Michael Quong, Cameron Straatsma, Michael Startsev, Michael Nielsen, Shawn Grieg, Nir Katchinskiy, and Brett Carnio. I hope our paths cross again someday.

I am infinitely thankful to have a loving and supportive family. They have always been there for me and gave me every opportunity to pursue my dreams. Finally, thanks to my wife, Laura, for her love, for keeping me sane, and for knowing when I needed space to work and when I needed to have fun!

Contents

1	Introduction	1
2	Background	8
2.1	Classical Electromagnetics in free-space	8
2.2	Linear Response of Materials	10
2.2.1	Dispersion and Absorption in Dielectrics	12
2.2.2	Electromagnetic Response of Metals	14
2.3	Nonlinear Response of Materials	15
2.4	Silicon-on-insulator Waveguides	17
2.4.1	Passive Routing	19
2.4.2	Nonlinear Optical Effects in Silicon-on-Insulator Waveguides	22
2.4.3	Four-wave Mixing	26
2.4.4	Free-carrier Dynamics in Silicon	29
2.5	Surface Plasmon Polariton (SPP) Waves	32
2.5.1	Planar: One-dimensional Confinement	32
2.5.2	Momentum matching between free-space and surface plasmon waves	34
2.5.3	Waveguides: Two-dimensional Confinement	37
2.6	The Ponderomotive Force	44
2.6.1	Derivation of the Ponderomotive Force	47
2.7	Impact Ionization and Avalanche Multiplication	49

3	Nanofabrication Process Development	56
3.1	Introduction	56
3.2	Silicon-on-Insulator Substrates	56
3.3	Integration of Nanoplasmonic Waveguides onto a Micron-Scale Characterization Beam	58
3.3.1	Background	58
3.3.2	Device Design	59
3.3.3	Fabrication Process Flow	60
3.4	Fabrication of Silicon-on-Insulator Nanophotonic Waveguides	75
3.5	Integration of Silicon-based Nanoplasmonic Waveguides with Silicon Photonic Waveguides	76
3.6	Fabrication of Trapezoid Fabry-Pérot Resonator Waveguides	78
3.7	Summary	80
4	Ultrafast Nonlinear Effects in Silicon-on-Insulator Waveguides	83
4.1	Introduction	83
4.2	Passive Routing	86
4.2.1	The Cut-Back Method	86
4.2.2	Measurements	86
4.3	Two-photon Absorption and Free-carrier Absorption	88
4.3.1	Power Transmission Saturation	88
4.3.2	Time Dynamics	93
4.4	Self-phase Modulation	98
4.5	Thermo-optic Tuning in a Fabry-Pérot Resonator via Nonlinear Loss	105
4.6	Summary	113
5	Third-Harmonic Generation in SOI Nanophotonic Structures	120
5.1	Introduction	120
5.2	The Influence of Waveguide Mode-Confinement on Third-Harmonic Generation	121
5.2.1	Phase-matching Considerations	121
5.2.2	Observation of Third-Harmonic Generation	124

5.2.3	Mode Confinement Dependence of Third-Harmonic Generation . . .	127
5.2.4	Third-Harmonic Power Scaling	127
5.2.5	Discussion of Third-Harmonic Spectral Reshaping	130
5.2.6	Calculation of Conversion Efficiency	134
5.3	Third-Harmonic Generation in Ultrashort, High-Confinement SOI Waveguides	136
5.4	Summary	140
6	Ultrafast and Strong-field Effects in Nanoplasmonic Waveguides	143
6.1	Introduction	143
6.2	Passive Operation	144
6.3	Active Device Design	151
6.3.1	Free-space Coupling to the Plasmonic Mode	154
6.3.2	Estimation of Electric Field Strength	158
6.4	Third-Harmonic Generation	159
6.5	White Light Emission	162
6.6	Ponderomotive Electron Acceleration	164
6.6.1	Modelling of Ponderomotive Acceleration	166
6.7	Impact Ionization and Avalanche Multiplication	171
6.8	White Light Emission Model	172
6.9	Ultrafast Nonlinear Dynamics	174
6.9.1	Analysis of the Free-carrier Time-Dynamics	177
6.10	Summary	180
7	Nanoplasmonic Devices: Active and Passive	187
7.1	Overview	187
7.2	Ultrafast All-Optical Switching and Modulation in Silicon-Based Nanoplasmonic Devices	187
7.2.1	Introduction	187
7.2.2	Ultrafast Free-Carrier Time-Dynamics in Silicon	188
7.2.3	The Five-Layer Plasmonic Slot Waveguide	190
7.2.4	The Plasmonic Nanoring Resonator	199

7.2.5	Summary	206
7.3	Nanostructured Nanoplasmonic Antennas Operating in the Near- to Mid- Infrared	206
7.3.1	Introduction	206
7.3.2	Simulation Technique	208
7.3.3	Fractal Geometry	208
7.3.4	Sierpiński Fractal Nanoplasmonic Antenna	211
7.3.5	Pascal Triangle Nanoplasmonic Antenna	214
7.3.6	The Nanoplasmonic Contour Bowtie Antenna	218
7.3.7	Summary	222
8	Conclusions	229
A	Finite Difference Time Domain Simulations	235
B	Nanofabrication Process Steps	243
C	Calculation of the Theoretical Third-Harmonic Generation Spectrum	250
D	The Pump-Probe Time-Domain Technique	253
E	Spectrometer Calibration	257
F	Wavelength Scan Program	259
G	Codes for Calculating Electron Trajectories	265

List of Tables

2.1	General Collection of Nonlinear Phenomena (adapted from [2])	18
2.2	Recombination times of various Si samples	31
4.1	Maximum intensity in Si core of waveguides with different widths at a peak input power, $P_{peak}^{\omega} = 75\text{W}$	93
6.1	Summary of Mode Propagation Characteristics	148
7.1	Geometric Properties of Antennas under Investigation	214

List of Figures

2.1	Schematic depiction of the Lorentz oscillator model and the linear polarization response to a time-varying electric field. Adapted from [2].	13
2.2	Schematic depiction of the Lorentz oscillator model and the nonlinear polarization response to a time-varying electric field. Adapted from [2]. . . .	16
2.3	Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ in SOI waveguides: (a) Quasi-TE mode, $w = 500\text{nm}$. (b) Quasi-TM mode, $w = 500\text{nm}$. (c) Quasi-TE mode, $w = 250\text{nm}$. (d) Quasi-TM mode, $w = 250\text{nm}$	20
2.4	Transverse magnetic and transverse electric effective waveguide mode indices for a fixed waveguide core height, $h = 340\text{nm}$, and varying waveguide widths, w	21
2.5	Fraction of the waveguide mode power that is confined to the Si core, for waveguides with a constant height, $h = 340\text{nm}$, and varying widths, w . . .	22
2.6	(a) Top-view schematic of two SOI photonic waveguides with $w = 250\text{nm}$, separated by a 100nm gap. (b) Finite-difference time-domain intensity (I) distributions of the waveguides depicted in (a), demonstrating the periodic coupling of mode energy between adjacent waveguides.	23
2.7	Schematic depiction of second harmonic generation. Here, two photons of the fundamental frequency, ω , combine to create a single photon with double the frequency, 2ω	24

2.8	Graph depicting the concept of self-phase modulation. The temporal front of the pulse produces an increasing refractive index (via the optical Kerr effect), which corresponds to a red-shift in the frequency components. The temporal back of the pulse produces a decreasing refractive index, which corresponds to a blue-shift in the frequency components. $\Delta\omega$ represents the frequency shift of the spectral components of the pulse due to SPM.	26
2.9	Schematic depiction of four-wave mixing, where a photon is produced at $\omega_3 = 2\omega_1 - \omega_2$	27
2.10	Schematic depiction of the band structure of Si, two-photon absorption and subsequent free-carrier absorption.	28
2.11	Schematic representation of THG. Here, three photons at the fundamental frequency, ω , combine to generate a single photon at triple the frequency, 3ω	29
2.12	Schematic depiction of a surface plasmon polariton (SPP) that is propagating along the interface between a metal (bottom) and a dielectric(top). The SPP consists of an oscillation of charge density coupled to electromagnetic oscillations that are confined to the interfaces and decay exponentially in intensity in the direction normal to the interfaces.	32
2.13	Conceptual plot of the surface plasmon polariton (SPP) dispersion relation (solid line), along with the dispersion relation of free-space photons and photons propagating through a prism. Notably, there are no points where the dispersion relations for a SPP and free-space photons intersect. (Adapted from [46])	35
2.14	Schematic depiction of: (a) Kretschmann configuration, (b) Otto configuration, (c) grating coupling, (d) slit Coupling, and (e) defect coupling for excitation of surface plasmon waves.	36
2.15	Schematic of (a) metal-insulator and (b) metal-insulator-metal nanoplasmonic waveguides.	40

2.16	Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for IM plasmonic waveguides with a width, (a) $w = 250\text{nm}$; and (b) $w = 100\text{nm}$. Mode profile intensity distributions at $\lambda = 1550\text{nm}$ for MIM plasmonic waveguides with widths, (c) $w = 250\text{nm}$; and (d) $w = 100\text{nm}$	40
2.17	Effective refractive index and propagation losses of modes in Au-Si-Au plasmonic waveguides with a constant height, $h = 340\text{nm}$ and a varying width, w	41
2.18	Fraction of the mode energy that is confined to the Si core in Au-Si-Au plasmonic waveguides with a constant height, $h = 340\text{nm}$ and a varying width, w	42
2.19	(a) Top-view schematic of two Au-Si-Au plasmonic waveguides with $w = 250\text{nm}$, separated by a 100nm Au film. (b) Finite-difference time-domain intensity (I) distributions of the waveguides depicted in (a), demonstrating effective confinement of the mode energy to the original waveguide.	43
2.20	Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for hybrid plasmonic waveguides with a height, $h = 340\text{nm}$, and widths, $w =$ (a) 250nm ; and (b) 100nm	44
2.21	Mode profile intensity (I) distribution of a long-range surface plasmon polariton waveguide with a gold film thickness of 20nm and a width, $w = 5\mu\text{m}$	45
2.22	Schematic depiction of cycle-by-cycle ponderomotive electron acceleration by a SPP field.	46
2.23	Schematic depiction of impact ionization process, leading to avalanche multiplication of free-carriers.	49
3.1	Cross-sectional SEM of a cleaved SOI wafer with a 465nm thick PECVD oxide.	58
3.2	Rendering of the envisioned sample and excitation scheme. Integrating nanoplasmonic devices onto a large characterization beam would allow for direct probing and detection with macroscopic optical components, including microscope objectives and optical fibres.	60

3.3	Nanofabrication process flow for the definition of nanoplasmonic waveguides with a Si core and Au cap. (a) Cleaned SOI wafer. (b) Electron-beam lithography exposure, followed by electron beam evaporation of Cr(5nm) + Au(60nm) + Cr(5nm) + SiO ₂ (40nm). (c) Evaporated films after lift-off. (d) A fully-defined nanoplasmonic waveguide after ICPRIE etching through the Si device layer. (e) A nanoplasmonic waveguide clad in a 465nm thick SiO ₂ . (f) Subsequent deposition of 150nm thick Cr etch mask.	61
3.4	(a) Schematic of a nanoplasmonic waveguide buried beneath a 465nm SiO ₂ cladding and a 150nm Cr etch mask. (b) The Cr etch mask consisting of a cleave guide adjacent to one waveguide end facet and a large area adjacent to the other end facet. (c) Etching through the BOX layer and 75μm into the Si substrate. (d) Schematic of the final structure consisting of nanoplasmonic waveguides integrated onto a characterization beam. The SiO ₂ cladding has been omitted for clarity.	63
3.5	(a) Optimization of the Cr adhesion layer thickness between the Si device layer and the Au nanoplasmonic waveguide cap. (a) Marginal adhesion with a 2.5nm thick Cr adhesion layer, where most of the Au cap adheres to the Si core, but other regions are detached. (b) Close-up of a region of poor adhesion.	64
3.6	Optimization of the EBL exposure for a 495k PMMA A6 resist with a thickness of 275nm, exposed at $V = 20\text{keV}$ and $d_{ap} = 20\mu\text{m}$. (a) Dose test for a nominal dose of $150\mu\text{C}/\text{cm}^2$; and (b) $300\mu\text{C}/\text{cm}^2$. Electron beam evaporation was used to deposit Cr(5nm) + Au(60nm) + Cr(5nm) + SiO ₂ (40nm) layers and a standard lift-off was performed in acetone. A final dose of $D = 450\mu\text{C}/\text{cm}^2$ was used to obtain a very high $\approx 100\%$ yield lift-off. . . .	65
3.7	Optimization of the design pattern rectangle width required to obtain a lift-off feature with a width of 100nm, at an exposure dose, $Dose = 450\mu\text{C}/\text{cm}^2$. Several widths were obtained during this process, including: (a) $w = 68\text{nm}$, (b) $w = 83\text{nm}$, (c) $w = 96\text{nm}$, and (d) $w = 112\text{nm}$	66

3.8	Lift-off features with $w \approx 100\text{nm}$ obtained using the final exposure parameters: (a) Top-view; and (b) 45° tilt-view of a cleaved lift-off feature, illustrating the topography.	67
3.9	Example of severe under-cut obtained when using a SiO_2+Au etch mask for a process that was designed for a PMMA mask.	68
3.10	Redevelopment of the unswitched Si Bosch-like process used for etching through the Si device layer. All parameters were kept fixed, except for the C_4F_8 gas flow. Scanning-electron micrographs are shown for C_4F_8 flow rates of: (a) 40sccm, (b) 45sccm, (c) 55sccm, (d) 60sccm, (e) 65sccm, (f) 70sccm, (g) 75sccm, and (h) 80sccm. Subsequent increases in the C_4F_8 flow rate did not yield any improvement in the sidewall roughness.	69
3.11	Typical etching results obtained with the redeveloped etching process, with a C_4F_8 flow rate of 80sccm.	70
3.12	(a) Several nanoplasmonic waveguides clad in a 465nm thick PECVD oxide. (b) Close-up of a single nanoplasmonic waveguide clad in a 465nm thick PECVD oxide. (c) Close-up of a single nanoplasmonic waveguide clad in 465nm thick PECVD oxide and a 150nm thick Cr etch mask.	71
3.13	Optimization of a Bosch process used for deep silicon etching. (a) Sidewalls obtained using a standard process revealed significant roughness near the etch mask. (b) Close-up of the roughness near the etch mask, where the etching scallops were barely distinguishable. (c) Development of a less aggressive etching process minimized the roughness near the etch mask. (d) Close-up of the sidewall near the etch mask, revealing well-defined scallops.	73
3.14	(a) Scanning electron micrograph of the Cr etch mask defined adjacent to a buried nanoplasmonic waveguide, prior to etching through the BOX layer or into the Si substrate. (b) Scanning electron micrograph of the final structure, imaged after deep etching, removal of the Cr etch mask, and cleaving. (c) Top-view SEM of the final structure. (d) Top-view optical microscope image of the final structure.	74

3.15	Cross-sectional scanning electron micrographs of end facets of waveguides integrated onto the characterization beam. (a) Nanoplasmonic waveguide with $w \approx 100\text{nm}$, (b) $w = 340\text{nm}$ photonic waveguide, and (c) $w = 340\text{nm}$ photonic and plasmonic waveguides integrated onto the same characterization beam.	75
3.16	Top-view SEM of a $w = 2.0\mu\text{m}$ silicon photonic waveguide.	76
3.17	Fabrication steps for interfacing Si photonic waveguides with nanoplasmonic waveguides. (a) Exposure of nanoplasmonic features for lift-off metallization. (b) Nanoplasmonic waveguide cap after lift-off. (c) Exposure of Si photonic waveguide etch mask. (d) Final structures after silicon device layer etching and resist removal.	77
3.18	Top-view SEM of $w = 300\text{nm}$ nanoplasmonic waveguides interfaced to tapered silicon photonic waveguides. (a) $L = 1\mu\text{m}$, (b) $L = 2\mu\text{m}$, (c) $L = 3\mu\text{m}$, and (d) a multiple input, multiple output nanoplasmonic device.	78
3.19	Optical microscope image of silicon photonic waveguides interfaced to nanoplasmonic structures.	79
3.20	Fabrication steps for obtaining TFPRs on the corner of a SOI wafer. (a) A $400\mu\text{m} \times 400\mu\text{m}$ array of Si photonic waveguides is defined at a 45° angle to the cleave planes of the sample. (b) The sample was cleaved in half, where the cleave was aligned to intersect the waveguide array. (c) Each half was cleaved again to intersect the waveguide arrays, yielding two samples with TFPRs (denoted by \checkmark) and two unusable samples (denoted by \times). (d) Top-view SEM of TFPR waveguides fabricated on a SOI substrate.	79
4.1	Schematic depiction of a SOI waveguide.	87
4.2	Plot of power transmitted through waveguides with fixed cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and lengths of $200\mu\text{m}$, $615\mu\text{m}$, $825\mu\text{m}$, and $1120\mu\text{m}$. An exponential decay function is fit to these data points and reveals a propagation length of $790\mu\text{m}$, corresponding to a linear propagation loss of $\alpha = 55\text{dB/cm}$. The y -intercept represents the power transmitted through a waveguide of infinitesimal length and provides the input and output coupling loss product.	88

4.3	Plot demonstrating the saturation of power transmission through a waveguide with cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and a length of $L = 1120\mu\text{m}$. The squares represent the average output power transmitted through the waveguide as a function of the peak power coupled to the waveguide. The solid line is an extrapolation of the low-power linear region of the measured data points and represents the average output power that would be observed in the absence of nonlinear loss mechanisms in the silicon core of the waveguide. The nonlinear loss is calculated by subtracting the measured data points from the linear extrapolation, and is represented by the triangles.	90
4.4	Power transmission saturation plots for four waveguides with fixed cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and varying lengths of $L = 200\mu\text{m}, 615\mu\text{m}, 825\mu\text{m}, 1120\mu\text{m}$	91
4.5	Quasi-transverse-magnetic mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$, for a fixed peak input power, $P_{peak}^\omega = 75\text{W}$, for waveguides with $h = 340\text{nm}$ and $w =$ (a) 260nm , (b) 340nm , (c) 470nm , (d) 540nm , (e) 760nm , and (f) 1000nm	92
4.6	Power transmission saturation plots for five waveguides with a fixed length, $L = 1120\mu\text{m}$, a fixed height, $h = 340\text{nm}$, and varying widths of $w = 260\text{nm}, 340\text{nm}, 470\text{nm}, 540\text{nm}, 760\text{nm}$	94
4.7	Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at $P_{peak}^\omega = 28\text{W}$. At a $t = 0\text{ps}$ time-delay, TPA produces a sharp drop in transmission. For a $t > 0\text{ps}$ time delay, FCA attenuates the probe beam by an amplitude that is proportional to the instantaneous free-carrier density. An exponential fit to the recovery tail yields a free-carrier recombination time of $\tau_r = 265\text{ps}$	95

4.8	Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at peak pump powers of $P_{peak}^\omega = 2.8\text{W}$ (black), $P_{peak}^\omega = 11.0\text{W}$ (red), $P_{peak}^\omega = 22.1\text{W}$ (green), and $P_{peak}^\omega = 33.1\text{W}$ (blue). The nonlinear loss imparted to the probe due to TPA and FCA from the pump increases with the pump power, enabling control of the modulation depth.	96
4.9	Free-carrier recovery times extracted from the recovery tail of pump-probe time domain traces for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at different peak pump powers.	97
4.10	Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 25\mu\text{m}$	98
4.11	Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 6.5\mu\text{m}$	99
4.12	Transmitted $\lambda = 1550\text{nm}$ pulse spectrum through a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 610\mu\text{m}$, taken at peak input powers of $P_{peak}^\omega = 7\text{W}$ (dotted line) and $P_{peak}^\omega = 152\text{W}$ (solid line). Both spectra are normalized and shown on a logarithmic y -axis scale.	101
4.13	(a) Surface depicting the transmitted pulse spectrum versus input peak power for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 610\mu\text{m}$. (b) Total transmitted spectral power versus input peak power, depicting transmission saturation due to TPA and FCA. Transmitted power is also calculated for the spectral ranges: (c) $1517\text{nm} \leq \lambda \leq 1519\text{nm}$, (d) $1535.7\text{nm} \leq \lambda \leq 1542\text{nm}$, (e) $1552.6\text{nm} \leq \lambda \leq 1559.6\text{nm}$, and (f) $1561\text{nm} \leq \lambda \leq 1570\text{nm}$	103
4.14	Surfaces depicting the transmitted pulse spectrum versus input peak power for waveguides with $h = 340\text{nm}$, $L = 610\mu\text{m}$, and $w =$ (a) 260nm , (b) 340nm , (c) 470nm , (d), 540nm , (e) 760nm , (f) 1000nm , (g) 1510nm , (h) 2050nm , and (i) 4150nm . Sections of the transmitted spectrum taken at $P_{peak}^\omega = 100\text{W}$ are plotted in white.	104

4.15	Plot of detected power at $\lambda = 1465\text{nm}$ versus peak input power, shown on a logarithmic x - and y -axis, for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 200\mu\text{m}$	105
4.16	Surfaces depicting the transmitted pulse spectrum versus input peak power for waveguides with $h = 340\text{nm}$, $L = 200\mu\text{m}$, and $w =$ (a) 260nm , (b) 340nm , (c) 470nm , (d), 540nm , (e) 760nm , (f) 1000nm , (g) 1510nm , (h) 2050nm , and (i) 4150nm . Section of the transmitted spectrum taken at $P_{peak}^\omega = 100\text{W}$ are plotted in white.	106
4.17	(a) Schematic depiction of the sample under test and the excitation and detection scheme. A microscope objective (shown on right) couples light into the TFPR and a lensed single mode optical fibre (shown on left), oriented perpendicularly to the angle of excitation, collects the transmitted radiation. A scanning electron micrograph of the $L = 7.80\mu\text{m}$ TFPR is shown in the inset. (b) Scanning electron micrograph of the fabricated sample, with multiple TFPR lengths.	108
4.18	(a) Top-view camera image of scattered third-harmonic radiation from the waveguide interfaces following a zig-zag pattern as the mode undergoes total internal reflection at the waveguide interfaces. (b) Time-averaged intensity (I) distribution in the waveguide obtained by finite-difference time-domain simulation, verifying the multiple total internal reflections as the mode propagates from the input facet to the output. The length of the depicted TFPR is $L = 31.1\mu\text{m}$	109
4.19	Measured broadband transmission spectrum from a TFPR with $L = 7.80\mu\text{m}$ (black) and the calculated spectrum obtained using finite-difference time-domain simulations.	110
4.20	Power transmission saturation plots for a TFPR with $L = 7.80\mu\text{m}$, displaying a linear region and a nonlinear loss region. The linear region is extrapolated and the experimental values are subtracted from this line to obtain the nonlinear loss.	111

4.21	(a) Broadband transmission spectrum of the TFPR as the femtosecond laser power coupled to the device is increased from $P_{ave}^\omega = 0\text{mW}$ to $P_{ave}^\omega = 11.3\text{mW}$. (b) Logarithmic x - and y -axis plots relating the input power to the wavelength shift and resonance amplitude attenuation. The wavelength shift has a slope of 2.01 ± 0.02 , while the attenuation does not show a linear trend.	112
4.22	(a) Broadband transmission spectrum for operating temperatures ranging from $298\text{K} \leq T \leq 478\text{K}$ in steps of $\Delta T = 20\text{K}$, obtained via finite-difference time-domain simulations, in the absence of nonlinear loss. The peak wavelength of the resonance red-shifts linearly with increasing temperature according to the relation $\Delta\lambda_{max} = 7.26 \times 10^{-2}\Delta T$ [nm]. (b) Temperature distribution for a thermal power of 3mW distributed uniformly in the TFPR.	114
5.1	Transverse-magnetic mode intensity (I) distributions for a SOI waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$ at (a) $\lambda = 1550\text{nm}$; and (b) $\lambda = 517\text{nm}$	122
5.2	Dispersion of modes in a $w = 260\text{nm}$, $h = 340\text{nm}$ SOI waveguide. Near-infrared TE (dotted line) and TM (solid line) waveguide modes have significantly different effective refractive indices, whereas the TE and TM modes (dashed line) at green wavelengths ($500\text{nm} \leq \lambda \leq 535\text{nm}$) have virtually identical refractive indices. Effective refractive indices were calculated with a mode solver.	123
5.3	Electric field (E_y) distribution of quasi-TM modes in a $w = 340\text{nm}$, $h = 340\text{nm}$ SOI waveguide at (a) $\lambda = 1550\text{nm}$ and (b) $\lambda = 517\text{nm}$, recorded in the xz -plane, as depicted in Fig. 2.15.	124
5.4	Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for: (a) the quasi-TM mode; and (b) the poorly-confined leaky mode.	126
5.5	Calculated third-harmonic intensity (I) distribution for: (a) the quasi-TM mode; (b) the poorly-confined, leaky mode; and (c) both modes excited simultaneously.	126

5.6	(a) Camera image of the interference pattern present in the waveguide taken at $20\times$ with a colour charge-coupled device camera. Mode beating with a period of $\Lambda = 3.05 \pm 0.85\mu\text{m}$ is observed. (b) Third-harmonic intensity (I) interference pattern with $\Lambda = 2.62 \pm 0.26\mu\text{m}$, calculated via FDTD simulations. (c) Artistic rendering of the sample under test along with the scattered green light that is captured with a microscope camera. The mesh plot contains the experimental data collected from the microscope camera. (d) Measured spectrum from the output facet of the SOI waveguide at $P_{peak}^\omega = 100\text{W}$	128
5.7	Intensity surfaces depicting the THG spectrum growth with increasing peak input power, P_{peak}^ω , for waveguides with $w =$ (a) 4150nm, (b) 2050nm, (c) 1510nm, (d) 1000nm, (e) 760nm, (f) 540nm, (g) 470nm, (h) 340nm, and (i) 260nm. The white line overlay portrays the third-harmonic spectrum measured at the dashed section, corresponding to $P_{peak}^\omega=100\text{W}$	129
5.8	Plots of third-harmonic spectrum peak wavelength (squares and solid line) and FWHM (circles and dotted line) as a function of waveguide width, w , for a fixed $P_{peak}^\omega = 100\text{W}$	130
5.9	Plots depicting the measured third-harmonic power as a function of the excitation $\lambda = 1550\text{nm}$ power, plotted on logarithmic x - and y - axes. Two data sets are plotted. The first represents the $\lambda = 1550\text{nm}$ power coupled to the input facet of the waveguide (circles), which scales with a slope of 2.20 ± 0.05 . The second represents the $\lambda = 1550\text{nm}$ power at the output facet of the waveguide, which is the most accurate measure of the power used to excite the THG, and scales with a slope of 3.02 ± 0.10	131
5.10	Intensity surfaces depicting the normalized third-harmonic spectrum with increasing peak input power for waveguides with $w =$ (a) 4150nm, (b) 2050nm, (c) 1510nm, (d) 1000nm, (e) 760nm, (f) 540nm, (g) 470nm, (h) 340nm, and (i) 260nm.	132

5.11	Transmitted fundamental spectra through a $L = 200\mu\text{m}$, $w = 260\text{nm}$, $h = 340\text{nm}$ waveguide for peak powers, $P_{peak}^\omega = 10\text{W}$ (blue) and $P_{peak}^\omega = 100\text{W}$ (red), along with a spectrum taken with no SOI waveguide. (b) Transmitted fundamental spectra through a $L = 200\mu\text{m}$, $w = 340\text{nm}$, $h = 340\text{nm}$ waveguide for peak powers, $P_{peak}^\omega = 10\text{W}$ (blue) and $P_{peak}^\omega = 100\text{W}$ (red), along with a spectrum taken with no SOI waveguide.	135
5.12	Third-harmonic generation conversion efficiency as a function of waveguide width, w , where each waveguide is excited with the same peak power, $P_{peak}^\omega = 100\text{W}$	136
5.13	(a) Short SOI waveguide with a width, $w = 340\text{nm}$, a height, $h = 340\text{nm}$, and lengths in the range, $1.8\mu\text{m} \leq L \leq 7.6\mu\text{m}$. Femtosecond $\lambda = 1550\text{nm}$ pulses are coupled to the $L = 5.2\mu\text{m}$ waveguide facet on the left-hand side. Green light emission is clearly visible (under normal room light illumination) at the waveguide end facets and is collected with an optical fibre on the right-hand side. (b) Measured and calculated third-harmonic spectra for a $L = 7.6\mu\text{m}$ waveguide. Green light emission is observed with a $20\times$ microscope camera from SOI waveguides with lengths, $L =$ (c) $1.8\mu\text{m}$, (d) $5.2\mu\text{m}$, (e) $6.5\mu\text{m}$, and (f) $7.6\mu\text{m}$	138
5.14	Intensity surfaces depicting the THG spectrum growth with increasing peak input power, P_{peak}^ω , for SOI waveguides with $L =$ (a) $3.7\mu\text{m}$, (b) $5.2\mu\text{m}$, (c) $6.5\mu\text{m}$, and (d) $7.6\mu\text{m}$. The white line overlay portrays the third-harmonic spectrum measured at the dashed section, corresponding to $P_{peak}^\omega = 100\text{W}$	139
5.15	Intensity surfaces depicting the normalized third-harmonic spectrum with increasing peak input power for SOI waveguides with $L =$ (a) $3.7\mu\text{m}$, (b) $5.2\mu\text{m}$, (c) $6.5\mu\text{m}$, and (d) $7.6\mu\text{m}$	139
6.1	Schematic depiction of a silicon-based nanoplasmonic waveguide interfaced to silicon photonic input and output couplers. The insets depict the mode profiles intensity (I) distributions for the photonic and plasmonic waveguides.	145

6.2	(a) Scanning electron micrograph of a silicon-based plasmonic waveguide interfaced to silicon photonic input and output tapered waveguides. The silicon-based plasmonic waveguide has a width, $w = 300\text{nm}$, a height, $h = 340\text{nm}$, and a length, $L = 4\mu\text{m}$. The inset reveals a $\sim 35\text{nm}$ misalignment between the photonic and plasmonic waveguides. (b) Scanning electron micrograph of a silicon-based plasmonic s-bend, with a transverse displacement of $\Delta x = 2\mu\text{m}$ and an overall length of $L = 6.3\mu\text{m}$	146
6.3	Simulation of interfacing between silicon photonic and silicon-based plasmonic waveguides. (a) Snapshot of electric fields (E_y) in the $x = 0\mu\text{m}$ plane, depicting the excitation of a plasmonic mode with a silicon photonic input coupler and the subsequent out-coupling to the output silicon photonic waveguide. (b) Snapshot of electric fields (E_y) in the $z = 330\text{nm}$ plane, highlighting the excitation of a plasmonic mode and the tight confinement near the Au-Si interface. (c) Electric field ($ E $) distribution of the silicon photonic mode taken $z = -500\text{nm}$ before the beginning of the plasmonic waveguide. (d) Electric field ($ E $) distribution at the interface between the silicon photonic and silicon-based plasmonic waveguides, showing strong diffraction at the Au interface. (e) Electric field ($ E $) distribution of the plasmonic mode, taken $z = 500\text{nm}$ after the beginning of the plasmonic waveguide.	147
6.4	Intensity (I) distributions of the modes supported by a Si-based nanoplasmonic with $w = 300\text{nm}$ and $h = 340\text{nm}$	149
6.5	Measured power transmission through $w = 300\text{nm}$, $h = 340\text{nm}$ silicon-based plasmonic waveguides. (a) Power transmitted through waveguides with lengths of $L = \{2.0, 3.0, 4.0, 5.0\}\mu\text{m}$ at $\lambda = 1550\text{nm}$. An exponential fit to the data points reveals a propagation length of $L_{prop} = 2.0\mu\text{m}$ and a coupling efficiency of 38.0% per interface. (b) Broadband propagation losses, measured in the range of $1520\text{nm} \leq \lambda \leq 1615\text{nm}$	150

6.6	(a) Schematic depiction of the Si-based nanoplasmonic waveguide with $w = 95\text{nm}$, $h = 340\text{nm}$, and $L = 5\mu\text{m}$. (b) Time-averaged electric field amplitude ($ E $) distribution in the xy -plane of the structure. (c) Time-averaged intensity (I) distribution in the yz -plane of the structure. (d) Time-averaged intensity (I) distribution in the xz -plane of the waveguide at $y = h/2 = 170\text{nm}$	152
6.7	(a) Schematic depiction of the conceived device, excitation scheme, and observed visible emission. (b) Scanning electron micrograph of a single nanoplasmonic waveguide before deposition of SiO_2 cladding. (c) Scanning electron micrograph of several nanoplasmonic waveguides fabricated on a narrow beam. A close-up of the cross-sectional geometry of the waveguide end-facet is shown in the inset. (d) Optical microscope image of visible emission from above the nanoplasmonic waveguides.	153
6.8	Intensity (I) distribution of focusing by a $NA=0.85$ microscope objective, shown on a logarithmic scale.	155
6.9	Electric field distributions at the focal plane of a $NA = 0.85$ microscope objective for: (a) E_x , (b) E_y , and (c) E_z . The intensity (I) distribution at the focal plane is shown in (d).	155
6.10	Deviation of the focal spot from the ideal Gaussian distribution with $w_0 = 580\text{nm}$ at the focus of a $NA = 0.85$ microscope objective.	156
6.11	Simulation of the excitation of the silicon-based nanoplasmonic waveguide with a $NA = 0.85$ microscope objective. (a) Time snapshot of focused free-space electric fields (E_y) incident on the input facet of the nanoplasmonic waveguide and coupling to a tightly confined mode at the Si-Au interface. Fields also couple to poorly-defined surface modes on the air-Au and Si-SiO ₂ interfaces, but become negligible within $\approx 2\mu\text{m}$. (b) Time snapshot of E_y taken in the $y = 330\text{nm}$ plane, demonstrating the coupling and enhancement of focused incident radiation to a plasmonic mode at the Au-Si interface. (c) Intensity (I) distribution of the excited plasmonic mode profile in the $z = 2\mu\text{m}$ plane.	157

6.12	Spectra of the ultrafast laser (Input), along with measured spectra of visible light collected from the end-facet of the silicon-based nanoplasmonic waveguide. The visible light consists of a strong THG ($500\text{nm} \leq \lambda \leq 530\text{nm}$) peak and a broad white light spectrum ($375\text{nm} \leq \lambda \leq 725\text{nm}$).	159
6.13	(a) Surface plot of third-harmonic power as a function of wavelength and input power measured at the output facet of the nanoplasmonic waveguide. The black overlaid spectrum is a 3ω spectrum measured at an average input power, $P_{in}^\omega = 1.36\text{mW}$, while the green spectrum is the 3ω spectrum calculated from the laser pulse spectrum. (b) Plots on logarithmic x - and y -axes of detected 3ω power versus input power for the entire 3ω spectrum (green) and individual wavelength, $\lambda = \{512, 517, 520\}\text{nm}$	160
6.14	Conversion efficiency as a function input power, plotted on logarithmic x - and y -axes. A linear fit to the data points reveals a slope of 1.92, which is very close to the expected value of 2.	162
6.15	(a) Surface plot of white light photon counts measured at the output facet of the nanoplasmonic waveguide as a function of wavelength and power. The black overlaid spectrum is measured at an average input power, $P_{in}^\omega = 1.42\text{mW}$, and the green curves represent the position of the 3ω components, which have been removed for clarity. (b) Plots of experimentally measured white light counts versus input power for the entire spectrum (black) and individual wavelengths, $\lambda = \{472, 537, 578\}\text{nm}$	163
6.16	Total white light counts versus input power, shown on various axes scales: (a) Linear x - and y -axes, along with an exponential growth fit. (b) Logarithmic x - and y -axes, with various linear fits, revealing an increasing slope with increasing input power. (c) Linear x -axis and natural-logarithmic y -axis, along with a linear fit.	165
6.17	Schematic depiction of physical processes relevant to the electron acceleration, including intense, asymmetric E_{SP} field propagation, TPA, cycle-by-cycle acceleration of the electron in the E_{SP} field leading to a net ponderomotive force, and electron and hole avalanche multiplication.	167

6.18	Sample calculated trajectories of 21 electrons excited via TPA at random points in time, with random positions and initial velocities. The green dots represent the starting point of the electron trajectory and the red dots denote the location of an electron colliding with the gold surface. . . .	167
6.19	(a) Electric field experienced by an electron during the course of its trajectory in response to the E_{SP} field, shown for three different E_{SP} field strengths. (b) Instantaneous velocity reached by the electron in response to the E_{SP} field, shown for three different E_{SP} field strengths. It should be noted that these calculations neglect collisions and should only be taken as valid before the electron undergoes its first collision (the first $\sim 28\text{nm}$ of its trajectory). The red and green curves are terminated when the electron has traversed the entire Si waveguide core and collides with the SiO_2 interface. (c) Calculated electron energy spectra for three different E_{SP} values in the same range as those used in the experiments. The electron energy represents the average energy of the electrons before undergoing their first collision. (d) Plot of the electron energy spectral peak and FWHM versus E_{SP} , taken from smoothed spectra.	170
6.20	(a) Schematic depiction of photon emission processes: (i) direct interband (D-I) recombination, (ii) intraband transitions (I-T) or Bremsstrahlung radiation, and (iii) indirect interband (I-I) recombination. (b) Plot of photon counts versus photon energy, as measured in the experiments, along with the fitted curves obtained using a multimechanism model including D-I (magenta), I-T (blue), and I-I (orange).	175
6.21	Artistic rendering of the ultrafast pump-probe characterization setup. . . .	175
6.22	Pump-probe time-dependent power transmission for a co-polarized configuration.	176

6.23	(a) Pump-probe time-dependent power transmission for the cross-polarized pump-probe configuration, for three exemplary pump powers. (b) A single pump-probe trace with a shaded area representing the total nonlinear loss experienced by the probe as the delay is varied. (c) Total nonlinear loss versus input pump power, plotted on logarithmic x - and y -axes. A linear fit to this data yields a slope of 2.52, confirming that observed excitation scales nonlinearly with the input power. (d) Pump-probe time-domain trace for a Si photonic waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, for $P_{peak}^\omega = 22.1\text{W}$	177
6.24	Recovery tail fit with: (a) A single exponential recovery; (b) A biexponential recovery; and (c) A power law recovery. (d) Power scaling of the two time-constants of the recovery tail, retrieved from a biexponential fit. . . .	178
7.1	(a) A schematic diagram showing the nanoplasmonic waveguide utilized as an ultrafast optical modulator and/or switch for infrared radiation. (b) Depicts an embodiment of the device realized using II-Si and silver (Ag) layers fabricated on top of a silicon dioxide (SiO_2) layer and a Si layers. The device can be fabricated using conventional metal and Si deposition and etching techniques, whereas the free-space region is fabricated via focused ion beam milling.	191
7.2	(a) Spatial distribution of time-averaged power, $\langle S \rangle$ normalized to 1W per unit width in the y -direction for 1550nm radiation within a five layer waveguide consisting of Ag(100nm) : Si(150nm) : free-space(50nm) : Si(150nm) : Ag(100nm) calculated using (a) the analytical model and (b) a finite element solution to Maxwell equations. The cross section is taken at the centre of the waveguide in panel (c)). (c) The corresponding mode profile intensity (I) distribution inside the waveguide calculated using finite element method numerical solution to Maxwell's equations.	194

7.3	(a) Real, $\eta_s(t)$, and imaginary, $\kappa_s(t)$, values of the refractive index in the first 6ps after excitation at various photoexcitation carrier densities ranging from $n_0/n_c = 0.05$ to 0.2. (b) $\beta_r(t)$ and $\beta_i(t)$ at the same photoexcitation carrier density range from $n_0/n_c = 0.05$ to 0.2 calculated for the $(a, d) = (50\text{nm}, 150\text{nm})$ waveguide configuration.	196
7.4	Minimum phase, β_r^{min} , and maximum attenuation, β_i^{max} , parameters for each set of (a, d) dimensions as a function of n_0/n_c from 0 to 0.2, for the waveguides having $(a, d) = (0\text{nm}, 150\text{nm})$, $(50\text{nm}, 150\text{nm})$, and $(100\text{nm}, 150\text{nm})$	197
7.5	FDTD intensity (I) distribution of $\lambda = 1550\text{nm}$ signal laser at (a) $n_0/n_c = 0$, (b) $n_0/n_c = 0.5$, and (c) $n_0/n_c = 1.1$ for a Ag(100nm):Si(150nm): free-space(50nm):Si(150nm):Ag(100nm) waveguide.	197
7.6	(a) The normalized output intensity of the ultrafast plasmonic switch as a function of time for three modulators. (a) Ag(100nm) : Si(300nm) : Ag(100nm) waveguide, (b) Ag(100nm) : Si(150nm) : free-space(50nm) : Si(150nm) : Ag(100nm), and (c) Ag(100nm) : Si(150nm) : free-space(100nm) : Si(150nm) : Ag(100nm). Each waveguide modulator configuration is calculated for photoinjected carrier densities of $n_0/n_c = 0$ (on state), 0.05, 0.1, 0.15, and 0.2. (d) On-off extinction ratios are calculated as function of the photoinjected carrier densities for the same nanoplasmonic waveguides.	199
7.7	Schematic depiction of the device geometry. The nanoring is designed to have a radius, $r = 560\text{nm}$, silver film thickness, $t_{Ag} = 100\text{nm}$, input coupler separation, $g_{Si} = 25\text{nm}$, modulation coupler separation, $g_{SiO_2} = 20\text{nm}$, and uniform waveguide widths, $w_{Si} = w_{SiO_2} = 100\text{nm}$	200
7.8	(a) Mode profile intensity (I) distribution of the excited mode in a Si-loaded plasmonic waveguide at $\lambda = 1515\text{nm}$. (b) Broadband transmission through Si bus plasmonic waveguide coupled to nanoring resonator. (c) Mode profile intensity (I) distribution of the excited mode in a SiO ₂ -loaded plasmonic waveguide at $\lambda = 800\text{nm}$. (d) Pump power ($\lambda = 800\text{nm}$) coupled to nanoring versus nanoring angle, θ (see inset). A skewed Gaussian function is fitted to the recorded points.	201

7.9	Refractive index of Si as a function of time and nanoring angle when excited by ultrafast above-bandgap pulses of $\tau_p = 10\text{fs}$ duration at $\lambda = 800\text{nm}$. The real component of the refractive index of II-Si in the ring is modelled for pump strengths of $n_0/n_c=0.05, 0.10, 0.15,$ and 0.20 in (a), (b), (c), and (d), respectively. The imaginary component of the refractive index of II-Si in the nanoring is modelled for pump strengths of $n_0/n_c=0.05, 0.10, 0.15,$ and 0.20 in (e), (f), (g), and (h), respectively.	203
7.10	(a) Intensity (I) distribution of the nanoring resonator in the “off” state without any pump. (b) Intensity (I) distribution of the ring resonator for a pump strength of $n_0/n_c=0.10$. (c) Intensity (I) distribution of the nanoring resonator in the “on” state, with a pump strength of $n_0/n_c = 0.22$. (d) Effect of the pump strength on the position of the transmission minimum.	205
7.11	(a) Dependence of the power transmission through II-Si bus waveguide on the pump strength. (b) Power transmission through the II-Si bus waveguide as a function of time for pump strengths of $n_0/n_c = \{0.05, 0.10, 0.15, 0.22\}$.	205
7.12	(a) Sierpiński triangle basis shape; and (b)-(d) first three iterations of Sierpiński triangle. (e)-(h) Modulus $M = 2, 3, 4$ and 5 Pascal triangles, respectively. (i) The relationship between modulus and Hausdorff dimension for a Pascal triangle.	210
7.13	Schematic representations of the simulation geometries: (a) bowtie antenna and (b) first-, (c) second- and (d) third-iteration Sierpiński fractal plasmonic antennas. The broadband electric field enhancement factor is shown in (e). The antenna colour in (a)-(d) corresponds to the line colour in (e).	210
7.14	Empirical relationships between antenna length and (a) resonant wavelength and (b) resonant enhancement factor for each antenna geometry under consideration. The colour of antennas on the right corresponds to the line colours in (a) and (b).	212

7.15	Intensity (I) distributions at resonance for (a) bowtie antenna and (b) first-, (c) second- and (d) third-iteration Sierpiński fractal plasmonic antennas. The line section shows the intensity (I) distribution along the x -axis of the antenna and within the antenna gap.	213
7.16	Schematic representation of (a) bowtie antenna, and bowtie antennas with Pascal triangle modulus (b) $M = 2$, (c) $M = 3$, and (d) $M = 4$ geometry. .	214
7.17	Broadband enhancement factor for a bowtie antenna with $L = 475\text{nm}$, along with $M = \{3, 4, 5, 6\}$ antennas with $L = 475\text{nm}$	215
7.18	Logarithmic scale intensity (I) distributions for (a) $M = 3$, (b) $M = 4$, and (c) $M = 5$ antennas. Intensity distributions are normalized to the input excitation.	216
7.19	(a) Resonant wavelength versus modulus for $L = 475\text{nm}$ (constant). A parabolic trend line is included. (b) Resonant wavelength versus antenna length for $M = 4$ (constant). The trend for a bowtie antenna is shown for comparison. (c) Paraboloid surface-fit to resonant wavelengths of 55 antennas that were simulated.	217
7.20	Schematic representations of the simulation geometries: (a) bowtie antenna and (b) contour bowtie antenna.	217
7.21	(a) Broadband enhancement factors for a bowtie antenna and contour bowtie antennas with $t = \{40, 60, 80\}\text{nm}$ and a length, $x = 475\text{nm}$. (b) λ_{res} versus antenna length for a bowtie antenna and a $t = 50\text{nm}$ contour bowtie antenna, each with a length, $x = 475\text{nm}$. (c) λ_{res} versus contour thickness, t , for an antenna length, $x = 475\text{nm}$. λ_{res} for a bowtie antenna with $x = 475\text{nm}$ is shown for reference. (d) Paraboloid fit to 70 simulated data points relating λ_{res} to $\{x, t\}$	219
7.22	Logarithmic scale intensity (I) distributions at spectral points of interest: (a) Secondary resonance at $\lambda = 1.2\mu\text{m}$, (b) anti-resonance at $\lambda = 1.5\mu\text{m}$, (c) primary resonance at $\lambda = 2.2\mu\text{m}$, and (d) above resonance at $\lambda = 3.9\mu\text{m}$.	219

7.23	(a) Schematic representation of a contour bowtie antenna with $t_v \neq t$. (b) λ_{res} versus t_v for $\{x, t\} = \{475, 30\}$ nm. λ_{res} for a bowtie antenna is shown for reference. (c) Schematic representation of a contour bowtie antenna with a shifted vertical edge. (d) λ_{res} versus edge displacement, d . The case when the vertical edge is removed altogether is shown for reference.	221
A.1	A unit cell of the FDTD Yee scheme.	236
B.1	Schematic depiction of the (a) passivation and (b) etching steps of the Bosch process.	248
C.1	(a) Femtosecond laser spectrum, measured in free-space. (b) Calculated third-harmonic spectrum.	251
D.1	Schematic representation of the pump-probe system.	254
D.2	Views of the pump-probe system.	256
E.1	(a) Plot of fibre-coupled power versus free-space laser power. (b) Plot of measured spectrometer versus fibre-coupled power.	258

List of Abbreviations and Symbols

Fundamental Constants

$c = 2.99702458 \times 10^8 \text{m/s}$	speed of light in free-space
$\varepsilon_0 = 8.85418782 \times 10^{-12} \text{F/m}$	permittivity of free-space
$\mu_0 = 4\pi \times 10^{-7} \text{H/m}$	permeability of free-space
$e = 1.602176565 \times 10^{-16} \text{C}$	elementary charge
$h = 6.62606057 \times 10^{-34} \text{J}\cdot\text{s}$	Planck constant
$k_B = 1.3806488 \times 10^{-23} \text{J/K}$	Boltzmann constant

Abbreviations

2D	two-dimensional
3D	three-dimensional
a.u.	arbitrary units
BOX	buried oxide
BTE	Boltzmann transport equation
CB	conduction band
CMOS	complementary metal-oxide-semiconductor
CW	continuous-wave
DC	direct current
DFG	difference frequency generation
D-I	direct interband
EBL	electron beam lithography
FBMS	fixed-beam moving stage
FCA	free-carrier absorption

FCD	free-carrier dispersion
FDTD	finite-difference time-domain
FOM	figure of merit
FWHM	full-width at half-maximum
FWM	four wave mixing
ICPRIE	inductively-coupled plasma reactive ion etching
I-I	indirect interband
IPA	isopropyl alcohol
I-T	direct interband
JFET	junction gate field-effect transistor
MI	metal-insulator
MIM	metal-insulator-metal
MIBK	methyl isobutyl ketone
NA	numerical aperture
NIR	near infrared
OSA	optical spectrum analyzer
PECVD	plasma-enhanced chemical vapour deposition
PhC	photonic crystal
PMMA	poly (methyl methacrylate)
R-C	resistive-capacitive
RF	radio frequency
RIE	reactive ion etching
RPM	revolutions per minute
SEM	scanning electron micrograph
SFG	sum frequency generation
SHG	second-harmonic generation
SMF	single-mode fibre
SNR	signal-to-noise ratio
SOI	silicon-on-insulator
SPM	self-phase modulation
SRH	Shockley-Read-Hall

SPP	surface plasmon polariton
TE	transverse electric
TEM	transverse electromagnetic
TFPR	trapezoid Fabry-Pérot resonator
TH	third-harmonic
THG	third-harmonic generation
TM	transverse magnetic
TPA	two-photon absorption
VB	valence band
XAM	cross-amplitude modulation
XPM	cross-phase modulation

Chapter 1

Introduction

A key revolution in information processing was achieved by developing devices to route electrons and modulate current, and developing techniques to integrate these devices to a very high density on a chip. With the prominent role of fibre optics in telecommunications, it is becoming equally important to gain the ability to route and control light, or photons. Although the principle of guiding light by refraction was achieved in 1842 [1], it wasn't until 1965 that the appropriate components and techniques were developed and a data transmission system based on fibre optics was demonstrated. Numerous further practical developments were required before the propagation losses were reduced to an acceptable level and fibres could be drawn to many kilometres in length [2]. Fibre optics has become a revolutionary technology and a significant percentage of all information transmission occurs through optical fibres.

This success has driven major efforts to develop integrated optical circuitry, which would feature similar functionality as integrated electronic circuits, but with higher data processing bandwidths and lower energy consumption. As the cardinal material of the electronics industry, silicon (Si) has amassed a comprehensive infrastructure of growth and processing techniques. This all-encompassing foundation has enabled rapid advances in Si nanophotonic devices for routing, filtering, buffering, and modulating near-infrared (NIR) electromagnetic radiation signals [3, 4]. Propagation losses have been reduced to approximately 1dB/cm in Si waveguides, enabling low-loss optical signal routing across chips and high quality-factor filtering. Industry has realized that Si photonic interconnects can be used to replace copper interconnects, alleviating resistive-capacitive time delays in computer processors and hybrid electronic-photonic circuitry is currently in advanced stages of research and development.

Despite these promising advances, the maximum integration density of Si photonic devices is limited by the diffraction limit, and a dimensional mismatch between electronic and photonic circuitry is apparent. For example, the 22 nm processing currently used to fabricate computer chips enables a 25 nm gate length, and further dimensional reductions to 14 nm processing are scheduled to occur within the next year. Such compact dimensions yield a significantly higher density of electronic devices than what is possible with photonic

waveguides, and hence, an alternate class of waveguides for high-density optical routing over short distances is desirable.

Nanoplasmonic waveguides enable light guiding at the interfaces of metallic structures with dimensions on the order of a few tens of nanometers [5]. With no stringent diffraction limit, these structures can be integrated to a similar density as electronic structures in regions where dense hybrid electrical-optical circuitry is required. Furthermore, they can enable both electrical and optical signals to be transmitted along the same structure [6], permitting advanced functionality. The strong coupling of light to the metallic features of nanoplasmonic devices provides electric field enhancement by up to three orders of magnitude, depending on the device geometry, permitting efficient access to sensitive light-matter interaction processes [7].

Unfortunately, these benefits come at the expense of higher propagation losses than photonic waveguides, due to the relatively high loss of metals at optical and near-infrared frequencies. In a subwavelength plasmonic waveguide, this limits the propagation distance of electromagnetic energy to the range of several microns to approximately $100\mu\text{m}$, depending on the waveguide design and the materials used. These losses limit the utility of plasmonic waveguides in large optical circuits, as it may be necessary to transmit the optical signal across a distance of several millimetres, and the signal would attenuate to the noise floor. Therefore, it is clear that nanoplasmonic waveguides have advantages and shortcomings, and that they must be used appropriately in integrated optical circuitry. In specific, it is crucial to interface subwavelength plasmonic waveguides to low-loss photonic bus waveguides for across-chip signal transmission. In this way, a more flexible design scheme enabling both high-density optical circuitry and low-loss signal transmission may be realized.

In the case of information processing, electric field enhancement enables strong non-linear interactions, which can be used to develop all-optical modulators. While promising passive plasmonic devices have been developed, there have been comparatively few demonstrations of ultrafast, all-optical plasmonic waveguide modulators [8]. Moreover, most of these have not embodied the subwavelength nature of plasmons, limiting their functionality in large-scale integrated circuitry. It has been shown that radiation-induced interband transitions in an aluminum waveguide can produce a plasmonic signal modulation of 7.5% on a femtosecond timescale [9]. Surface plasmon grating coupling has been modulated by excitation of free electrons in gold followed by nonequilibrium thermalization dynamics and by photogenerated free-carrier modulation in Si [10, 11]. It has also been shown that interference of a femtosecond pulse with a slowly depolarizing plasmonic system can be used to modulate the plasmonic signal [12]. Electric field enhancement can also be used to access extreme nonlinearities or strong-field phenomena, such as extreme ultraviolet fluorescence from noble gases or ponderomotive electron acceleration to energies of $\sim 2\text{keV}$ [13–15].

While the structures used in these investigations have effectively demonstrated the

principle of all-optical modulation in a plasmonic system, they do not truly embody the potential for ultrafast chip-scale plasmonic circuitry. Therefore, it is necessary to investigate ultrafast, nonlinear interactions in subwavelength devices that can be readily integrated with other technologies such as Si photonics or complementary metal-oxide-semiconductor (CMOS) electronics.

Thesis Objectives

The primary aim of this thesis is to demonstrate the potential for ultrafast all-optical nanoplasmonic circuitry that is almost entirely compatible with CMOS processing techniques ¹. These objectives are summarized as follows:

1. Develop high-yield nanofabrication processes for Si-based nanophotonic and nanoplasmonic structures.
2. Gain a detailed understanding of the nonlinear interactions and free-carrier dynamics in Si photonic waveguides.
3. Demonstrate passive routing in Si-based nanoplasmonic waveguides and interfacing to Si photonic waveguides.
4. Develop high-efficiency ultrafast visible light sources in Si nanophotonic and nanoplasmonic structures via third-harmonic generation.
5. Investigate strong-field interactions arising from ponderomotive electron acceleration in Si-based nanoplasmonic waveguides.
6. Demonstrate an all-optical electron avalanche process in Si-based nanoplasmonic waveguides, and its potential usage as an ultracompact integrated avalanche photodetector or ultrafast all-optical modulator.
7. Measure electron sweeping across Si-based nanoplasmonic waveguides due to ponderomotive electron acceleration, using pump-probe time-domain spectroscopy.
8. Investigate a new white light emission process in Si and its exponential power scaling, arising from collisional processes in the electron avalanche.
9. Demonstrate the potential for ultrafast all-optical modulation in nanoplasmonic structures based on free-carrier absorption and free-carrier dispersion.

¹With the exception of gold and silver features, which are not used in CMOS processing due to diffusion. They are, however, used for interconnects in military-specification electronic circuits. The gold features could be replaced with alternative plasmonic materials, such as aluminum, copper, or TiN, which are CMOS-compatible.

10. Investigate techniques to reduce the footprint area of nanoplasmonic antennas, making them more suitable for high-density integration with nanoplasmonic waveguides, for the purpose of coupling free-space radiation to the nanoplasmonic waveguide mode.

Thesis Organization

This thesis contains a total of eight chapters that present various aspects of the simulation, realization, and characterization of Si photonic and Si-based nanoplasmonic devices. Chapter 2 presents the relevant background theory of electromagnetics, nonlinear optics, silicon-on-insulator waveguides, plasmonics, ponderomotive electron acceleration, and impact ionization.

Substantial nanofabrication process development was required to realize the devices that were experimentally investigated, which is detailed in Chapter 3. Various aspects of the device design considerations are discussed, and the overall nanofabrication process flows are presented. Notably, each step of these processes required significant development and optimization, and the measures taken to achieve a high-yield process are examined.

The fabrication and characterization techniques were validated by investigating nonlinear interactions and ultrafast time-dynamics in Si photonic waveguides, which also established a performance baseline for subsequent nanoplasmonic devices. Several investigations into nonlinear optical effects and free-carrier dynamics are presented in Chapter 4, including two-photon absorption (TPA), free-carrier absorption, free-carrier plasma dispersion, and self-phase modulation.

In Chapter 5, broadband third-harmonic generation is demonstrated in Si photonic waveguides, and the power scaling, conversion efficiency, and interplay with other nonlinear interactions is studied. The potential for compact visible light sources on a Si chip is assessed by reducing the waveguide footprint area and record conversion efficiencies are demonstrated.

With a detailed understanding of nonlinear interactions in Si photonic waveguides, subsequent investigations were performed in sub-wavelength Si-based nanoplasmonic waveguides in Chapter 6. Simulations of ponderomotive acceleration of TPA-generated electrons reveal that electron energies can exceed the threshold for impact ionization, providing the conditions for an electron avalanche. Strong visible light emission in these structures is observed to scale exponentially with the input power, confirming the presence of the avalanche.

Additionally, ponderomotively-accelerated electrons are swept away from the vicinity of the nanoplasmonic mode at the saturation velocity of Si. The electron sweeping is visualized using pump-probe time-domain spectroscopy and confirms that the electrons are removed from the region being probed at the saturation velocity, corresponding to a timescale of approximately 2ps. Third-harmonic generation conversion efficiencies com-

parable to what was observed in the Si photonic waveguides is demonstrated, but in a footprint area reduced by a factor of 4.

The final chapter presents 5 types of nanoplasmonic devices:

- A Si-based ultrafast all-optical nanoplasmonic waveguide modulator. This structure works on the principle of free-carrier absorption and enables a signal at $\lambda = 1550\text{nm}$ to be modulated on a timescale of 5ps when above bandgap pulses excite free-carriers in the Si features of the structure.
- A Si-based ultrafast all-optical nanoplasmonic nanoring resonator modulator. This structure operates on the principle of free-carrier plasma dispersion and modulates a signal at $\lambda = 1515\text{nm}$ on a 3ps timescale when above bandgap pulses excite free-carriers in the Si features of the structure. This device has the advantage of being more compact and more energy efficient than a straight nanoplasmonic waveguide operating on the principle of free-carrier absorption.
- A nanoplasmonic antenna with Sierpiński fractal geometry.
- A nanoplasmonic antenna with Pascal's triangle fractal geometry.
- A contour nanoplasmonic bowtie antenna.

Each of the three classes of nanoplasmonic antennas operates at a resonant wavelength that is red-shifted compared to a standard bowtie antenna, allowing for more compact antenna designs.

References

- [1] D. Colladon. Sur les réflexions d'un rayon de lumière à l'intérieur d'une veine liquide parabolique. *Comptes Rendus de l'Academie des Sciences*, 15:800–802, 1842.
- [2] T. Li. *Optical Fiber Communications: Fiber Fabrication*. Academic Pr, Orlando, 1985.
- [3] L. Pavesi and D. J. Lockwood. *Silicon Photonics*. Springer, Heidelberg, 2004.
- [4] D. J. Lockwood. *Silicon Photonics II: Components and Integration*. Springer, Heidelberg, 2011.
- [5] D. J. Dikken, M. Spasenović, E. Verhagen, D. van Oosten, and L. K. Kuipers. Characterization of bending losses for curved plasmonic nanowire waveguides. *Opt. Express*, 18:16112–16119, 2010.
- [6] J. A. Dionne, L. A. Sweatlock, M. T. Sheldon, A. P. Alivisatos, and H. A. Atwater. Silicon-based plasmonics for on-chip photonics. *Selected Topics in Quantum Electronics, IEEE Journal of*, 16:295–306, 2010.
- [7] S. A. Maier. *Plasmonics: Fundamentals and Applications*. Springer, New York, 2007.
- [8] M. Kauranen and A. V. Zayats. Nonlinear plasmonics. *Nature Photon.*, 6:737–748, 2012.
- [9] K. F. MacDonald, Z. L. Samson, M. I. Stockman, and N. I. Zheludev. Ultrafast active plasmonics: transmission and control of femtosecond plasmon signals. *Nature Photon.*, 3:55–58, 2008.
- [10] N. Rotenberg, M. Betz, and H. M. van Driel. Ultrafast all-optical coupling of light to surface plasmon polaritons on plain metal surfaces. *Phys. Rev. Lett.*, 105:017402, 2010.
- [11] J. N. Caspers, N. Rotenberg, and H. M. van Driel. Ultrafast silicon-based active plasmonics at telecom wavelengths. *Opt. Express*, 18:19761–19769, 2010.
- [12] T. Utikal, M. I. Stockman, A. P. Heberle, M. Lippitz, and H. Giessen. All-optical control of the ultrafast dynamics of a hybrid plasmonic system. *Phys. Rev. Lett.*, 104:113903, 2010.

- [13] S. E. Irvine, A. Dechant, and A. Y. Elezzabi. Generation of 0.4-keV femtosecond electron pulses using impulsively excited surface plasmons. *Phys. Rev. Lett.*, 93:184801, 2004.
- [14] M. Siviş, M. Duwe, B. Abel, and C. Ropers. Extreme-ultraviolet light generation in plasmonic nanostructures. *Nature Phys.*, 9:304–309, 2013.
- [15] M. Siviş and C. Ropers. Generation and bistability of a waveguide nanoplasma observed by enhanced extreme-ultraviolet fluorescence. *Phys. Rev. Lett.*, 111:085001, 2013.

Chapter 2

Background

The results presented in this thesis explore the interaction of near-infrared femtosecond pulses with silicon (Si) photonic and Si-loaded nanoplasmonic waveguides. For sufficiently intense laser pulses, this interaction incorporates aspects of waveguides, surface plasmon polariton waves, nonlinear optics, electron acceleration, and avalanche electron multiplication. This chapter begins with an overview of Maxwell's equations in free-space and a discussion of how these equations can be adapted to account for the response of electrons in materials. This response falls into two regimes: (1) For sufficiently low excitation intensity, the interaction is a simple linear modification of Maxwell's equations in free-space; and (2) for higher excitation intensity ($> 10^{14}\text{W}/\text{cm}^2$), the interaction becomes more complex and a nonlinear polarization term must be taken into account. Subsequently, an overview of each of the relevant nonlinear optical effects observed in silicon-on-insulator (SOI) photonic waveguides is discussed. The next section presents the classical theory of surface plasmon waves on planar interfaces and extends these concepts to surface plasmon waveguides. Notably, for a plasmonic waveguide with asymmetric geometry, the highly confined mode profile is characterized by a high gradient in the electric field, which gives rise to a nonlinear force, known as the ponderomotive force. A derivation of this force is given in the subsequent section, and its application in electron acceleration is discussed. The final section is devoted to a collisional process known as impact ionization that can occur in semiconductors and can facilitate avalanche electron multiplication.

2.1 Classical Electromagnetics in free-space

The most comprehensive description of classical electromagnetic phenomena is summarized by Maxwell's equations, which can be expressed in differential form as [1]:

$$\nabla \cdot \mathbf{D} = \rho \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \quad (2.4)$$

In these equations, \mathbf{E} is the electric field strength, \mathbf{D} is the electric flux density, \mathbf{H} is the magnetic field strength, \mathbf{B} is the magnetic flux density, ρ is the volume density of free charges, and \mathbf{J} is the density of free currents. Equation 2.1 states that free electric charges are the source of divergence in the electric flux density, whereas Eq. 2.2 implies that the magnetic flux density never diverges (which would not be true if magnetic monopoles existed). Equation 2.3 and 2.4 are a set of coupled differential equations that describe time-varying electromagnetic fields and wave propagation. Equation 2.3 signifies that the instantaneous temporal rate of change of the local magnetic flux density determines the instantaneous curl in the local electric fields. Equation 2.4 states that the local surface current and/or the instantaneous temporal rate of change of the electric flux density determines the instantaneous curl in the local magnetic field. The spatio-temporal coupling between Eqs. 2.3-2.4 provides a basis for understanding and modelling simple wave propagation, interference, diffraction, and Gaussian optics.

In free-space, the electric flux density is related to the electric field strength by a universal constant:

$$\mathbf{D} = \varepsilon_0 \mathbf{E}, \quad (2.5)$$

where $\varepsilon_0 \approx 8.8542 \times 10^{-12} \text{F/m}$ is the permittivity of free-space. Similarly, the magnetic flux density is related to the magnetic field strength by another universal constant:

$$\mathbf{B} = \mu_0 \mathbf{H}, \quad (2.6)$$

where $\mu_0 = 4\pi \times 10^{-7} \text{H/m}$ is the permeability of free-space.

In the absence of free charges or free currents, Eqs. 2.1-2.4 can be rewritten in terms of only the electric field strength and the magnetic field strength:

$$\nabla \cdot \mathbf{E} = 0 \quad (2.7)$$

$$\nabla \cdot \mathbf{H} = 0 \quad (2.8)$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad (2.9)$$

$$\nabla \times \mathbf{H} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (2.10)$$

Applying the curl operator to Eq. 2.9, using the vector identity $\nabla \times (\nabla \times \mathbf{\Psi}) = \nabla(\nabla \cdot \mathbf{\Psi}) - \nabla^2 \mathbf{\Psi}$, and taking $\nabla \cdot \mathbf{E} = 0$ produces:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} \quad (2.11)$$

$$\nabla^2 \mathbf{E} = \mu_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) \quad (2.12)$$

$$\nabla^2 \mathbf{E} = \mu_0 \frac{\partial}{\partial t} \left(\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \quad (2.13)$$

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (2.14)$$

where $c = 1/\sqrt{\mu_0 \varepsilon_0} = 2.998 \times 10^8 \text{ m/s}$ is the speed of light in free-space. Performing a similar procedure on Eq. 2.10 produces an analogous result:

$$\nabla^2 \mathbf{H} = \frac{1}{c^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.15)$$

Equations 2.14-2.15 take the form of the wave equation and provide simpler relations that may be advantageous when modelling electromagnetic interactions in structures with no free charges or free currents.

Maxwell's equations (Eqs. 2.1-2.4) and the simpler \mathbf{E} and \mathbf{H} wave equations (Eqs. 2.14-2.15) are powerful tools for analyzing classical electromagnetic interactions. These equations are scale-invariant and accurately model light-matter interactions with structures consisting of arbitrary materials and geometries. Although the discussion presented in this section considered electromagnetic interactions in the absence of a medium (i.e. free-space), the next section will extend this discussion to incorporate broadband light-matter interaction.

2.2 Linear Response of Materials

When electromagnetic waves interact with a medium, the behaviour becomes more complex. Electric fields interact with bound and/or free electrons in the material, whereas magnetic fields interact with magnetic moments of electrons and/or atoms. Small contributions can also arise from ions or molecules with permanent dipoles, although these are negligible in the materials investigated in this thesis. In general, each material will respond to a time-varying electric field in a unique way via induced electric dipole moments, and the volume average of this response is represented by the polarization, \mathbf{P} . Similarly, each material will respond to a time-varying magnetic field in a unique way via induced magnetic moments, and the volume average of this response is represented by the magnetization, \mathbf{M} .

An induced polarization and/or magnetization acts as a source of electromagnetic waves, enabling waves to be transmitted and reflected at an interface, and enabling wave propagation through a medium. For low electric field strengths, the induced electric dipole responds linearly to the incident electric field and radiates an electromagnetic wave with the same frequency as the excitation. Through this interaction, the fields radiated by the dipole can incur a phase change and can be attenuated (due to energy transfer to forms other than electromagnetic). These interactions are described by the linear refractive index and the linear absorption.

In general, it can be said that both the incident electric field strength (i.e. the field that would be present in the absence of the electron response of the medium) and the

local polarization of the material contribute to the local electric flux density. In the same manner, the magnetic flux density depends on the incident magnetic field strength and the local magnetization of the material. These contributions are summarized as follows:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad (2.16)$$

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (2.17)$$

In the special case of a linear, isotropic material, the polarization and magnetization are related to the electric and magnetic fields via:

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E} \quad (2.18)$$

$$\mathbf{M} = \chi_m \mathbf{H}, \quad (2.19)$$

where χ_e and χ_m are the electric and magnetic susceptibilities, respectively. A simple substitution of these relations into Eqs. 2.16-2.17 then produces:

$$\mathbf{D} = \varepsilon_0 (1 + \chi_e) \mathbf{E} \quad (2.20)$$

$$\mathbf{B} = \mu_0 (1 + \chi_m) \mathbf{H}. \quad (2.21)$$

Two new parameters, known as the relative permittivity and relative permeability, can then be defined as:

$$\varepsilon_r = 1 + \chi_e \quad (2.22)$$

$$\mu_r = 1 + \chi_m. \quad (2.23)$$

Therefore, Eqs. 2.16-2.17 can be expressed as:

$$\mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E} = \varepsilon \mathbf{E} \quad (2.24)$$

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H} = \mu \mathbf{H}, \quad (2.25)$$

which are analogous to the relations for free-space (Eqs. 2.5-2.6). Using a similar approach as the derivation of Eq. 2.14, the wave equation in a medium can be written as:

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (2.26)$$

Equations 2.16-2.26 predict the interaction of monochromatic radiation with linear, isotropic dielectrics. However, the presence of resonant electronic excitations in naturally-occurring materials produces a wavelength-dependence in the electromagnetic properties, known as *dispersion* and *absorption*. Furthermore, the electromagnetic properties of some materials depends on crystallographic orientation, and the response is *anisotropic*. Finally, for sufficient electric field strength or magnetic field strength, the linear relations of Eqs.

2.18 and 2.19 lose accuracy, and the behaviour is referred to as *nonlinear*. This thesis does not deal with magnetic materials, and from hereon in it is assumed that $\mu = \mu_0$.

2.2.1 Dispersion and Absorption in Dielectrics

Although a fundamental understanding of light-matter interactions requires a rigorous treatment via quantum electrodynamics, intuitive descriptions based on classical mechanics can be used to accurately model most common optical materials. In a dielectric medium, valence electrons are bound to a host nucleus. The nucleus has a much higher inertia than the valence electrons and it can be assumed that it remains stationary in response to an electromagnetic excitation at an angular frequency, ω . For example, the mass of the nucleus of a silicon atom is approximately $m_n = 4.66 \times 10^{-26}$ kg, which is much larger than the mass of an electron, $m_e = 9.11 \times 10^{-31}$ kg. The bound electrons respond via the Lorentz force (i.e. $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$), and follow the electromagnetic excitation in their temporal displacement. For a sufficiently weak electric field, this displacement will be relatively small, and the restoring force of the nucleus can be modelled by Hooke's Law, $F_{spring} = -k\Delta x$, where k is the spring constant, Δx is the displacement from the equilibrium position, and F_{spring} is the resulting spring force. This is represented schematically in Fig. 2.1, where an atom is exposed to an electric field and an electron is displaced from the nucleus at $x = x_n$ to a new position, $x = x_e$. In this schematic representation, the electron is displaced from the nucleus by a distance, $\Delta x = x_e - x_n$, that is linearly proportional to the strength of the incident field, responds to the time-varying amplitude and direction of the incident field, and returns to its equilibrium position once the electric field ceases. The dashed line in the Fig. 2.1 represents the linear relationship between the electric field strength (horizontal axis) and the polarization of the system (vertical axis).

In this case, the resonant frequency of the simple harmonic system is determined by a resonant electron excitation at an angular frequency, ω_0 . Furthermore, the electron motion can be damped due to internal collisions in the dielectric and lost energy in the form of radiation. Summarizing these in the equation of motion yields:

$$\frac{d\mathbf{p}}{dt} = \sum \mathbf{F} \quad (2.27)$$

$$m_e^* \frac{d^2 \mathbf{r}}{dt^2} = \mathbf{F}_{driving} + \mathbf{F}_{damping} + \mathbf{F}_{spring} \quad (2.28)$$

$$m_e^* \frac{d^2 \mathbf{r}}{dt^2} = -e\mathbf{E} - m_e^* \gamma_0 \frac{d\mathbf{r}}{dt} - m_e^* \omega_0^2 \mathbf{r} \quad (2.29)$$

$$m_e^* \frac{d^2 \mathbf{r}}{dt^2} + m_e^* \gamma_0 \frac{d\mathbf{r}}{dt} + m_e^* \omega_0^2 \mathbf{r} = -e\mathbf{E}, \quad (2.30)$$

where \mathbf{p} is the electron momentum vector, \mathbf{r} is the electron position vector, $\mathbf{F}_{driving}$ is the driving force of the electric field acting on the electron, $\mathbf{F}_{damping}$ is the damping force

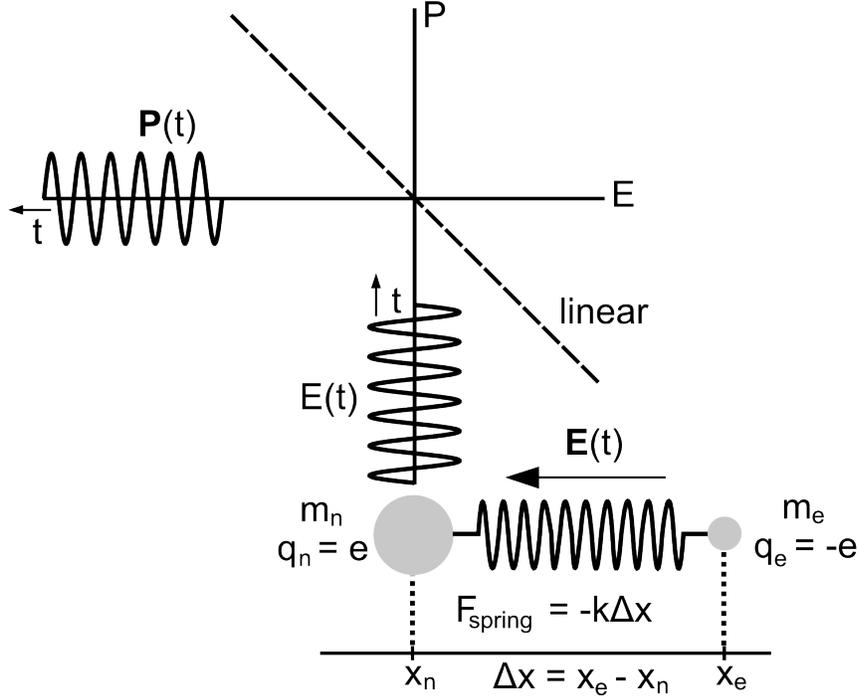


Figure 2.1: Schematic depiction of the Lorentz oscillator model and the linear polarization response to a time-varying electric field. Adapted from [2].

due to internal collisions and radiative losses, m_e^* is the effective mass of electrons, γ_0 is the oscillation damping coefficient, ω_0 is the resonant frequency of the system, and $e = 1.602 \times 10^{-19} \text{C}$ is the elementary charge. For a sufficiently small electron displacement, the excitation electric field, $\mathbf{E}(t) = \mathbf{E}_0 e^{-i\omega t}$, will drive an electron trajectory, $\mathbf{r}(t) = \mathbf{r}_0 e^{-i\omega t}$, where \mathbf{E}_0 is the electric field amplitude and \mathbf{r}_0 is the equilibrium position of the electron. Substituting this into Eq. 2.30 produces:

$$-m_e^* \omega^2 \mathbf{r}_0 e^{-i\omega t} - i m_e^* \gamma_0 \dot{\mathbf{r}}_0 e^{-i\omega t} + m_e^* \omega_0^2 \mathbf{r}_0 e^{-i\omega t} = -e \mathbf{E}_0 e^{-i\omega t}, \quad (2.31)$$

and rearranging yields:

$$\mathbf{r}_0(\omega) = -\frac{e}{m_e^* (\omega_0^2 - \omega^2 - i\gamma_0 \omega)} \mathbf{E}_0 \quad (2.32)$$

For a bulk material with N electrons per unit volume, the polarization can be written as, $\mathbf{P} = -N e \mathbf{r}_0$. Solving for \mathbf{r}_0 and substituting this into Eq. 2.32 produces:

$$\mathbf{P}(\omega) = \left(\frac{N e^2}{m_e^*} \right) \left(\frac{1}{\omega_0^2 - \omega^2 - i\gamma_0 \omega} \right) \mathbf{E}(\omega). \quad (2.33)$$

In general, materials can exhibit multiple resonances at different frequencies. Since a linear system has been assumed, the principle of superposition is valid and the overall polarization density can be written as a summation of the appropriate Lorentz oscillator

terms:

$$\mathbf{P}(\omega) = \left(\frac{Ne^2}{m_e^*} \right) \sum_j \frac{a_j}{\omega_j^2 - \omega^2 - i\gamma_j} \mathbf{E}(\omega), \quad (2.34)$$

where ω_j is the resonant frequency of a particular resonance, γ_j is the respective damping coefficient, and a_j is the relative strength of the resonance. Notably, this expression has the same form as Eq. 2.18, and the electric susceptibility can be written as:

$$\chi_e = \frac{Ne^2}{\varepsilon_0 m_e^*} \sum_j \frac{a_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega}. \quad (2.35)$$

Finally, the frequency-dependent permittivity can be determined via Eq. 2.23:

$$\varepsilon(\omega) = \varepsilon_0 \left(1 + \sum_j \frac{a_j \omega_p^2}{\omega_j^2 - \omega^2 - i\gamma_j \omega} \right), \quad (2.36)$$

where $\omega_p = \sqrt{\frac{Ne^2}{\varepsilon_0 m_e^*}}$ is the plasma frequency, which is the natural collective oscillation frequency of the electrons.

2.2.2 Electromagnetic Response of Metals

In the case of dielectrics, light matter interaction is determined primarily by the valence electrons. However, for metals, a high density ($\sim 10^{22} \text{cm}^{-3}$) of conduction electrons dominates the interaction. In contrast to the spring-like force that was used to model the bound valence electrons of a dielectric, the conduction electrons of a metal are free. Therefore, the restoring force term in Eq. 2.29, \mathbf{F}_{spring} , can be neglected, and Eq. 2.30 can be rewritten as:

$$m_e^* \frac{d^2 \mathbf{r}}{dt^2} + m_e^* \gamma_0 \frac{d\mathbf{r}}{dt} = -e\mathbf{E}. \quad (2.37)$$

Following a similar procedure as for dielectrics, the permittivity function of the metal can be derived as:

$$\varepsilon(\omega) = \varepsilon_0 \left(1 - \frac{a_0 \omega_p^2}{\omega^2 + i\gamma_0 \omega} \right). \quad (2.38)$$

While this expression is often sufficient for modelling the interaction of light with metals, there are exceptions. Resonant bound electron excitations can exist at ultraviolet, visible, or near-infrared wavelengths, depending on the metal, and can introduce significant deviations from this relation. Both the bound and conduction electrons of the metal can be modelled by combining the Drude model with the Lorentz oscillator model as follows:

$$\varepsilon(\omega) = \varepsilon_0 \left(1 - \frac{a_0 \omega_p^2}{\omega^2 + i\gamma_0 \omega} + \sum_j \frac{a_j \omega_p^2}{\omega_j^2 - \omega^2 - i\gamma_j \omega} \right). \quad (2.39)$$

By including a Lorentz oscillator term for each resonant excitation of the metal, an overall electromagnetic response can be obtained.

2.3 Nonlinear Response of Materials

As the amplitude of the exciting electric field approaches 10^9V/m , it becomes comparable to the amplitude of the electric fields binding electrons to nuclei ($\sim 3 \times 10^{10}\text{V/m}$), and the electron displacement becomes too large for the restoring force to be accurately modelled by Hooke's law. This is depicted schematically in Figure 2.2. In contrast to Figure 2.1, the polarization (or, similarly, the electron displacement) no longer follows the electric field in a linear fashion (shown as the dashed line), and instead they are related by a nonlinear function, depicted by the dotted line. When operating in this regime, simple principles of electromagnetics, including wave propagation, reflection, refraction, and superposition become more complicated. Deviation from a linear restoring force also gives rise to the generation of new frequency components and nonlinear loss mechanisms.

The Nonlinear Polarization

Modelling the nonlinear interaction of light with a medium requires modifications to Maxwell's equations, making them nonlinear partial differential equations. The most intuitive way to do this is through the polarization vector. Equation 2.18 described a linear relation between the susceptibility and the polarization, which can equivalently be written as $\mathbf{P}^{(1)}(t) = \chi^{(1)}\mathbf{E}(t)$, where $\mathbf{P}^{(1)}(t)$ is the linear polarization and $\chi^{(1)}$ is the linear susceptibility. Generalizing to a nonlinear susceptibility, additional polarization terms can be added [2]:

$$\mathbf{P}(t) = \mathbf{P}^{(1)}(t) + \mathbf{P}^{(NL)} \quad (2.40)$$

$$\mathbf{P}(t) = \mathbf{P}^{(1)}(t) + \mathbf{P}^{(2)}(t) + \mathbf{P}^{(3)}(t) + \dots \quad (2.41)$$

$$\mathbf{P}(t) = \chi^{(1)}\mathbf{E}(t) + \chi^{(2)}\mathbf{E}^2(t) + \chi^{(3)}\mathbf{E}^3(t) + \dots \quad (2.42)$$

where $\mathbf{P}^{(n)}$ is the n^{th} -order polarization vector and $\chi^{(n)}$ is the n^{th} -order nonlinear susceptibility. In general, $\chi^{(1)}$ is a second-rank tensor, $\chi^{(2)}$ is a third-rank tensor, $\chi^{(3)}$ is a fourth-rank tensor, and so on. The real component of $\chi^{(1)}$ is related to the refractive index, whereas the imaginary is related to loss or gain properties of the medium. In general, the nonlinear susceptibilities are a function of the excitation wavelength(s) and resonant nonlinearities can be observed. In the case of a bulk, centrosymmetric crystal, all even polarization terms, ($\chi^{(2)}, \chi^{(4)}, \dots$) are identically zero.

The second-order term, $\chi^{(2)}$, gives rise to mixing between two waves that overlap in space and time and polarize the same system. In the case of monochromatic light, second-harmonic generation (SHG) and rectification can be observed. In the presence of two or

more frequencies, sum frequency generation (SFG) and difference frequency generation (DFG) occur. If one of these frequencies is zero (a DC field), the nonlinearity is known as the electro-optic effect.

The third-order term, $\chi^{(3)}$, produces third-harmonic generation (THG), third-order frequency mixing, the optical Kerr effect, and two photon absorption (TPA). Notably, the optical Kerr effect is responsible for other nonlinear effects, including self- and cross-phase modulation and four wave mixing.

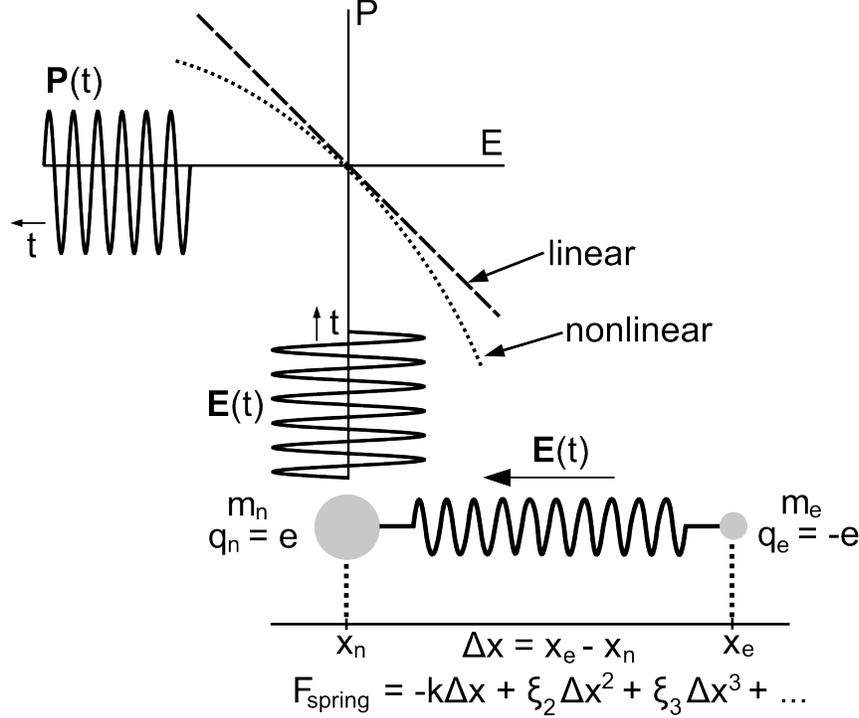


Figure 2.2: Schematic depiction of the Lorentz oscillator model and the nonlinear polarization response to a time-varying electric field. Adapted from [2].

To facilitate a systematic discussion of nonlinear optical effects, the notation of Butcher and Cotter is adopted and will be used throughout the remainder of this thesis. In this notation, the nonlinear polarization can be expressed as [2]:

$$\mathbf{P}_{\omega_\sigma}^{(n)} = \varepsilon_0 \sum_{\omega} K(-\omega_\sigma; \omega_1, \dots, \omega_n) \chi^{(n)}(-\omega_\sigma; \omega_1, \dots, \omega_n) \mathbf{E}_{\omega_1} \cdots \mathbf{E}_{\omega_n}, \quad (2.43)$$

where $K(-\omega_\sigma; \omega_1, \dots, \omega_n)$ is the degeneracy factor and the summation accounts for all distinct sets of $\omega_1, \omega_2, \dots, \omega_n$. The term in parenthesis, i.e. $(-\omega_\sigma; \omega_1, \dots, \omega_n)$, indicates that n photons with frequencies, $\omega_1, \dots, \omega_n$ interact in the nonlinear process to produce a photon of frequency, ω_σ , traveling in the same direction. For example, THG is represented as $\chi^{(3)}(-3\omega; \omega, \omega, \omega)$, which indicates that three photons of the same frequency interact in the nonlinear process to produce a photon with triple the frequency, traveling in the same direction. This summation is also relevant when considering nonlinearities excited by femtosecond pulses consisting of a broad spectrum of frequencies. The degeneracy

factor, K , is defined as:

$$K(-\omega_\sigma; \omega_1, \dots, \omega_n) = 2^{l+m-n} p \quad (2.44)$$

where $l=1$ if $\omega_\sigma \neq 0$ (i.e. optical rectification) and zero otherwise, m is the number of exciting frequencies that are static (DC), n is the order of the nonlinearity, and p is the number of distinct permutations of $\omega_1, \dots, \omega_n$. Oftentimes, it is not necessary to employ the full notation of Eq. 2.43 and a simpler version is more intuitive:

$$P_\sigma^{(n)} = \varepsilon_0 \sum_\omega K(-\omega_\sigma; \omega_1, \dots, \omega_n) \chi^{(n)}(-\omega_\sigma; \omega_1, \dots, \omega_n) E_1 \cdots E_n \quad (2.45)$$

The nonlinear terms in Maxwell's equations can then be accounted for by rewriting Eq. 2.16 as [2]:

$$\mathbf{D} = \varepsilon \mathbf{E} = \varepsilon_0 (\mathbf{I} + \chi^{(1)}) \mathbf{E} + \mathbf{P}^{(NL)}, \quad (2.46)$$

where \mathbf{I} is the identity tensor. Applying this relation to Eq. 2.26, the nonlinear wave equation is obtained:

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2} \quad (2.47)$$

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \left(\frac{\partial^2 \mathbf{P}^{(1)}}{\partial t^2} + \frac{\partial^2 \mathbf{P}^{(NL)}}{\partial t^2} \right) \quad (2.48)$$

A general collection of nonlinear optical effects is presented in Table 2.1. This thesis focuses on nonlinear optical interactions in Si-based nanophotonic and nanoplasmonic structures and the relevant nonlinear effects in Si nanophotonic waveguides are discussed in the subsequent sections.

2.4 Silicon-on-insulator Waveguides

High index contrast Si waveguides have a number of characteristics that make them attractive for complex integrated optical circuits that can be used as modular components in fibre optical networks or in hybrid electrical-optical circuitry [3]. Firstly, SOI waveguides are virtually lossless at telecommunication wavelengths ($1260\text{nm} \leq \lambda \leq 1675\text{nm}$) and propagation losses of 3.6dB/cm have been achieved [4]. The high refractive index of Si, $n = 3.48$, at $\lambda = 1550\text{nm}$ [5] allows for compact cross-sectional dimensions of the waveguide core, typically in the vicinity of width \times height = $w \times h = 400\text{nm} \times 220\text{nm}$. Furthermore, Si has a relatively high second-order nonlinear refractive index, $n_2 = 3.2 \times 10^{-14} \text{cm}^2/\text{W}$ [6], or equivalently, a high $\chi^{(3)}$ nonlinear coefficient, $\chi^{(3)} = 8n_0 n_2 / 3 = 2.97 \times 10^{-15} \text{cm}^2/\text{W}$, making SOI waveguides useful as modulators or for frequency mixing [2]. Finally, the growth and processing techniques of Si are well-established and high-quality, standardized SOI wafers are readily available. All of these characteristics have enabled the rapid development of Si photonic devices over the last two decades [7]. Furthermore, hybrid

Process	Order	$-\omega_\sigma; \omega_1, \dots, \omega_n$	K
Linear refractive index,			
absorption/gain	1	$-\omega; \omega$	1
Second harmonic generation	2	$-2\omega; \omega, \omega$	$\frac{1}{2}$
Optical rectification	2	$0; \omega, -\omega$	$\frac{1}{2}$
Sum/Difference Frequency			
Mixing, parametric amplification			
and oscillation	2	$-\omega_3; \omega_1, \pm\omega_2$	1
Pockels effect (linear			
electro-optics effect)	2	$-\omega; 0, \omega$	2
Third harmonic generation	3	$-3\omega; \omega, \omega, \omega$	$\frac{1}{4}$
Third-order sum/difference-			
frequency mixing	3	$-\omega_3; \pm\omega_1, \omega_2, \omega_2$	$\frac{3}{4}$
Optical Kerr effect, intensity-			
dependent refractive index	3	$0; \omega, -\omega, \omega$	$\frac{3}{4}$
Single Frequency Two-photon			
absorption	3	$-\omega; -\omega, \omega, \omega$	$\frac{3}{4}$
Two-frequency two-photon			
absorption	3	$-\omega_1; -\omega_2, \omega_2, \omega_1$	$\frac{3}{2}$
Four-wave mixing	3	$-\omega_4; \omega_1, \omega_2, \omega_3$	$\frac{3}{2}$
DC Kerr effect (quadratic			
electro-optic effect)	3	$-\omega; 0, 0, \omega,$	3
Dc-induced second-			
harmonic generation	3	$-2\omega; 0, \omega, \omega$	$\frac{3}{2}$
Anti-Stokes Raman scattering	3	$-\omega_{AS}; \omega_P, \omega_P, \omega_S$	$\frac{3}{4}$
Optical Kerr effect, stimulated			
Raman/Brillouin scattering	3	$-\omega_S; \omega_P, -\omega_P, \omega_S$	$\frac{3}{2}$
n th harmonic generation	n	$-n\omega; \omega, \dots, \omega$	2^{1-n}

Table 2.1: General Collection of Nonlinear Phenomena (adapted from [2])

electronic-photonic computer chips have attracted major industrial interest and are currently in advanced stages of research and development.

2.4.1 Passive Routing

Silicon-on-insulator waveguides are the main class of waveguide used in Si photonics and a vast array of devices have been demonstrated for routing, filtering, buffering, and modulating near-infrared electromagnetic signals. Single mode waveguides with cross-sectional dimensions on the order of $w \times h = 400\text{nm} \times 200 - 350\text{nm}$ are most commonly used, as they enable the smallest mode profile effective area, A_{eff} , which is important for high-density optical circuitry and nonlinear functionality. These waveguides support two fundamental modes: (1) the quasi-transverse-electric (quasi-TE) mode, where the primary polarization component is parallel to the substrate; and (2) the quasi-transverse-magnetic (quasi-TM) mode, where the primary polarization component is normal to the substrate. The quasi-TE mode profile for a waveguide with a height, $h = 340\text{nm}$, and a width, $w = 500\text{nm}$, is shown in Fig. 2.3(a), and the quasi-TM mode profile is shown in Fig. 2.3(b)¹. The quasi-TE mode profile for a waveguide with half the width, $w = 250\text{nm}$ is shown in Fig. 2.3(c), and the quasi-TM mode profile is shown in Fig. 2.3(d). Notably, as the waveguide dimensions decrease, the relative intensity of the mode components outside of the Si core become stronger. The effective refractive index of the quasi-TE and quasi-TM mode profiles in SOI waveguides with $h = 340\text{nm}$ are plotted as a function of w in Fig. 2.4. For $w \geq 700\text{nm}$, the effective refractive index of the quasi-TM mode, n_{eff}^{TM} , is approximately constant at $n_{eff}^{TM} \approx 2.62 - 2.76$. As w is decreased below 700nm , the n_{eff}^{TM} decreases rapidly, as the Si core is no longer able to confine the electromagnetic energy and a large fraction of the mode propagates outside of the Si core. For example, for the quasi-TM mode of $w = 500\text{nm}$ waveguide, 62.2% of the mode energy is confined to the Si core, whereas this number is reduced to 45.7% for the quasi-TM mode of the $w = 250\text{nm}$ waveguide. As such, the propagating energy also interacts optically with the SiO_2 substrate and air cladding and n_{eff}^{TM} decreases rapidly as more of the energy interacts with the lower-index materials. In a similar manner, the effective refractive index of the quasi-TE mode, n_{eff}^{TE} is relatively constant for $w \geq 700\text{nm}$, at $n_{eff}^{TE} = 2.92 - 3.09$, and for $w \leq 700\text{nm}$, n_{eff}^{TE} decreases significantly. For $w = 250\text{nm}$, values of $n_{eff}^{TE} = 1.33$ and $n_{eff}^{TM} = 2.03$ are obtained. For $w = 365\text{nm}$, the two modes have the same effective refractive index, $n_{eff}^{TE} = n_{eff}^{TM} = 2.33$ and they are said to be degenerate.

The fraction of electromagnetic power propagating through the Si core is plotted as a function of w for the quasi-TM mode profile in Fig. 2.5. This parameter follows essentially the same trend as the effective refractive indices. For $w \geq 700\text{nm}$, the core power fraction is relatively constant at 65.5%-68.2%. However, for $w \leq 700\text{nm}$, the core power fraction

¹The mode profiles in this thesis were obtained using Lumerical FDTD Solutions, where the waveguide geometry is meshed using the finite difference algorithm, a matrix eigenvalue problem based on Maxwell's equations is setup, and the matrix is solved using sparse matrix techniques.

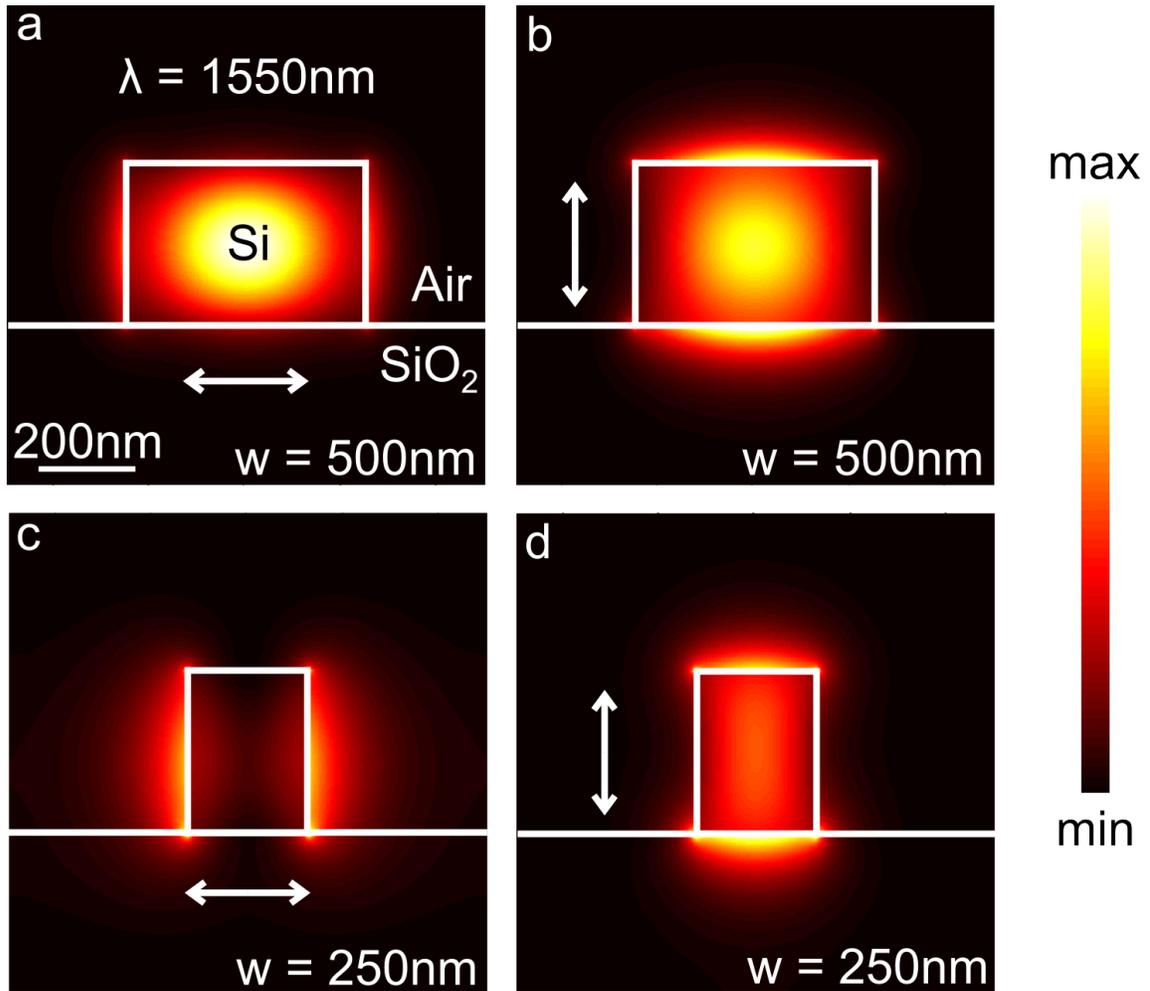


Figure 2.3: Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ in SOI waveguides: (a) Quasi-TE mode, $w = 500\text{nm}$. (b) Quasi-TM mode, $w = 500\text{nm}$. (c) Quasi-TE mode, $w = 250\text{nm}$. (d) Quasi-TM mode, $w = 250\text{nm}$.

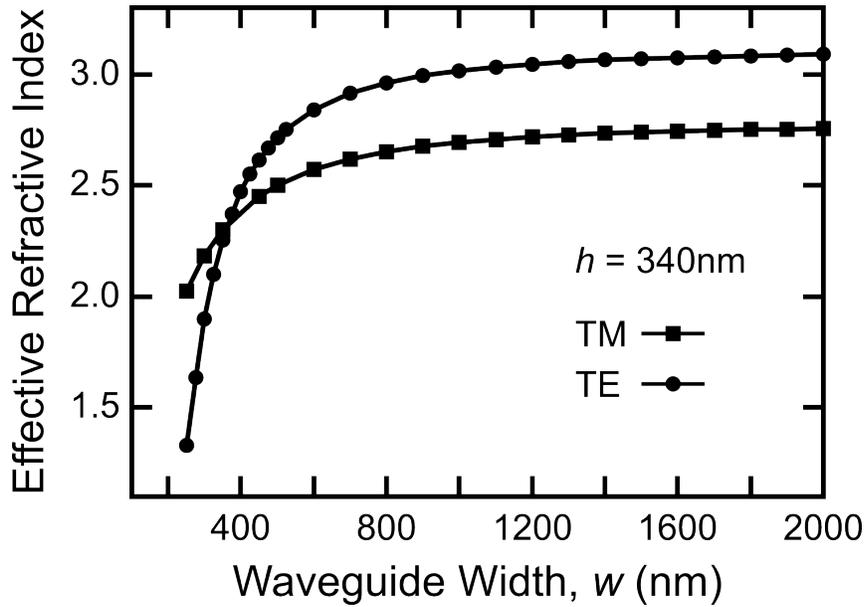


Figure 2.4: Transverse magnetic and transverse electric effective waveguide mode indices for a fixed waveguide core height, $h = 340\text{nm}$, and varying waveguide widths, w .

decreases rapidly, and for $w = 150\text{nm}$, it diminishes to a value of 19.3%. Further reduction of w would produce an effective refractive less than the surrounding environment ($n = 1$), and therefore, the Si core would no longer confine the mode. Such a mode is known as a “leaky” waveguide mode. Low confinement to the Si core is undesirable for two main reasons:

1. As the confinement decreases, the waveguide is more prone to cross-talk and the integration density must be kept lower. An extreme case of high-density optical circuitry is shown in Fig. 2.6. Figure 2.6(a) depicts two Si waveguides with $w = 250\text{nm}$ that are separated by a distance of 100nm . This geometry is simulated via 2D finite-difference time-domain (FDTD) simulations, and the resulting intensity distribution is shown in Fig. 2.6(b). When a mode is excited in Waveguide 1 (on the right-hand side), the mode energy easily couples to Waveguide 2 (on the left-hand side), and after propagating a distance of $3.41\mu\text{m}$, virtually all of the mode energy propagates in Waveguide 2. It then couples back to Waveguide 1, and this periodic exchange of energy continues along the length of the waveguides. This coupling length is virtually identical to the value calculated with coupled mode theory [8]:

$$L_c = \frac{\lambda}{2(N_s - N_a)} \quad (2.49)$$

$$= \frac{1.55\mu\text{m}}{2(3.021 - 2.788)} \quad (2.50)$$

$$= 3.33\mu\text{m}, \quad (2.51)$$

where N_s is the effective refractive index of the symmetric mode and N_a is the effective refractive index of the asymmetric mode. Obviously, such cross-talk is not desirable when a high signal-to-noise ratio must be maintained.

2. It becomes extremely challenging to design a waveguide that has very small core dimensions (enabling a high optical intensity) that effectively confines a large fraction of the energy to the core, which is desirable for developing photonic circuitry with nonlinear functionality.

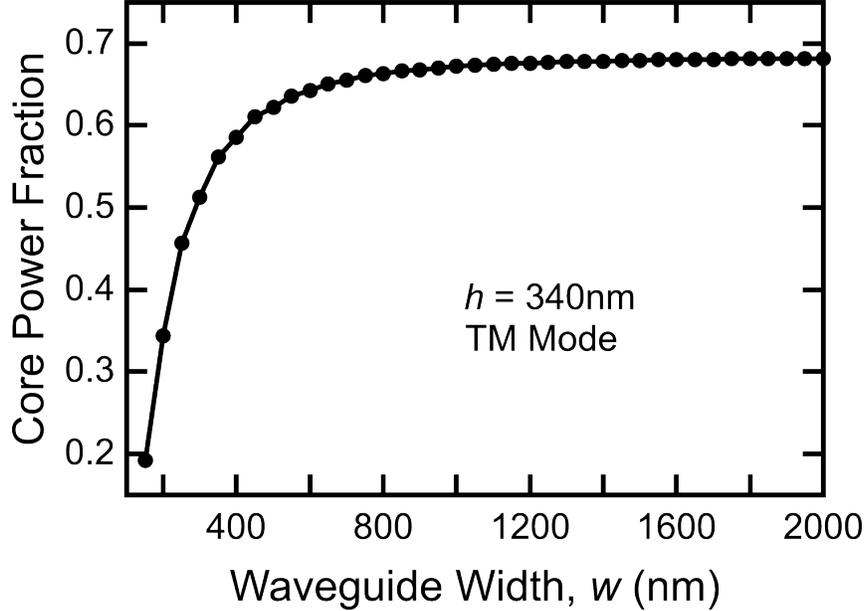


Figure 2.5: Fraction of the waveguide mode power that is confined to the Si core, for waveguides with a constant height, $h = 340\text{nm}$, and varying widths, w .

2.4.2 Nonlinear Optical Effects in Silicon-on-Insulator Waveguides

Second-order Nonlinearities

As a centrosymmetric semiconductor, bulk Si has vanishing even-order nonlinear coefficients, i.e. $\chi^{(2)}, \chi^{(4)}$, etc. Regardless, it is possible to excite second-order nonlinearities at the surfaces of Si structures. Silicon photonic crystal nanocavities have been used to focus incident electromagnetic waves and generate both second and third harmonics [9]. In this case, generation of the second-harmonic is made possible by substantial optical power near the Si-air interfaces where the centrosymmetry is broken. Alternatively, a straining layer can be deposited on top of a device, introducing asymmetry to the lattice near the straining layer, and breaking the centrosymmetry [10, 11]. Such straining layers have been used to induce a second-order nonlinear coefficient of $\chi^{(2)} \approx 15\text{pmV}^{-1}$ in SOI waveguides, and SHG efficiencies of $\eta_{SHG} = 5 \times 10^{-8}$ have been demonstrated [12].

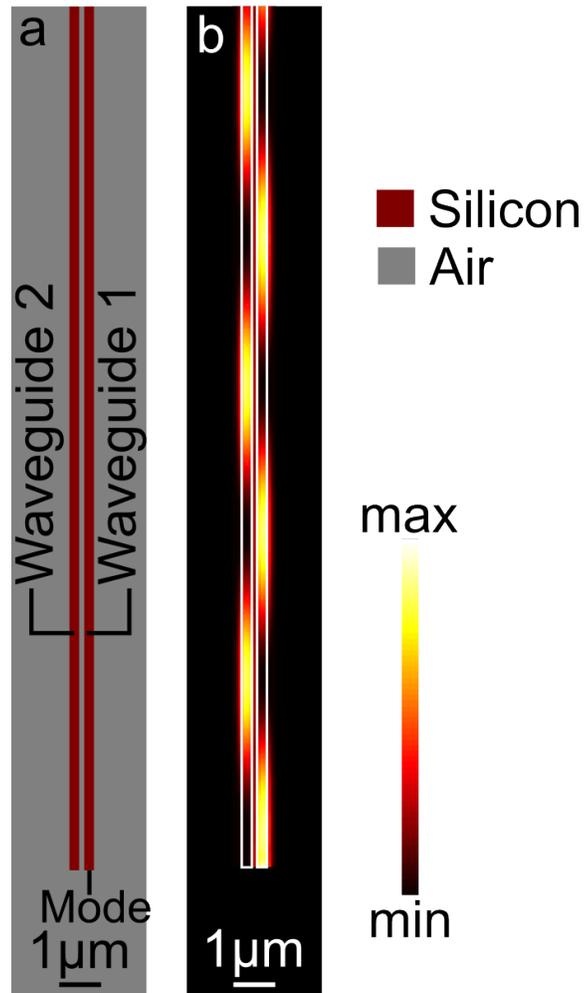


Figure 2.6: (a) Top-view schematic of two SOI photonic waveguides with $w = 250\text{nm}$, separated by a 100nm gap. (b) Finite-difference time-domain intensity (I) distributions of the waveguides depicted in (a), demonstrating the periodic coupling of mode energy between adjacent waveguides.

Although this is an interesting approach, the introduction of a straining layer places a constraint on device geometry design and was not desirable for this thesis. In particular, an extra straining layer would introduce significant fabrication challenges and may not function as intended, due to additional plasmonic films deposited on top of the straining layer. The SHG process is shown schematically in Fig. 2.7.

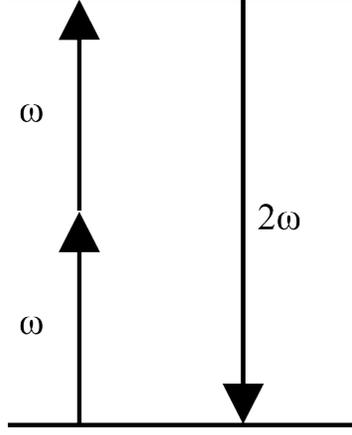


Figure 2.7: Schematic depiction of second harmonic generation. Here, two photons of the fundamental frequency, ω , combine to create a single photon with double the frequency, 2ω .

Third-order Nonlinear Effects in Silicon

In the special case of monochromatic light, two third-order polarization effects are present in Si. In the case of linear polarization, these can be expressed as [13]:

$$P^{(3)}(\omega) = \varepsilon_0 \frac{3}{4} \chi^{(3)}(-\omega; \omega, -\omega, \omega) |E_\omega|^2 E_\omega, \quad (2.52)$$

$$P^{(3)}(3\omega) = \varepsilon_0 \frac{1}{4} \chi^{(3)}(-3\omega; \omega, \omega, \omega) E_\omega^3. \quad (2.53)$$

Equation 2.52 describes degenerate two-photon absorption and the optical Kerr effect via its imaginary and real components, respectively, and Eq. 2.53 portrays third-harmonic generation (THG). Further details of these effects are presented in the subsequent sections.

The Optical Kerr Effect

The optical Kerr effect is related to the real component of $\chi^{(3)}$ and describes the intensity-dependent refractive index, n_2 . In specific, n_2 is related to $\chi^{(3)}$ via [2]:

$$n_2 = \frac{3\chi^{(3)}}{8n_0}. \quad (2.54)$$

If the intensity of electromagnetic energy in a particular region of material is I , then the refractive index in that region will be:

$$n(I) = n_0 + n_2 I. \quad (2.55)$$

Experimental values of n_2 in Si vary from 4×10^{-14} to $9 \times 10^{-14} \text{cm}^2 \text{W}^{-1}$ [14, 15]. The optical Kerr effect gives rise to several subordinate effects: self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM).

Self-phase Modulation

As outlined above, the Kerr nonlinearity creates an intensity-dependent refractive index. As an intense wave propagates a distance, L , through the nonlinear waveguide, it incurs a phase-shift, given by:

$$\phi = -n(I)k_0 L = -\frac{2\pi(n + n_2 I)L}{\lambda_0}. \quad (2.56)$$

Therefore, the phase change can be calculated:

$$\Delta\phi = -\frac{2\pi n_2 I L}{\lambda_0}. \quad (2.57)$$

When used in a phase-sensitive geometry such as a Mach-Zehnder interferometer, this effect can be used for phase-sensitive signal modulation.

There are further consequences of SPM if an ultrashort pulse is propagated through a waveguide, as the pulse will induce a time-dependent change in the effective mode index of the waveguide. This will produce a shift in the instantaneous phase of the pulse, which results in frequency shifting. The temporal front of the pulse will experience an increasing refractive index, which corresponds to a red shift, whereas the temporal back of the pulse experiences a decreasing refractive index, which corresponds to a blue shift. The further the pulse propagates, the more it broadens, until its peak intensity diminishes to an amplitude where the nonlinear contribution is not significant. The SPM process is depicted schematically in Fig. 2.8, where $\Delta\omega$ denotes the frequency shift arising from SPM. Due to the combination of long length and tight modal confinement, this effect can become quite pronounced in Si waveguides [16–21], and is typically a precursor to supercontinuum generation.

Cross-phase Modulation

Cross-phase modulation is similar to SPM except that the Kerr-induced phase shift from a strong pump pulse is used to broaden a second, possibly weaker pulse [22–24], which can have a different frequency spectrum. Notably, by controlling the relative time delay between the pump and probe, it is possible to blue-shift, red-shift, or symmetrically

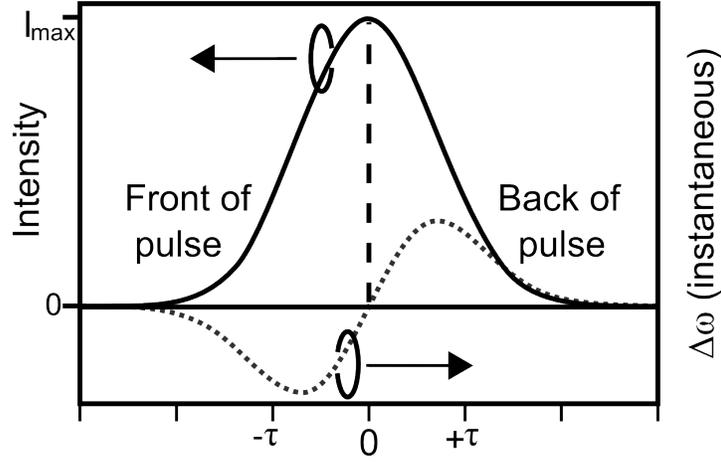


Figure 2.8: Graph depicting the concept of self-phase modulation. The temporal front of the pulse produces an increasing refractive index (via the optical Kerr effect), which corresponds to a red-shift in the frequency components. The temporal back of the pulse produces a decreasing refractive index, which corresponds to a blue-shift in the frequency components. $\Delta\omega$ represents the frequency shift of the spectral components of the pulse due to SPM.

broaden the spectrum of the probe (i.e. a second pulse).

2.4.3 Four-wave Mixing

Four-wave mixing can occur in a Si waveguide when at least two frequency components are present, ω_1 and ω_2 [15, 25–28]. In the case when only two frequencies are present, two new frequencies are generated, $\omega_3 = \omega_1 - (\omega_2 - \omega_1) = 2\omega_1 - \omega_2$ and $\omega_4 = \omega_2 + (\omega_2 - \omega_1) = 2\omega_2 - \omega_1$. A schematic depiction of the case when $\omega_3 = 2\omega_1 - \omega_2$ is shown in Fig. 2.9. To describe FWM, Eq. 2.52 can be rewritten as:

$$P^{(3)}(2\omega_1 - \omega_2) = \varepsilon_0 \frac{3}{4} \chi^{(3)}(-2\omega_1 + \omega_2; \omega_1, \omega_1, -\omega_2) E_{\omega_1}^2 E_{\omega_2}^*. \quad (2.58)$$

Notably, FWM is a phase-sensitive process, and therefore it is easiest to observe when the involved frequencies are close together, so that their waveguide effective mode indices are approximately equal and they can interact over long distances.

Two-photon Absorption (TPA)

Two-photon absorption is a resonant nonlinear loss mechanism in Si. Whereas the optical Kerr effect introduces a second-order nonlinear refractive index contribution (i.e. $n = n_0 + n_2 I$), TPA introduces a second-order nonlinear absorption contribution, i.e. $\alpha = \alpha_{lin} + \beta_{TPA} I$, where α_{lin} is the linear absorption and β_{TPA} is the TPA coefficient. Reported values of β_{TPA} range from 4×10^{-12} to $9 \times 10^{-12} \text{W}^{-1} \text{m}$ [14, 16, 29–31].

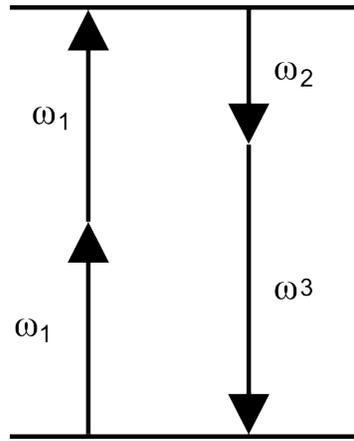


Figure 2.9: Schematic depiction of four-wave mixing, where a photon is produced at $\omega_3 = 2\omega_1 - \omega_2$.

The photon energy corresponding to $\lambda = 1550\text{nm}$ is calculated as follows:

$$E = \frac{hc}{\lambda} = \frac{(4.136 \times 10^{-15} \text{eV} \cdot \text{s})(2.998 \times 10^8 \text{m/s})}{1550 \times 10^{-9} \text{m}} = 0.80 \text{eV} \quad (2.59)$$

Silicon has an indirect bandgap of 1.12 eV and a direct bandgap of 3.40 eV. Thus, TPA is possible via the indirect bandgap at $\lambda = 1550\text{nm}$ and direct TPA will not be possible at this wavelength. A phonon must assist the two photons to conserve momentum in the indirect transition, which is shown schematically in Figure 2.10.

When a pulse propagates through a SOI waveguide, power will be absorbed via TPA and the pulse intensity will diminish along the propagation direction. The TPA-induced signal attenuation depends on the local intensity and the distance propagated as follows:

$$\frac{dI}{dz} = -\beta_{TPA} I^2. \quad (2.60)$$

When free-carriers are excited via TPA, the local optical properties change. Free-carrier absorption (FCA) causes further losses and free-carrier dispersion (FCD) can change the instantaneous phase of signals interacting with a free-carrier plasma. These are discussed in the next section.

Free-carrier Absorption (FCA) and Free-carrier Plasma Dispersion (FCD)

Once electrons have been excited to the conduction band via TPA, an additional loss mechanism can take place. Subsequent incident photons can be absorbed by conduction-band electrons and excited to a higher conduction band energy state, which is known as FCA. This process is depicted schematically in Figure 2.10. As a pulse propagates through a SOI waveguide and excites free-carriers via TPA, the remainder (i.e. the temporally

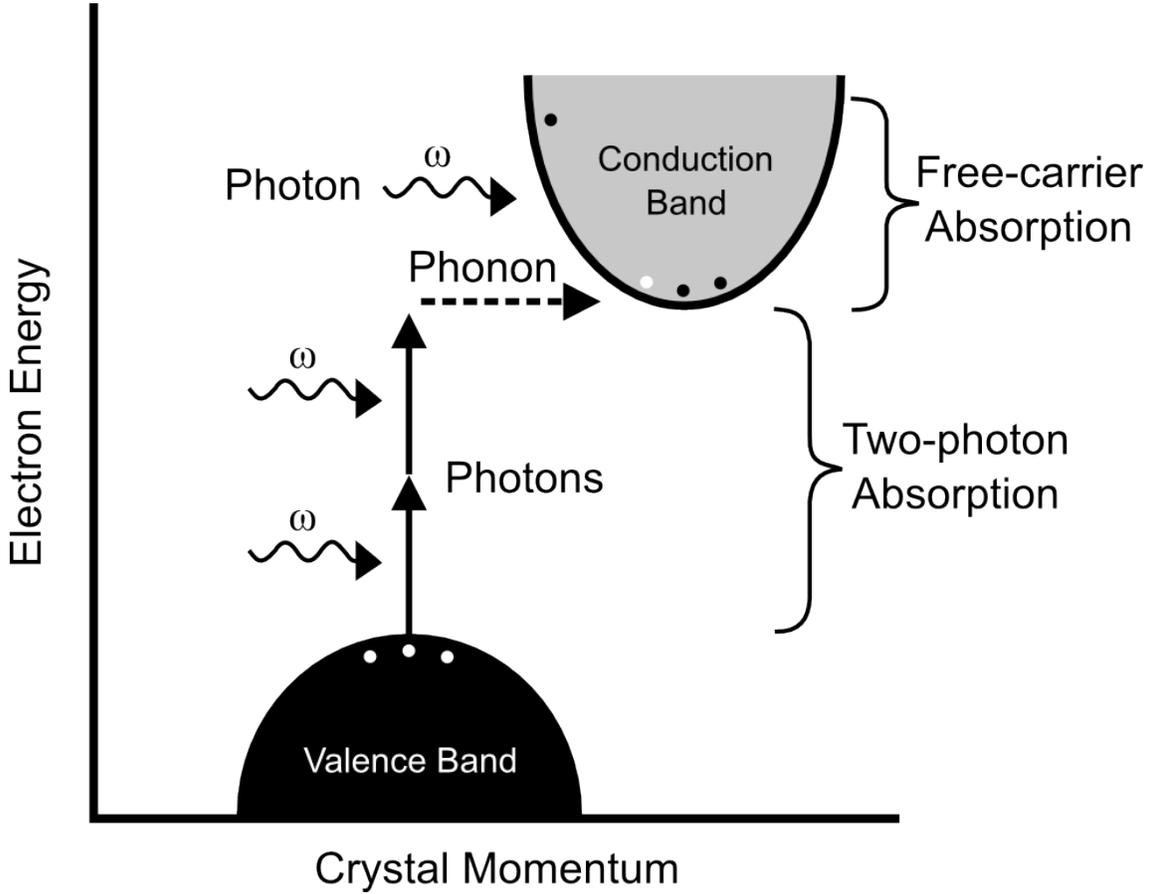


Figure 2.10: Schematic depiction of the band structure of Si, two-photon absorption and subsequent free-carrier absorption.

trailing portion) of the pulse can undergo absorption due to FCA. This is expressed as:

$$\frac{dI}{dz} = -\sigma_{FCA}NI, \quad (2.61)$$

where σ_{FCA} is the FCA cross-section and N is the free-carrier density. Typical measured values for σ_{FCA} are in the vicinity of $\sigma_{FCA} = 1.45 \times 10^{-17} \text{cm}^2$ [32]. For a pulse of sufficient intensity, a combination of TPA and FCA will attenuate the pulse until its peak intensity diminishes below the threshold for observing the nonlinear interaction. As an absorption mechanism, FCA introduces a change to the imaginary component of the refractive index, $n_0 = n + ik$, via [10]:

$$\Delta k = \frac{e^3 \lambda^3}{16\pi^3 c^3 \epsilon_0 n} \left[\frac{\Delta N_e}{(m_e^*)^2 \mu_e} + \frac{\Delta N_h}{(m_h^*)^2 \mu_h} \right], \quad (2.62)$$

where n is the background refractive index, $m_e^* = 0.26m_0$ and $m_h^* = 0.39m_0$ are the conductivity effective masses, $m_0 = 9.11 \times 10^{-31} \text{kg}$ is the free-electron rest mass, $\mu_e = 1400 \text{cm}^2 \text{V}^{-1} \text{s}^{-1}$ is the mobility of electrons, $\mu_h = 450 \text{cm}^2 \text{V}^{-1} \text{s}^{-1}$ is the mobility of holes, ΔN_e is the number density of electrons, and ΔN_h is the number density of holes.

The generation of a free-carrier plasma will also introduce subsequent dispersive contributions to the material. The change in the real component of the refractive index can be expressed as:

$$\Delta n = -\frac{e^2 \lambda^2}{8\pi^2 c^2 \epsilon_0 n} \left[\frac{\Delta N_e}{m_e} + \frac{\Delta N_h}{m_h} \right]. \quad (2.63)$$

Notably, the introduction of a large free-carrier population will give rise to a large decrease in the refractive index. Therefore, free-carriers excited by the peak of a pulse will produce a rapidly decreasing refractive index for the temporal back of the pulse, resulting in a blue-shift to the spectral power of that portion of the pulse.

Third-harmonic Generation (THG)

Third-harmonic generation occurs when three photons interact with the valence electrons of an atom and emit one photon that carries three times the energy. This process is shown schematically in Fig. 2.11. For telecommunication wavelengths ($1260\text{nm} \leq \lambda \leq 1675\text{nm}$), the third-harmonic (TH) lies in the violet-blue-green ($420\text{nm} \leq \lambda \leq 558\text{nm}$) region of the visible spectrum, which is absorbed strongly in Si. The strong absorption of visible light in Si means that free-carriers will be excited, and FCA due to the absorbed TH photons will attenuate the signal in the same manner as TPA-generated free-carriers. Furthermore, Si is strongly dispersive in the visible-NIR spectral region. Therefore, it is not possible to accommodate phasematched THG in Si waveguides, and THG must come from predominantly localized interactions. Despite being studied in bulk Si several decades ago, THG has only been observed once in a SOI waveguide [33]. This is a topic that is studied extensively in this thesis and will be presented in more detail in Chapters 5-6.

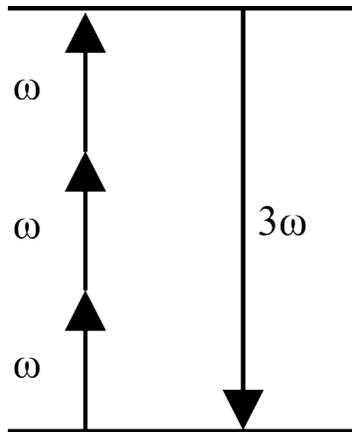


Figure 2.11: Schematic representation of THG. Here, three photons at the fundamental frequency, ω , combine to generate a single photon at triple the frequency, 3ω .

2.4.4 Free-carrier Dynamics in Silicon

As stated earlier, free-carriers can be excited in SOI waveguides via TPA. These free-carriers recombine through a number of different pathways which, in general, can be categorized as *intrinsic* or *extrinsic* [34].

Intrinsic Recombination

Intrinsic recombination is not dependent on crystal growth technique or doping; it is present in a perfect crystal. Two distinct intrinsic recombination processes can be observed: (1) radiative band-band recombination, where energy is converted to a photon; and (2) Auger band-band recombination, which is non-radiative and energy is transferred to a third carrier. Due to the indirect bandgap of Si, radiative band-band recombination occurs with a very low probability, and is not a prominent recombination mechanism.

Extrinsic Recombination

Extrinsic recombination is known as Shockley-Read-Hall (SRH) recombination and relies on defect states at intermediate energy levels in the band gap. It should be noted that three main processes can occur at defect states: (1) recombination of electron-hole pairs; (2) generation of electron-hole pairs; and (3) trapping, where a free-carrier is captured at the defect centre, and subsequently released to its original band.

Recombination at defect centres can occur in two manners: (1) a conduction band electron and a valence band hole are captured at the defect energy level and annihilate one another; or (2) the conduction band electron first transitions to the defect energy level and then to the conduction band, where it annihilates a hole. Shockley-Read-Hall recombination can occur through naturally occurring lattice defects, impurities, or surface states, which result from dangling bonds at the wafer surface.

The interface of a Si wafer is host to a large density of partially-bonded Si atoms. These dangling bonds give rise to a continuous distribution of interface states throughout the bandgap. The surface recombination velocity depends strongly on how the surface of the wafer has been prepared. For example, surface passivation with a thermally grown SiO₂ or SiN_x layer greatly reduces the density of dangling bonds and minimizes the surface recombination velocity, making only bulk recombination significant.

In general, the recombination time can be expressed as a parallel combination of all of these recombination mechanisms, as follows:

$$\tau_r^{-1} = (\tau_{SRH}^{-1} + \tau_{Auger}^{-1} + \tau_{rad}^{-1}) + \tau_{surface}^{-1} \quad (2.64)$$

$$= \tau_{bulk}^{-1} + \tau_{surface}^{-1}. \quad (2.65)$$

Here, τ_r is the effective recombination time, τ_{SRH} is the SRH recombination time, τ_{Auger} is the Auger recombination time, τ_{rad} is the radiative recombination time, τ_{bulk} is the

effective bulk recombination time, and $\tau_{surface}$ is the effective surface recombination time. Although the details of each recombination mechanism are intricate, here the main interest is in the the time-dynamics that arise from τ_r .

In general, free-carriers can be excited via TPA in response to an intensity excitation with an arbitrary temporal waveform, $I(t)$, and generated at a rate that is proportional to $I^2(t) \cdot \beta_{TPA}/(2h\nu)$. Furthermore, it can be assumed that free-carriers recombine with a time-constant, τ_r . Therefore, the rate equation can be written as:

$$\frac{dN(t)}{dt} = \frac{\beta_{TPA}}{2h\nu} I^2(t) - \frac{N(t)}{\tau_r}. \quad (2.66)$$

For the case of a continuous-wave intensity input, the steady-state free-carrier density reduces to:

$$N = \frac{\beta_{TPA} I^2 \tau_r}{2h\nu}. \quad (2.67)$$

For switching schemes that rely on free-carrier generation, the electron-hole recombination time is an important parameter that limits the modulation bandwidth. The free-carriers can be generated very quickly, on a timescale that is limited mainly by the duration of the pump pulse, which can be $\tau_p \leq 100\text{fs}$. Thus, modulation achieved by the introduction of free-carriers can occur very quickly. However, in order for the device to recover to its original state, the electron-hole pairs must recombine. Depending on the surface passivation conditions, the recombination time is typically on the order of $100\text{ns} \leq \tau_r \leq 100\mu\text{s}$ in bulk Si. When considering Si nanophotonic devices, the recombination time is reduced considerably and is typically in the range of $100\text{ps} \leq \tau_r \leq 10\text{ns}$. In the case of a Si photonic waveguide with cross-sectional dimensions, $w \times h = 400\text{nm} \times 220\text{nm}$, the excited free-carriers rapidly diffuse at $36\text{cm}^2/\text{s}$ [35] and can become uniformly distributed across the waveguide cross-section on a timescale of approximately $\Delta t = 20\text{ps}$. Therefore, a majority of the free-carriers are able to interact with the waveguide sidewalls and interfaces on a picosecond timescale, where a high surface recombination velocity facilitates much faster overall recombination.

A number of processing techniques have been employed to reduce the recombination time further, such as using porous Si [36], using polymorphous Si [37], ion-implanting the Si with helium [38], oxygen [39], argon [40], or Si [41], diffusing species into the Si, or sweeping free-carriers out of the device by integrating a p-i-n diode across the photonic structure and applying a reverse bias, generating a high DC field [42]. Some typical recombination times are summarized in Table 2.2.

2.5 Surface Plasmon Polariton (SPP) Waves

Fundamentally, a SPP is a quasiparticle that can be excited at the interface between two materials (typically a dielectric and a metal) and consists of a photon coupled to an oscillation of charge density co-propagating along the interface [45]. This excitation

Sample	Recombination Time	References
Bulk Si	100ns-100 μ s	[34]
Nanophotonic Waveguides	450ps-3ns	[42, 43]
Porous Si	200ps	[36]
Poly Si	135ps	[37]
P-I-N diode sweeping	12.2ps	[42]
Ion-Implanted Si	600fs-70ps	[38–41, 44]

Table 2.2: Recombination times of various Si samples

occurs when an incident electromagnetic wave (free to propagate in three dimensions) becomes coupled to the conduction electrons of the metal and is effectively confined to a two-dimensional surface mode.

2.5.1 Planar: One-dimensional Confinement

The simplest structure that supports surface plasmon modes is the planar interface between a metal and a dielectric. The SPP excitation can be understood quantitatively by invoking Maxwell's equations at the interface. The geometry depicting SPP propagation at a metal-air interface is shown in Fig. 2.12, where the interface is situated at $z = 0$. The region located at $z < 0$ is metal, and will be denoted by the subscript, m , whereas the region located at $z > 0$ is dielectric, and will be denoted by the subscript, d . It is assumed that the excitation propagates in the $+x$ -direction and that the transverse dimensions of the interface are sufficiently large that there is no \mathbf{E} variation in the y -direction. Therefore, the six electromagnetic field components can be written as:

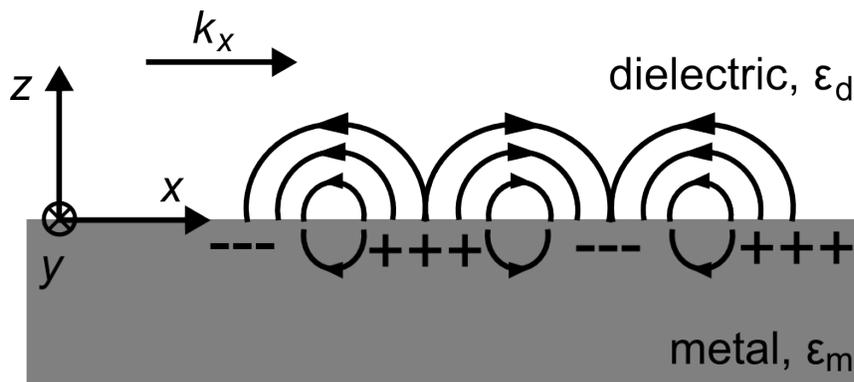


Figure 2.12: Schematic depiction of a surface plasmon polariton (SPP) that is propagating along the interface between a metal (bottom) and a dielectric(top). The SPP consists of an oscillation of charge density coupled to electromagnetic oscillations that are confined to the interfaces and decay exponentially in intensity in the direction normal to the interfaces.

$$\mathbf{E} = \begin{cases} (E_{xd}, E_{yd}, E_{zd})e^{i(k_x x - \omega t)}e^{-\alpha_d z}, & z > 0 \\ (E_{xm}, E_{ym}, E_{zm})e^{i(k_x x - \omega t)}e^{\alpha_m z}, & z < 0 \end{cases} \quad (2.68)$$

$$\mathbf{H} = \begin{cases} (H_{xd}, H_{yd}, H_{zd}) e^{i(k_x x - \omega t)} e^{-\alpha_d z}, & z > 0 \\ (H_{xm}, H_{ym}, H_{zm}) e^{i(k_x x - \omega t)} e^{\alpha_m z}, & z < 0 \end{cases} \quad (2.69)$$

In these equations, $\alpha_d > 0$ and $\alpha_m > 0$ are enforced to ensure exponential field decay away from the interface. Continuity of the tangential field components requires that $E_{xd} = E_{xm} = E_{x0}$, $E_{yd} = E_{ym} = E_{y0}$, $H_{xd} = H_{xm} = H_{x0}$, and $H_{yd} = H_{ym} = H_{y0}$. Assuming the materials are non-magnetic, the normal component of \mathbf{B} must also be continuous across the interface, and $H_{zd} = H_{zm} = H_{z0}$. Finally, Eq. 2.2 requires that $H_{xd} = H_{zd} = 0$, and therefore, $H_{xm} = H_{zm} = 0$. Summarizing these observations, the field components can be written as:

$$\mathbf{E} = \begin{cases} (E_{x0}, E_{y0}, E_{zd}) e^{i(k_x x - \omega t)} e^{-\alpha_d z}, & z > 0 \\ (E_{x0}, E_{y0}, E_{zm}) e^{i(k_x x - \omega t)} e^{\alpha_m z}, & z < 0 \end{cases} \quad (2.70)$$

$$\mathbf{H} = \begin{cases} (0, H_{y0}, 0) e^{i(k_x x - \omega t)} e^{-\alpha_d z}, & z > 0 \\ (0, H_{y0}, 0) e^{i(k_x x - \omega t)} e^{\alpha_m z}, & z < 0 \end{cases} \quad (2.71)$$

Substituting these expressions into Eq. 2.4 produces the following relationships:

$$\frac{\partial H_y}{\partial z} = i\varepsilon_m \omega E_x \quad (2.72)$$

$$0 = -i\varepsilon_m \omega E_y \quad (2.73)$$

$$\frac{\partial H_y}{\partial x} = -i\varepsilon_m \omega E_z, \quad (2.74)$$

and it can be concluded that $E_{y0} = 0$. Substituting the field components into Eq. 2.73 then yields:

$$-\alpha_d H_{y0} = i\varepsilon_d \omega E_{x0} \quad (2.75)$$

$$\alpha_m H_{y0} = i\varepsilon_m \omega E_{x0}. \quad (2.76)$$

Dividing these two expressions produces:

$$\frac{\alpha_d}{\alpha_m} = -\frac{\varepsilon_d}{\varepsilon_m}. \quad (2.77)$$

Next, the dispersive nature of surface plasmons can be understood by applying the wave equation:

$$\nabla^2 \mathbf{E} = \varepsilon \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (2.78)$$

Substituting Eqs. 2.70-2.71 gives the following relationship for $z > 0$:

$$\begin{aligned} (-k_x^2 + \alpha_d^2) \cdot (E_{x0}, 0, E_{z0}) e^{i(k_x x - \omega t)} e^{-\alpha_d z} &= -\varepsilon_d \mu_0 \omega^2 (E_{x0}, 0, E_{z0}) \\ &\cdot e^{i(k_x x - \omega t)} e^{-\alpha_d z} \end{aligned} \quad (2.79)$$

$$-k_x^2 + \alpha_d^2 = -\varepsilon_d \mu_0 \omega^2 \quad (2.80)$$

and the following relationship for $z < 0$:

$$\begin{aligned} (-k_x^2 + \alpha_m^2) \cdot (E_{x0}, 0, E_{z0}) e^{i(k_x x - \omega t)} e^{\alpha_m z} &= -\varepsilon_m \mu_0 \omega^2 (E_{x0}, 0, E_{z0}) \\ &\cdot e^{i(k_x x - \omega t)} e^{\alpha_m z} \end{aligned} \quad (2.81)$$

$$-k_x^2 + \alpha_m^2 = -\varepsilon_m \mu_0 \omega^2. \quad (2.82)$$

Rearranging Eq. 2.80 and 2.82 and dividing the two produces:

$$\frac{\alpha_d^2}{\alpha_m^2} = \frac{k_x^2 - \varepsilon_d \mu_0 \omega^2}{k_x^2 - \varepsilon_m \mu_0 \omega^2}. \quad (2.83)$$

Finally, substituting Eq. 2.77 into this result yields:

$$\frac{\varepsilon_d^2}{\varepsilon_m^2} = \frac{k_x^2 - \varepsilon_d \mu_0 \omega^2}{k_x^2 - \varepsilon_m \mu_0 \omega^2} \quad (2.84)$$

$$\omega^2 = \frac{k_x^2 \left(\frac{\varepsilon_m^2 - \varepsilon_d^2}{\varepsilon_m^2 \varepsilon_d - \varepsilon_d^2 \varepsilon_m} \right)}{\mu_0 \left(\frac{(\varepsilon_m + \varepsilon_d)(\varepsilon_m + \varepsilon_d)}{\varepsilon_m \varepsilon_d (\varepsilon_m - \varepsilon_d)} \right)} \quad (2.85)$$

$$\omega = ck_x \sqrt{\varepsilon_d \varepsilon_m} \sqrt{\frac{\varepsilon_d + \varepsilon_m}{\varepsilon_d \varepsilon_m}}. \quad (2.86)$$

This expression is considerably different than the energy-momentum relationship of an incident free-space photon (i.e. $\omega = ck_0$), and naturally, this implies additional complexities when it comes to exciting SPPs with free-space photons.

2.5.2 Momentum matching between free-space and surface plasmon waves

A plot of the real component of k_x along with the light line is shown in Fig. 2.13. From this figure, it is evident that the dispersion relation of free-space photons lies to the left of the dispersion relation for SPPs. Therefore, when a photon is propagating through free-space (or any other dielectric) and is incident on a metal interface, it does not carry sufficient momentum to excite a SPP. In order for a SPP to be excited, momentum must be added to the photon as it is incident upon the interface. There are several well-established methods of supplying the additional momentum, which are described in the following sections.

The Kretschmann Configuration

This configuration consists of a dielectric prism that is coated with metal and possibly another low-index dielectric, as shown in Fig. 2.14(a). Free-space radiation is passed through the prism, multiplying its momentum by the magnitude of the refractive index of

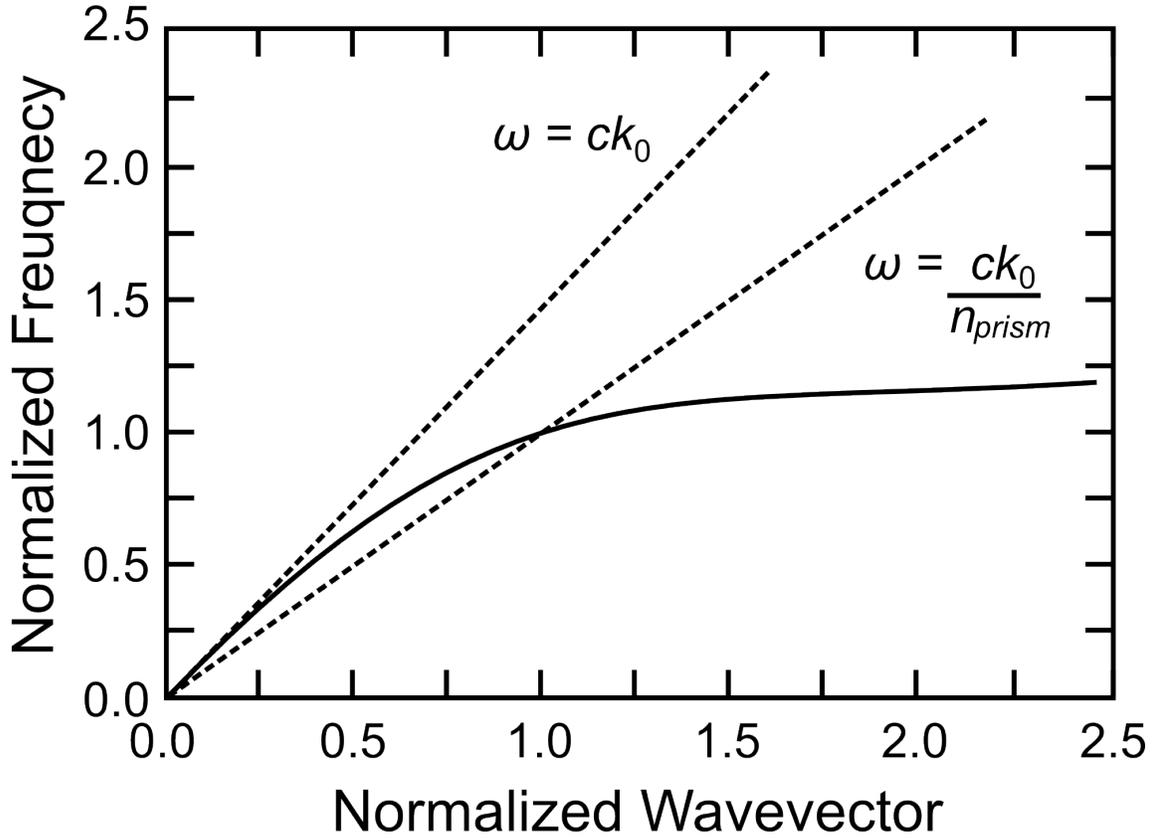


Figure 2.13: Conceptual plot of the surface plasmon polariton (SPP) dispersion relation (solid line), along with the dispersion relation of free-space photons and photons propagating through a prism. Notably, there are no points where the dispersion relations for a SPP and free-space photons intersect. (Adapted from [46])

the prism, n_{prism} . For an appropriate choice of prism material, the photon momentum in the prism will be greater than the momentum of SPPs on the metal-free-space interface for a certain range of wave vectors, as shown in Fig. 2.13. Furthermore, there is now a point where the two lines intersect, signifying that the photons can be used to excite SPPs. By controlling the angle at which the TM radiation is incident on the prism-metal interface, it is possible to control the component of the momentum that is projected onto the interface. Therefore, by careful adjustment of the angle of incidence, θ_{SP} , it is possible to match the projected momentum to the surface plasmon momentum, providing the conditions necessary to excite a surface plasmon wave. This can be deduced from the following expression:

$$k_x = k_0 \sin(\theta_{SP}) = \frac{\omega}{c} n_{prism} \sin(\theta_{SP}) \quad (2.87)$$

If the metal film is sufficiently thin ($\sim 45\text{nm}$), the radiation can evanescently couple through the metal and excite a surface plasmon on the top metal-dielectric interface.

The Otto Configuration

The Otto configuration works on essentially the same principle as the Kretschmann configuration, and is depicted schematically in Fig. 2.14(b). Electromagnetic radiation is passed through a prism, increasing its momentum. In this case, the metal film is not deposited directly on the prism surface, but is separated from the metal film by a low refractive index thin film, so that the momentum-matched radiation can couple evanescently through the low refractive index material onto the metal surface.

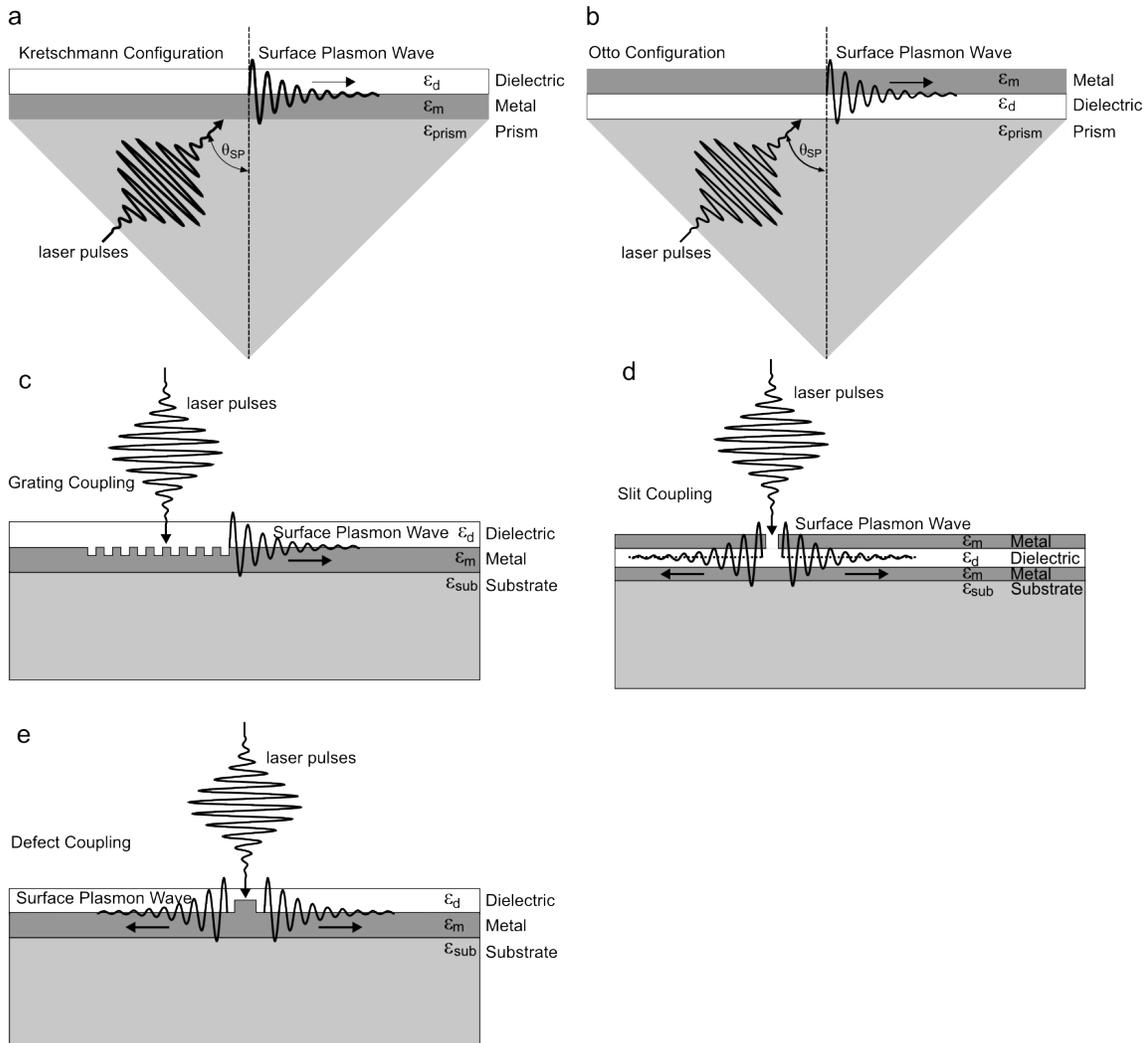


Figure 2.14: Schematic depiction of: (a) Kretschmann configuration, (b) Otto configuration, (c) grating coupling, (d) slit Coupling, and (e) defect coupling for excitation of surface plasmon waves.

Diffraction

Alternatively, diffraction can be used to add momentum to the free-space radiation. In the case of a diffraction grating, the momentum of electromagnetic radiation incident on

the grating will be:

$$k_x = k_0 \sin(\theta_{SP}) \pm \frac{2\pi m_i}{a}, \quad (2.88)$$

where a is the grating period and m_i is an integer. This scheme is shown in Fig. 2.14(c). Notably, other geometric features that give rise to diffraction can also be used for SPP coupling. Two other examples are a single slit in a metal film or a ridge defect in a metal film, shown in Figs. 2.14(d),(e), respectively.

Surface Plasmon Field Decay

Further consideration of Eqs. 2.80 and 2.82 gives insight into the field decay away from the interface. First it is assumed that the surface plasmon electric field can be written as, $\mathbf{E}_{SP} = \eta \mathbf{E}_0 e^{i(k_x x - \omega t)} e^{-\alpha_d z}$, where η represents the factor by which the free-space electric fields are enhanced as they become tightly confined to the surface plasmon mode. Assuming that the Kretschmann geometry is used, Eq. 2.87 is substituted into Eq. 2.80, as follows:

$$\alpha_d^2 = \frac{\omega^2}{c^2} n_{prism}^2 \sin^2(\theta_{SP}) - \frac{\omega^2}{c^2} n_d^2 \quad (2.89)$$

$$\alpha_d = \frac{\omega}{c} \sqrt{n_{prism}^2 \sin^2(\theta_{SP}) - n_d^2}. \quad (2.90)$$

For example values of $\lambda = 1550\text{nm}$, $\theta = 45^\circ$, $n_{prism} = 1.5$, and $n_d = 1$, a field decay length scale of $\alpha_d = 698\text{nm}$ is calculated. In the same manner, Eq. 2.87 can be used in Eq. 2.82, producing:

$$\alpha_m = \frac{\omega}{c} \sqrt{n_{prism}^2 \sin^2(\theta_{SP}) - n_m^2}, \quad (2.91)$$

where n_m is the real component of the refractive index of the metal.

2.5.3 Waveguides: Two-dimensional Confinement

In the previous section, it was assumed that surface plasmons were excited on an infinite planar interface between a metal and a dielectric. Spatial reduction of this interface to a finite, and in particular, to a sub-wavelength scale in the y -dimension, gives the structure functionality as a waveguide with a tightly confined mode profile and enhanced electric fields. In contrast to SOI waveguides, plasmonic waveguides can be reduced to sub-100nm dimensions while maintaining high mode confinement, making them an attractive device for dense optical circuitry with nonlinear functionality. Notably, the enhanced electric fields enable efficient access to nonlinear optical effects and unique surface effects that are not typically present in photonic waveguides.

The geometry of nanoplasmonic waveguides makes it challenging to obtain the mode parameters and field distributions through analytical solutions of Maxwell's equations. Therefore, a numerical mode solver is typically used to obtain the effective refractive index of the propagating mode (which determines the momentum of the mode) and the

propagation loss (which determines the useful device length)². As discussed previously, an important aspect of planar surface plasmon waves is momentum matching between the free-space laser radiation and the surface plasmon wave. In a similar manner, there are specific conditions that must be satisfied to achieve efficient coupling to plasmonic waveguides.

Coupling

A conceptually similar configuration could be realized by fabricating the plasmonic waveguide on top of a prism (or dielectric wedge), enabling coupling in the same manner as the Kretschmann or Otto configurations. In the majority of cases, it is not practical to fabricate plasmonic waveguides on a prism, and a planar substrate is used. In this case, momentum matching is often achieved via diffraction effects. Placing a diffraction grating directly adjacent to the input and output facets of the plasmonic waveguide enables normally-incident laser radiation to couple to the plasmonic mode. Alternatively, a sub-wavelength slit coupler, similar to the one in Fig. 2.14(d) can be fabricated adjacent to the input or output of the plasmonic waveguide, facilitating coupling in this manner. A significant diffraction effect can also occur when laser radiation is normally incident on the waveguide end facets, making it possible to end-fire coupled radiation from free-space to nanoplasmonic waveguide modes.

Although fundamentally similar to planar surface plasmons, waveguide plasmons can also be analyzed using techniques from photonic waveguides and linear circuitry. When analyzed from this perspective, the efficiency of coupling to a surface plasmon waveguide becomes easier to quantify and several alternative coupling schemes become apparent. In the case of a plane wave incident on a nanoplasmonic waveguide end-facet, the plane wave will propagate with a characteristic impedance of $Z_1 = 377\Omega$ in free-space. Moreover, the plasmonic mode profile has a characteristic impedance. For example, the impedance of a plasmonic two-wire optical transmission line was calculated to be $Z_2 = (216 - 5.5i)\Omega$ at $\lambda = 830\text{nm}$ [46], which is comparable to analogous structures at radio frequencies. Using these example impedances, the reflection coefficient due to impedance mismatch can be calculated via:

$$\Gamma = \frac{Z_2 - Z_1}{Z_2 + Z_1} = -0.27 - 0.00118i \quad (2.92)$$

$$|\Gamma| = 0.27 \quad (2.93)$$

Here, the negative real component of the reflection coefficient signifies that the reflected wave receives a 180° phase shift. In addition to impedance matching, the distribution of electromagnetic fields of the free-space radiation should be matched as closely as

²The mode profiles in this thesis were obtained using Lumerical FDTD Solutions, where the waveguide geometry is meshed using the finite difference algorithm, a matrix eigenvalue problem based on Maxwell's equations is setup, and the matrix is solved using sparse matrix techniques.

possible to those of the plasmonic mode profile. This is fundamentally challenging for sub-diffraction waveguide core dimensions and, as a result, coupling efficiencies to sub-diffraction plasmonic waveguides are typically less than 10%. A mode overlap factor can be introduced to account for the field profile mismatch, which is given by:

$$f(A, B) = \frac{|\int \int A(x, y)B^*(x, y)dx dy|^2}{\int \int |A(x, y)|^2 dx dy \cdot \int \int |B(x, y)|^2 dx dy}, \quad (2.94)$$

where $A(x, y)$ is the spatial distribution of electromagnetic energy of the input beam (typically a Gaussian distribution), $B(x, y)$ is the spatial distribution of electromagnetic energy of the plasmonic mode, and the area integrals are carried out over an infinite area, which can be truncated to the area containing a majority of the electromagnetic energy.

These same circuit principles can be extended to end-fire excitation of plasmonic waveguides using other structures, such as optical fibres or photonic waveguides. For large long-range SPP waveguides with a width, $w = 5\mu\text{m}$, the field distribution and impedance are matched very closely to a single-mode optical fibre, and a high coupling efficiency of 98% has been demonstrated [47]. Naturally, excitation of subwavelength plasmonic waveguides directly from free-space is much less efficient, and is on the order of 2-3%. High-efficiency excitation of nanoplasmonic waveguides via end-fire excitation with photonic waveguides is possible, and efficiencies in the range of 30% – 80% have been demonstrated [48, 49]. Therefore, on-chip radiation can be easily and efficiently coupled to nanoplasmonic waveguide modes.

An overview of several popular classes of plasmonic waveguides is presented in the following sections.

Metal-Insulator (MI) and Metal-Insulator-Metal (MIM) Waveguides

Adding nanoscale noble metal features to a SOI waveguide produces a Si-loaded nanoplasmonic waveguide. Adding a metal cap to the SOI waveguide, as shown in Fig. 2.15(a), produces the MI nanoplasmonic waveguide. Alternatively, adding two adjacent gold features produces the MIM nanoplasmonic waveguide, as shown in Fig. 2.15(b). As in the case of a SOI waveguide, the silicon cores of these two nanoplasmonic waveguides have cross-sectional dimensions, width, w , and height, h . In addition, for the MI waveguide, the gold cap has a thickness, t , and for the MIM waveguide, the gold sides have widths, w_{Au} .

The mode profile for a MI nanoplasmonic waveguide with, $w = 250\text{nm}$, $h = 340\text{nm}$, and $t = 60\text{nm}$, is shown in Fig. 2.16(a). Scaling down the width to $w = 100\text{nm}$ produces the mode profile shown in Fig. 2.16(b). Despite the ultracompact core dimensions, 80% of the mode energy is confined to the Si core for $w = 250\text{nm}$ and 45% for $w = 100\text{nm}$. In this case, the asymmetry of the geometry enables the mode to expand into free-space as w is decreased. However, the confinement is still 75% better than SOI waveguides with

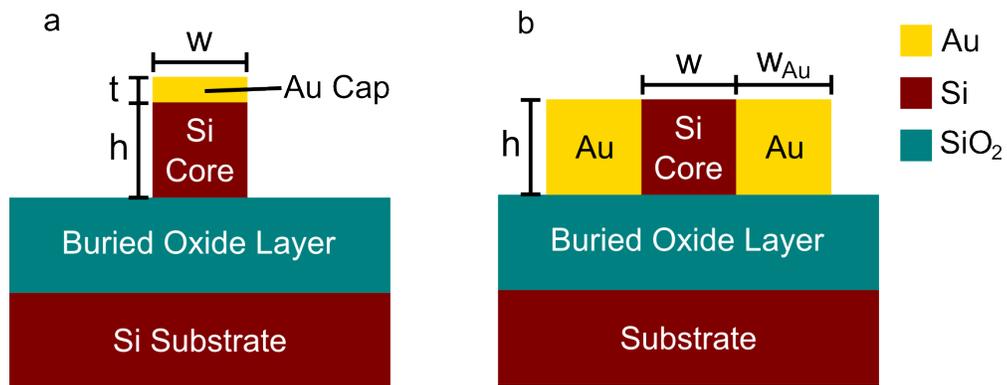


Figure 2.15: Schematic of (a) metal-insulator and (b) metal-insulator-metal nanoplasmonic waveguides.

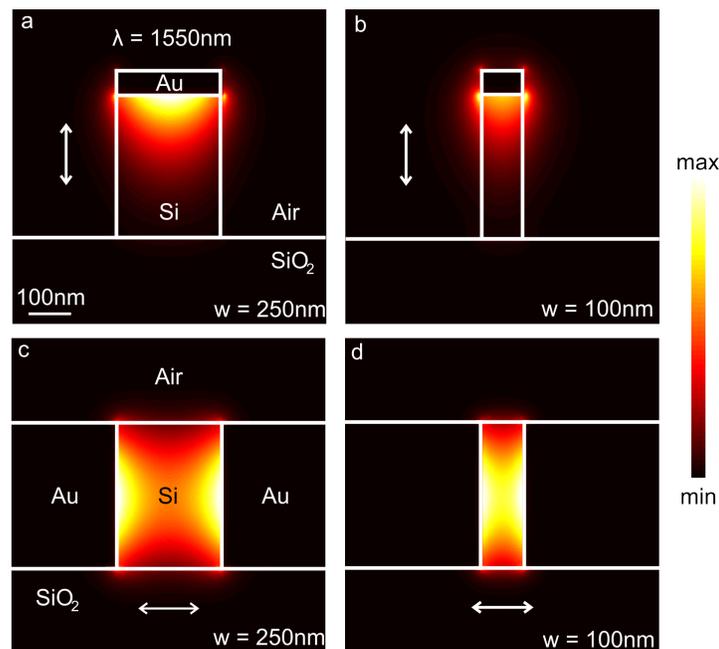


Figure 2.16: Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for IM plasmonic waveguides with a width, (a) $w = 250\text{nm}$; and (b) $w = 100\text{nm}$. Mode profile intensity distributions at $\lambda = 1550\text{nm}$ for MIM plasmonic waveguides with widths, (c) $w = 250\text{nm}$; and (d) $w = 100\text{nm}$.

the same dimensions. As a result of the geometric asymmetry, the mode intensity decays with a length scale of $\alpha_0 = 135\text{nm}$ away from the Au-Si interface. A MIM nanoplasmonic waveguide with $w = 250\text{nm}$, $h = 340\text{nm}$, and $w_{Au} = 300\text{nm}$, produces the mode profile shown in Fig. 2.16(c). Scaling down the width to $w = 100\text{nm}$ produces the mode profile shown in Fig. 2.16(d). In this case, electromagnetic energy in the Si core is confined on both sides. Such confinement enables $\approx 91 - 95\%$ of the electromagnetic energy to be confined to the Si core, regardless of w .

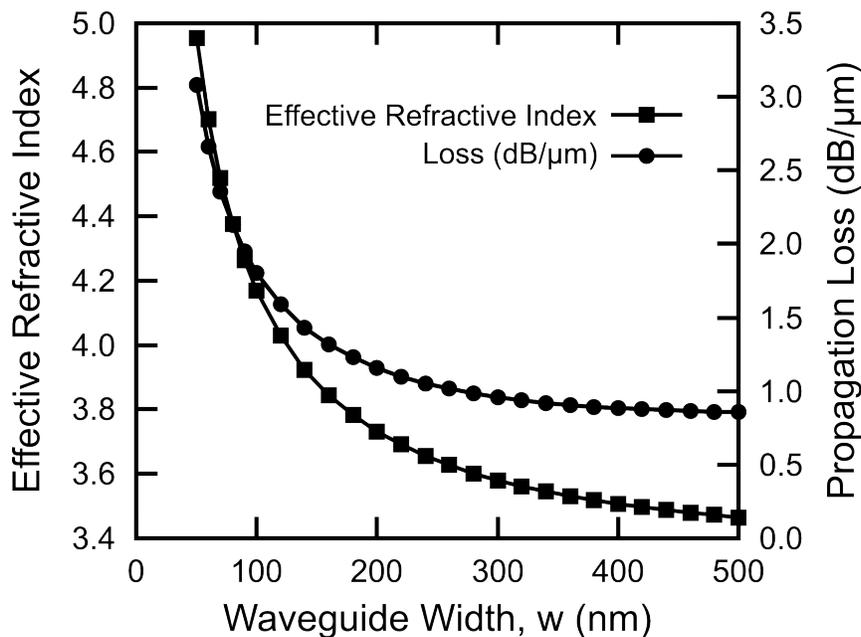


Figure 2.17: Effective refractive index and propagation losses of modes in Au-Si-Au plasmonic waveguides with a constant height, $h = 340\text{nm}$ and a varying width, w .

A decrease in w increases the fraction of the mode energy that interacts with the Au sidewalls, which has two implications: (1) the effective refractive index of the mode will increase; and (2) the overall Ohmic losses of the mode will increase, decreasing the propagation length of the mode. These implications are summarized in Fig. 2.17(a), which shows the effective refractive index and propagation losses as a function of w . Decreasing w from $w = 500\text{nm}$ to $w = 50\text{nm}$ yields an increase in the effective refractive index from $n_{eff} = 3.47$ to $n_{eff} = 4.95$. Similarly, a MIM waveguide with $w = 500\text{nm}$ has a propagation loss of $0.86\text{dB}/\mu\text{m}$, corresponding to a propagation length of $L_{prop} = 5.06\mu\text{m}$, whereas a MIM waveguide with $w = 50\text{nm}$ has a propagation loss of $3.08\text{dB}/\mu\text{m}$, or a propagation length of $L_{prop} = 1.41\mu\text{m}$. Therefore, there exists a trade-off between mode confinement and propagation loss, which exists in most plasmonic systems.

In a similar manner as for the SOI waveguide, the core power fraction is calculated for the MIM nanoplasmonic waveguide, as shown in Fig. 2.17(b). In this case, the fraction of power confined to the Si core remains relatively constant in the range of $91\% - 95\%$. Such high confinement at nanoscale dimensions makes this structure useful for high-density optical circuitry. This is demonstrated by performing mode-coupling simulations in the

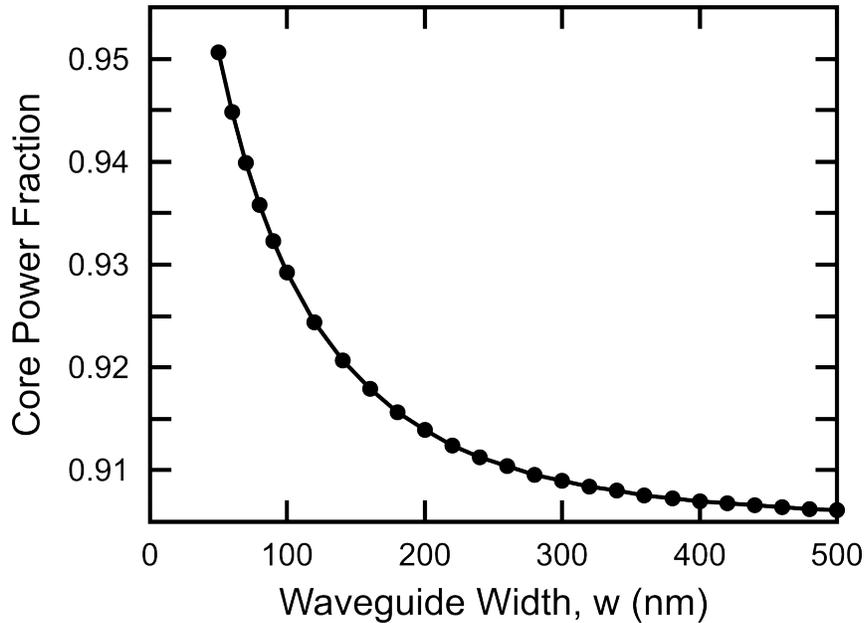


Figure 2.18: Fraction of the mode energy that is confined to the Si core in Au-Si-Au plasmonic waveguides with a constant height, $h = 340\text{nm}$ and a varying width, w .

same manner as in Fig. 2.6, which is depicted schematically in Fig. 2.19(a). In this case, the two Si waveguides are clad on both sides by a Au film. In this way, the 100nm gap between the waveguides is filled with Au, making it virtually impossible for the fields to penetrate from Waveguide 1 to Waveguide 2. The simulated intensity distribution, shown in Fig. 2.19(b), demonstrates that within a length scale of several propagation lengths of the plasmonic mode, virtually all of the mode energy remains in Waveguide 1.

Hybrid Mode Plasmonic Waveguides

If a low-index dielectric (such as SiO_2) is used to separate the Si from the Au in the MI nanoplasmonic waveguide of Fig. 2.15(a), it is possible to excite hybrid photonic-plasmonic modes, and these waveguides are referred to as “hybrid plasmonic waveguides.” These modes show characteristics of both plasmonic waveguides and photonic waveguides and enable longer propagation lengths compared to the MI nanoplasmonic waveguide. If a $t_{\text{SiO}_2} = 50\text{nm}$ layer of SiO_2 is inserted between the Si and Au features of the MI nanoplasmonic waveguide with $w = 250\text{nm}$, the mode profile shown in Fig. 2.20(a) is obtained. If instead, the waveguide has $w = 100\text{nm}$, this yields the mode profile shown in Fig. 2.20(b). Notably, the $w = 250\text{nm}$ hybrid plasmonic waveguide has a propagation length, $L_{prop} = 25.8\mu\text{m}$, which is approximately a factor of 6 greater than the MI nanoplasmonic waveguide. Similarly, for $w = 100\text{nm}$, a propagation length of $L_{prop} = 9.65\mu\text{m}$ is obtained, which is approximately 4 times greater than for the MI nanoplasmonic waveguide.

Examining the mode profiles of Fig. 2.20, it is observed that a large fraction of the mode energy is confined to the SiO_2 layer. For the $w = 250\text{nm}$ waveguide, approximately

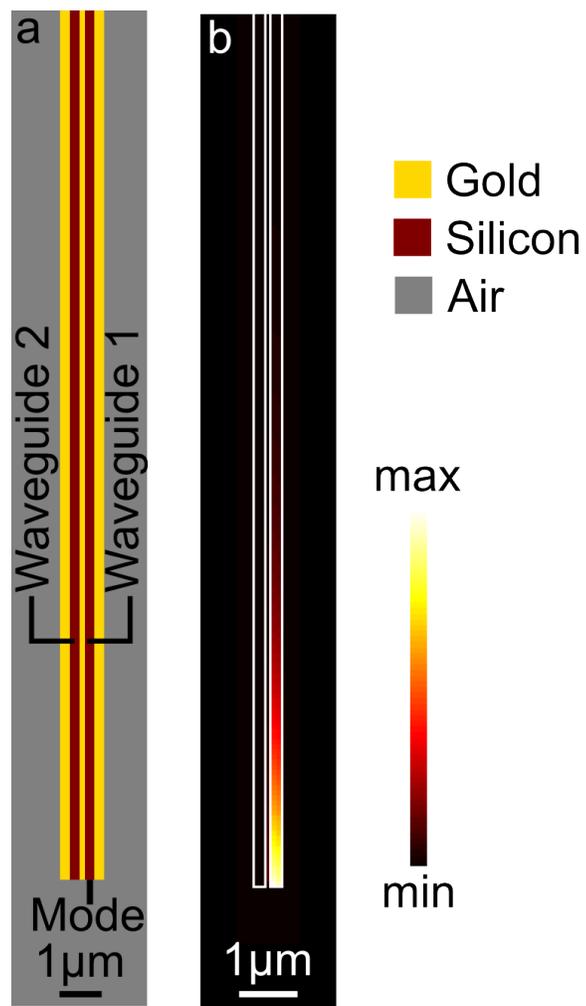


Figure 2.19: (a) Top-view schematic of two Au-Si-Au plasmonic waveguides with $w = 250\text{nm}$, separated by a 100nm Au film. (b) Finite-difference time-domain intensity (I) distributions of the waveguides depicted in (a), demonstrating effective confinement of the mode energy to the original waveguide.

36.8% of the energy is confined to the SiO_2 , whereas 36.0% propagates in the Si layer. For the $w = 100\text{nm}$ waveguide, 42.9% of the energy propagates in the SiO_2 layer, whereas only 25.6% propagates in the Si core. While these characteristics are desirable for high-confinement passive routing in a very small volume (i.e. the SiO_2 layer), they are highly undesirable for nonlinear circuitry for two reasons: (1) Nonlinear interactions are much more pronounced in Si than in SiO_2 , as its $\chi^{(3)}$ nonlinear coefficient is greater by a factor of approximately 10^2 . Therefore, for nonlinear circuitry it is desirable for the energy to propagate in the Si core. For a very small w , the fraction of energy propagating in the Si core is much less than for the MI waveguide, putting it at a disadvantage for nonlinear circuitry. (2) Plasmonic enhancement of electric fields is often much greater than what is predicted theoretically, due to surface roughness between two adjacent materials. For this reason, it is highly desirable for the enhancement to occur at the interface between the metal and the nonlinear medium, i.e. Si. For the hybrid plasmonic waveguide, the enhancement occurs at the metal- SiO_2 interface. For these reasons, it is concluded that the hybrid plasmonic waveguide is a superior choice for passive routing, whereas the MI waveguide exhibits distinct advantages for nonlinear circuitry.

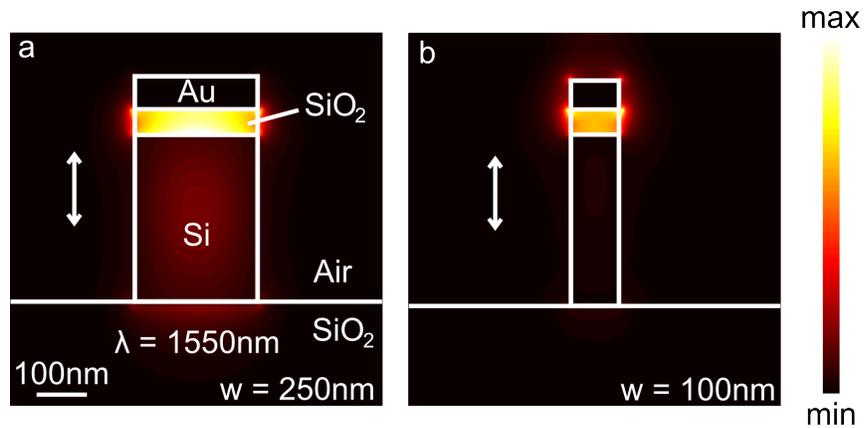


Figure 2.20: Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for hybrid plasmonic waveguides with a height, $h = 340\text{nm}$, and widths, $w =$ (a) 250nm ; and (b) 100nm .

Long-range Surface Plasmon Polariton Waveguides

Long-range surface plasmon polariton waveguides typically consist of a thin ($\sim 20\text{nm}$) noble metal stripe embedded in a low-index cladding. Provided that the indices of the substrate and cladding are very closely matched, it is possible to excite plasmonic modes that can propagate long distances on the order of mm. A mode profile for a $w \times h = 5\mu\text{m} \times 20\text{nm}$ Au stripe embedded in a SiO_2 film is shown in Fig. 2.21, and is found to propagate a distance, $L_{prop} = 135\mu\text{m}$. Notably, using a silver film will increase this distance to several mm. While the losses of this mode are low, the mode is not confined in any way, and it does not offer advantages over photonic waveguides in this regard. While this plasmonic waveguide is not useful for high-density circuitry, it has found applications

in biosensing and low quantum noise plasmonic lasers [50, 51].

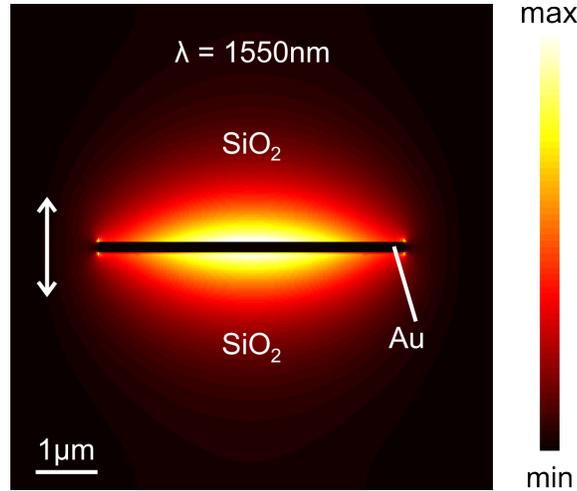


Figure 2.21: Mode profile intensity (I) distribution of a long-range surface plasmon polariton waveguide with a gold film thickness of 20nm and a width, $w = 5\mu\text{m}$.

2.6 The Ponderomotive Force

The ponderomotive force is a nonlinear electromagnetic force that can be understood from the Lorentz force acting on a charged particle. In a solid state system, the ponderomotive force is most significant in the vicinity of a highly confined electromagnetic field, such as the plasmonic fields at a metal-dielectric interface or the ones shown in the MI nanoplasmonic waveguide mode profiles in Fig. 2.16(a),(b). Here, we consider the ponderomotive force acting in vacuum, or in a medium in the ballistic regime, where collisions with the crystal lattice can be neglected. These conditions are satisfied when the interaction length under consideration is less than the mean free path of electrons in the crystal. For example, electrons in crystalline Si have a mean free path of 28nm, which is the maximum interaction length that the following analysis is accurate for.

Electrons can be optically injected in close proximity to a plasmonic field through several nonlinear optical interactions arising from the excitation radiation, including multi-photon ionization or tunnelling ionization in the plasmonic metal, or multi-photon absorption in an adjacent semiconductor layer (e.g. TPA in an adjacent Si layer). When an electron is placed in the oscillating electromagnetic field, it accelerates in response to the electric and magnetic fields, where the contribution from the magnetic field is typically much weaker. When the field reverses direction, the velocity of the electron will also change directions. A special condition arises in the case of a highly asymmetric oscillating electromagnetic field. Depending on the position and velocity of the electron, the electron can be accelerated away from the metal film. Because the field is tightly confined to the metal film, it is accelerated into a region where the electric field is significantly weaker. When the electric field reverses direction, the Lorentz force imparted on that particle

will be weaker than what it experienced in the previous half-cycle of the electromagnetic wave. Therefore, the distance travelled away from the metal film in the first half-cycle will be greater than the distance travelled towards the metal film in the second half-cycle. This interaction continues to build in a cycle-by-cycle manner, and the electron is rapidly accelerated away from metal film. Eventually, it travels far enough from the metal film that its interaction with the electromagnetic fields is negligible and it continues to move away from the metal film with a fixed velocity. Ponderomotive electron acceleration is schematically depicted in Fig. 2.22, where the solid line represents the time-dependent electric field experienced by the electron and the dotted line represents its trajectory. After several oscillations of the electric field, the electron has been imparted with a net velocity increase.

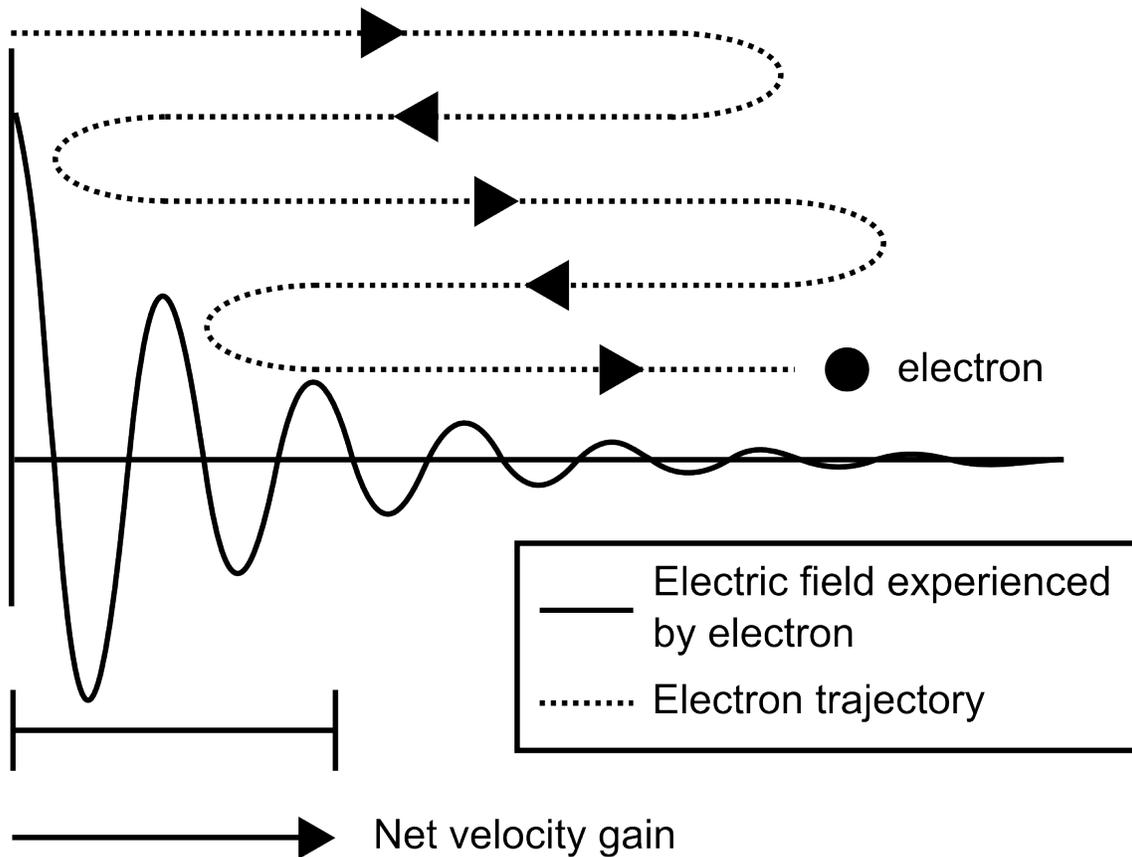


Figure 2.22: Schematic depiction of cycle-by-cycle ponderomotive electron acceleration by a SPP field.

In the case of a prism coated with a thin metal film (i.e. the Kretschmann configuration), free electrons generated by multi-photon ionization or tunnelling ionization in the metal film can be accelerated to energies of $\sim 2\text{keV}$ [52–54]. Electron acceleration has also been performed using several nanoplasmonic structures, such as nanotips [55], nanorods [56], and nanoparticles [57]. In the case where a plasmon is propagating along a metal-semiconductor interface, free-carriers can be excited by direct or multi-photon absorption in the semiconductor, which can occur with a much higher probability than multiphoton

ionization or tunnelling ionization in the metal. In this way, free-carriers can experience ponderomotive acceleration in a *solid*. In contrast to the the case of acceleration *in vacuo*, the accelerated free-carriers will undergo frequent collisions with valence electrons of the solid. Such an interaction has never before been investigated, and a significant portion of this thesis is devoted to this topic.

2.6.1 Derivation of the Ponderomotive Force

In deriving an expression for the ponderomotive force acting on an electron, it is necessary to consider the equation of motion of an electron of charge, e , placed in an oscillating plasmonic field:

$$m_e^* \frac{d\mathbf{v}}{dt} = -e[\mathbf{E}_{\mathbf{SP}}(\mathbf{r}, t) + \mathbf{v} \times \mathbf{B}_{SP}(\mathbf{r}, t)], \quad (2.95)$$

where \mathbf{v} , $\mathbf{E}_{\mathbf{SP}}(\mathbf{r}, t)$ and $\mathbf{B}_{SP}(\mathbf{r}, t)$ are the time-dependent velocity of the electron and the plasmonic electric and magnetic fields that it experiences, respectively. Here, it has been assumed that the electron can be accelerated in an arbitrary medium, where its mass is replaced by an effective mass, m_e^* . The contribution of the $\mathbf{E}_{\mathbf{SP}}$ field is dominant in determining the electron trajectory, and the $\mathbf{v} \times \mathbf{B}_{SP}(\mathbf{r}, t)$ term is subsequently treated as a second-order perturbation.

In the vicinity of a plasmonic structure, an electron can experience an enhanced and spatially inhomogeneous plasmonic electric field, $\mathbf{E}_{SP}(\mathbf{r}, t)$, of the form,

$$\mathbf{E}_{\mathbf{SP}}(\mathbf{r}, t) = \eta \mathbf{E}_0(\mathbf{r}) \cos(\omega t), \quad (2.96)$$

where ω is the angular frequency, η is the enhancement factor (relative to the free-space fields that excited the plasmonic structure), and $\mathbf{E}_0(\mathbf{r})$ is the spatially-dependent plasmonic electric field amplitude.

Neglecting the second-order $\mathbf{v} \times \mathbf{B}_{\mathbf{SP}}(\mathbf{r}, t)$ term, the first-order velocity response of an electron, \mathbf{v}_1 , to a plasmonic electric field, $\mathbf{E}_{SP}(\mathbf{r} = \mathbf{r}_0)$, at an initial position, \mathbf{r}_0 , can be written as:

$$m_e^* \frac{d\mathbf{v}_1}{dt} = -e \mathbf{E}_0(\mathbf{r}_0) \cos(\omega t). \quad (2.97)$$

Integrating Eq. 2.97, the first-order velocity is obtained:

$$\mathbf{v}_1 = -\frac{e}{(m_e^* \omega)} \mathbf{E}_0(\mathbf{r}_0) \sin(\omega t). \quad (2.98)$$

Similarly, the incremental change in position can be written as:

$$\delta \mathbf{r} = \frac{e}{m_e^* \omega^2} \mathbf{E}_0(\mathbf{r}_0) \cos(\omega t). \quad (2.99)$$

Next, the $\mathbf{v} \times \mathbf{B}_{\mathbf{SP}}(\mathbf{r}, t)$ contribution is added as a second-order perturbation to the Lorentz

equation of motion. Expanding $\mathbf{E}_{\text{SP}}(\mathbf{r})$ about r_0 yields:

$$\mathbf{E}_{\text{SP}}(\mathbf{r}) = \mathbf{E}_{\text{SP}}(\mathbf{r}_0) + (\delta\mathbf{r}_1 \cdot \nabla)\mathbf{E}_{\text{SP}}|_{\mathbf{r}=\mathbf{r}_0} + \dots \quad (2.100)$$

The second-order expression for the magnetic field can be obtained from Eq. 2.3, $\nabla \times \mathbf{E}_{\text{SP}} = -\partial\mathbf{B}_{\text{SP},1}/\partial t$, as

$$\mathbf{B}_{\text{SP},1} = -\frac{1}{\omega}\nabla \times \eta\mathbf{E}_0(\mathbf{r}_0)\sin(\omega t). \quad (2.101)$$

Now, the second-order expression for the Lorentz force equation can be written as:

$$m_e^* \frac{d\mathbf{v}_2}{dt} = -e[(\delta\mathbf{r}_1 \cdot \nabla)\mathbf{E}_{\text{SP}} + \mathbf{v}_1 \times \mathbf{B}_{\text{SP},1}]. \quad (2.102)$$

Substituting Eqs. 2.98, 2.99, and 2.101 into Eq. 2.102 produces the following expression for the instantaneous nonlinear force:

$$m_e^* \frac{d\mathbf{v}_2}{dt} = -\frac{e^2}{m_e^*\omega^2} [((\eta\mathbf{E}_0(\mathbf{r}_0)\cos(\omega t)) \cdot \nabla)\eta\mathbf{E}_0(\mathbf{r})\cos(\omega t) + \quad (2.103)$$

$$+(-\eta\mathbf{E}_0(\mathbf{r}_0)\sin(\omega t)) \times (-\nabla \times \eta\mathbf{E}_0(\mathbf{r}_0)\sin(\omega t))]. \quad (2.104)$$

Since the electron's centre of oscillation drifts towards a region of a weaker field over many cycles, the high-frequency oscillating $\cos^2(\omega t)$ and $\sin^2(\omega t)$ terms can be time-averaged to yield 1/2, and Eq. 2.104 is rearranged to produce:

$$m_e^* \left\langle \frac{d\mathbf{v}_2}{dt} \right\rangle = -\frac{e^2\eta^2}{2m_e^*\omega^2} [(\mathbf{E}_0 \cdot \nabla)\mathbf{E}_0 + \mathbf{E}_0 \times (\nabla \times \mathbf{E}_0)]. \quad (2.105)$$

Using the vector identity, $\nabla E_0^2 = 2(\mathbf{E}_0 \cdot \nabla)\mathbf{E}_0 + \mathbf{E}_0 \times (\nabla \times \mathbf{E}_0)$, the equation of motion for the centre of oscillation reduces to

$$m_e^* \left\langle \frac{d\mathbf{v}_2}{dt} \right\rangle = -\frac{e^2\eta^2}{4m_e^*\omega^2} \nabla E_0^2. \quad (2.106)$$

Finally, the ponderomotive force can be written in terms of the plasmonic field as:

$$\mathbf{F}_{\text{Pond}} = -\frac{e^2}{4m_e^*\omega^2} \nabla E_{\text{SP}}^2, \quad (2.107)$$

and the corresponding ponderomotive potential as:

$$U_{\text{Pond}} = \frac{e^2}{4m_e^*\omega^2} E_{\text{SP}}^2. \quad (2.108)$$

From this, it can be concluded that the oscillation centre experiences a nonlinear ponderomotive force that is proportional to the gradient of the intensity of the plasmonic field. To achieve a large F_{Pond} , both large field amplitudes and steep gradients are required. These criteria are easily satisfied by the MI nanoplasmonic waveguide mode profile, as shown

in Figs. 2.16(a),(b), which decays on length scale of $\alpha_0 = 135\text{nm}$. For ponderomotive acceleration in a Si structure, where $m_e^* = 0.26m_0 = 2.37 \times 10^{-31}\text{kg}$, $\lambda = 1550\text{nm}$, and $E_{SP} = 5 \times 10^9\text{V/m}$, a ponderomotive potential of $U_{Pond} = 2.75\text{eV}$ is calculated. Therefore, by coupling intense $\lambda = 1550\text{nm}$ pulses to a MI waveguide, free-carriers can be excited via TPA, and will subsequently experience ponderomotive acceleration in the steep field gradient of the plasmonic mode profile. Notably, the free-carriers will experience this acceleration in the Si crystal lattice, leading to frequent collisions with valence electrons of the Si lattice.

Since $U_{Pond} \propto \omega^{-2}$, ponderomotive acceleration becomes stronger as the wavelength of radiation is increased. However, as the wavelength is increased, the operation shifts further from the plasma frequency of noble metals, the SPP field confinement decreases, and the intensity gradient decreases. Therefore, there exists an interplay between the radiation wavelength and the field confinement that will enable the highest ponderomotive acceleration. Ultimately, these factors influence the energy spectrum of electrons accelerated by the SPP field.

2.7 Impact Ionization and Avalanche Multiplication

Impact ionization is a high-energy collision process that exists in semiconductors [58]. When a sufficiently high electric field is applied to the semiconductor, free-carriers will be accelerated. Free-carriers that attain a high energy undergo scattering events with valence band electrons, which are bound to the nucleus. During these collisions, the kinetic energy of the free-carrier can be transferred to the valence band electron, exciting it to the conduction band and leaving a new electron-hole pair. The new electron-hole pair can subsequently be accelerated in the electric field and undergo a collision with a bound electron in the same manner. This process cascades in an avalanche mode and a very small number of free-carriers can be multiplied drastically [59]. Avalanche multiplication of free-carriers via impact ionization is depicted schematically in Figure 2.23.

Notably, there exists a threshold energy, E_t , that must be attained by the electrons and imparted to the valence band electrons in order for impact ionization to occur. In the case of Si, this threshold has been measured to be $E_t = 2.3\text{eV}$ [58].

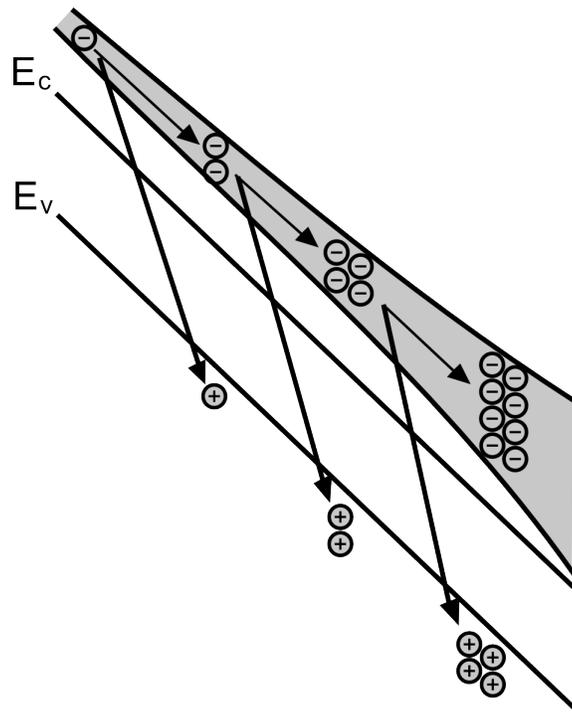


Figure 2.23: Schematic depiction of impact ionization process, leading to avalanche multiplication of free-carriers.

References

- [1] M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light*. Cambridge University Press, London, 1999.
- [2] P. N. Butcher and D. Cotter. *The Elements of Nonlinear Optics*. University Press, Cambridge, 1990.
- [3] D. Van Thourhout, T. Spuesens, S.K. Selvaraja, Liu Liu, G. Roelkens, R. Kumar, G. Morthier, P. Rojo-Romeo, F. Mandorlo, P. Regreny, O. Raz, C. Kopp, and L. Grenouillet. Nanophotonic devices for optical interconnect. *Selected Topics in Quantum Electronics, IEEE Journal of*, 16:1363–1375, 2010.
- [4] V. Yuriy and S. McNab. Losses in single-mode silicon-on-insulator strip waveguides and bends. *Opt. Express*, 12:1622–1631, 2004.
- [5] E. D. Palik. *Handbook of optical constants of solids*. Academic Press, San Diego, 1998.
- [6] A. R. Motamedi, A. H. Nejadmalayeri, A. Khilo, F. X. Kärtner, and E. P. Ippen. Ultrafast nonlinear optical studies of silicon nanowaveguides. *Opt. Express*, 20:4085–4101, 2012.
- [7] R. Soref. The past, present, and future of silicon photonics. *IEEE J. Sel. Topics*, 12:1678–1687, 2006.
- [8] B. E. Little and W. P. Huang. Coupled-mode theory for optical waveguides. *Prog. Electromagn. Res.*, 10:217–270, 1995.
- [9] M. Galli, D. Gerace, K. Welna, T. F. Krauss, L. O’Faolain, G. Guizzetti, and L. C. Andreani. Low-power continuous-wave generation of visible harmonics in silicon photonic crystal nanocavities. *Opt. Express*, 18:26613–26624, 2010.
- [10] G. T. Reed and C. E. J. Png. Silicon optical modulators. *Materials Today*, 8, 2005.
- [11] R. S. Jacobsen, K. N. Andersen, P. I. Borel, J. Fage-Pedersen, L. H. Frandsen, O. Hansen, M. Kristensen, A. V. Lavrinenko, G. Moulin, H. Ou, C. Peucheret, B. Zsigri, and A. Bjarklev. Strained silicon as a new electro-optics material. *Nature*, 441:199–202, 2006.

- [12] M. Cazzanelli, F. Bianco, E. Borga, G. Pucker, M. Ghuliyani, E. Degoli, E. Luppi, V. Veniard, S. Ossicini, D. Modotto, S. Wabnitz, R. Pierobon, and L. Pavesi. Second-harmonic generation in silicon waveguides trained by silicon nitride. *Nat. Mater.*, pages 148–154, 2011.
- [13] P. Dumon, G. Priem, L. R. Nunes, W. Begaerts, D. van Thourhout, P. Bienstman, T. K. Liang, M. Tsuchiya, P. Jaenen, S. Beckx, J. Wouters, and R. Baets. Linear and nonlinear nanophotonic devices based on silicon-on-insulator wire waveguides. *Jpn. J. Appl. Phys*, 45:6589–6602, 2006.
- [14] M. Dinu, F. Quochi, and H. Garcia. Third-order nonlinearities in silicon at telecom wavelengths. *App. Phys. Lett.*, 82:2954, 2003.
- [15] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, T. Tsuchizawa, T. Watanabe, J. I. Takahashi, and S. I. Itabashi. Four-wave mixing in silicon wire waveguides. *Opt. Express*, 13:4629–4637, 2005.
- [16] H. K. Tsang, C. S. Wong, T. K. Liang, I. E. Day, S. W. Roberts, A. Harpin, J. Drake, and M. Asghari. Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength. *Appl. Phys. Lett.*, 80:416–418, 2002.
- [17] O. Boyraz, T. Indukuri, and B. Jalali. Self-phase-modulation induced spectral broadening in silicon waveguides. *Opt. Express*, 12:829–834, 2004.
- [18] G. W. Rieger, K. S. Virk, and J. F. Young. Nonlinear propagation of ultrafast 1.5 μm pulses in high-index-contrast silicon-on-insulator waveguides. *Applied Physics Letters*, 84:900–902, 2004.
- [19] A. Cowan, G. Rieger, and J. Young. Nonlinear transmission of 1.5 μm pulses through single-mode silicon-on-insulator waveguide structures. *Opt. Express*, 12:1611–1621, 2004.
- [20] E. Dulkeith, Y. A. Vlasov, X. Chen, N. C. Panoiu, and Jr. R. M. Osgood. Self-phase-modulation in submicron silicon-on-insulator photonic wires. *Opt. Express*, 14:5524–5534, 2006.
- [21] I. W. Hsieh, X. Chen, J. I. Dadap, N. C. Panoiu, R. M. Osgood, S. J. McNab, and Y. A. Vlasov. Ultrafast-pulse self-phase modulation and third-order dispersion in silicon photonic wire-waveguides. *Opt. Express*, 14:12380–12387, 2006.
- [22] R. Dekker, A. Driessen, T. Wahlbrink, C. Moormann, J. Niehusmann, and M. Först. Ultrafast kerr-induced all-optical wavelength conversion in silicon waveguides using 1.55 μm femtosecond pulses. *Opt. Express*, 14:8336–8346, 2006.

- [23] O. Boyraz, P. Koonath, V. Raghunathan, and B. Jalali. All optical switching and continuum generation in silicon waveguides. *Opt. Express*, 12:4094–4102, 2004.
- [24] I. W. Hsieh, X. Chen, J. I. Dadap, N. C. Panoiu, Jr. R. M. Osgood, S. J. McNab, and Y. A. Vlasov. Cross-phase modulation-induced spectral and temporal effects on co-propagating femtosecond pulses in silicon photonic wires. *Opt. Express*, 15:1135–1146, 2007.
- [25] R. Espinola, J. Dadap, Jr. R. Osgood, S. McNab, and Y. Vlasov. C-band wavelength conversion in silicon photonic wire waveguides. *Opt. Express*, 13:4341–4349, 2005.
- [26] M. Foster, A. Turner, J. Sharping, B. Schmidt, M. Lipson, and A. Gaeta. Broad-band optical parametric gain on a silicon photonic chip. *Nature*, 441:960–963, 2006.
- [27] Q. Lin, J. Zhang, P. M. Fauchet, and G. P. Agrawal. Ultrabroadband parametric generation and wavelength conversion in silicon waveguides. *Opt. Express*, 14:4786–4799, 2006.
- [28] M. A. Foster, A. C. Turner, R. Salem, M. Lipson, and A. L. Gaeta. Broad-band continuous-wave parametric wavelength conversion in silicon nanowaveguides. *Opt. Express*, 15:12949–12958, 2007.
- [29] D. J. Moss, L. Fu, I. Littler, and B. J. Eggleton. Ultrafast all-optical modulation via two-photon absorption in silicon-on-insulator waveguides. *Electron. Lett.*, 41:320–321, 2005.
- [30] T. Liang, L. Nunes, T. Sakamoto, K. Sasagawa, T. Kawanishi, M. Tsuchiya, G. Priem, D. van Thourhout, P. Dumon, R. Baets, and H. Tsang. Ultrafast all-optical switching by cross-absorption modulation in silicon wire waveguides. *Opt. Express*, 13:7298–7303, 2005.
- [31] T. K. Liang, L. R. Nunes, M. Tsuchiya, K. S. Abedin, T. Miyazaki, D. van Thourhout, W. Bogaerts, P. Dumon, R. Baets, and H. K. Tsang. High speed logic gate using two-photon absorption in silicon waveguides. *Opt. Commun.*, 265:171–174, 2006.
- [32] R. Dekker, N. Usechak, M. Forst, and A. Driessen. Ultrafast nonlinear all-optical processes in silicon-on-insulator waveguides. *J. Phys. D: Appl. Phys.*, 40:R249–R271, 2007.
- [33] C. Corcoran, C. Monat, C. Grillet, D. J. Moss, B. J. Eggleton, T. P. White, L. O’Faolain, and T. F. Krauss. Green light emission in silicon through slow-light enhanced third-harmonic generation in photonic-crystal waveguides. *Nature Photon.*, 3:206–210, 2009.

- [34] S. Rein. *Lifetime Spectroscopy: A Method of Defect Characterization in Silicon for Photovoltaic Applications*. Springer, Verlag, 2005.
- [35] S. M. Sze and K. K. Ng. *Physics of Semiconductor Devices*. Wiley-Interscience, Hoboken, 2006.
- [36] P. Apiratikul, A. M. Rossi, and T. E. Murphy. Nonlinearities in porous silicon optical waveguides at 1550 nm. *Opt. Express*, 17:3396–3406, 2009.
- [37] K. Preston, P. Dong, B. Schmidt, and M. Lipson. High-speed all-optical modulation using polycrystalline silicon microring resonators. *Appl. Phys. Lett.*, 92:151104, 2008.
- [38] Y. Liu and H. K. Tsang. Nonlinear absorption and raman gain in helium-ion-implanted silicon waveguides. *Opt. Lett.*, 31:1714–1716, 2006.
- [39] F. E. Doany, D. Grischkowsky, and C. C. Chi. Carrier lifetime versus ion-implanted dose in silicon on sapphire. *Appl. Phys. Lett.*, 50:460, 1987.
- [40] T. Tanabe, K. Nishiguchi, A. Shinya, E. Kuramochi, H. Inokawa, M. Notomi, K. Yamada, T. Tsuchizawa, T. Watanabe, H. Fukuda, H. Shinojima, and S. Itabashi. Fast all-optical switching using ion-implanted silicon photonic crystal nanocavities. *Appl. Phys. Lett.*, 90:031115, 2007.
- [41] P. J. Foster, J. K. Doylend, P. Mascher, A. P. Knights, and P. G. Coleman. Optical attenuation in defect-engineered silicon rib waveguides. *J. Appl. Phys.*, 99:073101, 2006.
- [42] A. C. Turner-Foster, M. A. Foster, J. S. Levy, C. B. Poitras, R. Salem, A. L. Gaeta, and M. Lipson. Ultrashort free-carrier lifetime in low-loss silicon nanowaveguides. *Opt. Express*, 18:3582–3591, 2010.
- [43] V. R. Almeida, C. A. Barrios, R. R. Panepucci, and M. Lipson. All-optical control of light on a silicon chip. *Nature*, 431:1081–1084, 2004.
- [44] M. Först, J. Niehusmann, T. Plötzing, J. Bolten, T. Wahlbrink, C. Moormann, and H. Kurz. High-speed all-optical switching in ion-implanted silicon-on-insulator microring resonators. *Opt. Lett.*, 32:2046–2048, 2007.
- [45] S. A. Maier. *Plasmonics: Fundamentals and Applications*. Springer, New York, 2007.
- [46] J. S. Huang, T. Feichtner, P. Biagioni, and B. Hecht. Impedance matching and emission properties of nanoantennas in an optical nanocircuit. *Nano Letters*, 9:1897–1902, 2009.
- [47] P. Berini, R. Charbonneau, N. Lahoud, and G. Mattiussi. Characterization of long-range surface-plasmon-polariton waveguides. *J. Appl. Phys.*, 98:043109, 2005.

- [48] S. Sederberg, V. Van, and A. Y. Elezzabi. Monolithic integration of plasmonic waveguides into a complimentary metal-oxide-semiconductor- and photonic-compatible platform. *Appl. Phys. Lett.*, 96:121101, 2010.
- [49] M. Nielsen and A. Y. Elezzabi. Nanoplasmonic distributed bragg reflector resonators for monolithic integration on a complementary metal-oxide-semiconductor platform. *Appl. Phys. Lett.*, 103:051107, 2013.
- [50] I. De Leon and P. Berini. Amplification of long-range surface plasmons by a dipolar gain medium. *Nature Photon.*, 4:382–387, 2010.
- [51] O. Krupin, H. Asiri, C. Wang, R. N. Tait, and P. Berini. Biosensing using straight long-range surface plasmon waveguides. *Opt. Express*, 21:698–709, 2013.
- [52] S. E. Irvine, A. Dechant, and A. Y. Elezzabi. Generation of 0.4-keV femtosecond electron pulses using impulsively excited surface plasmons. *Phys. Rev. Lett.*, 93:184801, 2004.
- [53] S. E. Irvine and A. Y. Elezzabi. Ponderomotive electron acceleration using surface plasmon waves excited with femtosecond laser pulses. *Appl. Phys. Lett.*, 86:264102, 2005.
- [54] P. Dombi, S. E. Irvine, P. Racz, M. Lenner, N. Kroo, G. Farkas, A. Mitrofanov, A. Baltuska, T. Fuji, F. Krausz, and A. Y. Elezzabi. Observation of few-cycle, strong-field phenomena in surface plasmon fields. *Opt. Express*, 18:24206–24212, 2010.
- [55] G. Herink, D. R. Solli, M. Guide, and C. Ropers. Field-driven photoemission from nanostructures quenches the quiver motion. *Nature*, 483:190–193, 2012.
- [56] P. M. Nagel, J. S. Robinson, B. D. Harteneck, T. Pfeifer, M. J. Abel, J. S. Prell, D. M. Neumark, R. A. Kaindl, and S. R. Leone. Surface plasmon assisted electron acceleration in photoemission from gold nanopillars. *Chem. Phys.*, 414:106–111, 2013.
- [57] P. Dombi, A. Horl, P. Racz, I. Marton, A. Trugler, J. R. Krenn, and U. Hohenester. Ultrafast strong-field photoemission from plasmonic nanoparticles. *Nano Letters*, 13:674–678, 2013.
- [58] P. A. Wolf. Theory of electron multiplication in silicon and germanium. *Phys. Rev.*, 95:1415–1420, 1954.
- [59] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Appl. Opt.*, 35:1956–1976, 1996.

Chapter 3

Nanofabrication Process Development

3.1 Introduction

Nanofabrication process development made up a significant portion of the work that went into producing the results presented in this thesis. Since different aspects of the device characterization will be discussed in separate chapters, this chapter serves as a central location for all of the nanofabrication processes developed, and the physical specifications of the devices. It should be noted that this chapter focuses mainly on the development and use of the nanofabrication processes and not the fundamental process mechanisms. A brief description of each nanofabrication technique is presented in Appendix B, and where possible, references to comprehensive review papers have been listed for the benefit of the reader.

This chapter begins with a description of the silicon-on-insulator (SOI) substrates used and then proceeds to discuss four main fabrication process flows that were developed. The first (and most involved) is the integration of nanoplasmonic waveguides onto a large “characterization beam.” Many of the measures taken to optimize the processes are shown in detail in this section. The subsequent sections then present a brief summary of the processes used to fabricate silicon (Si) photonic waveguides, integrate Si photonic waveguides with nanoplasmonic waveguides, and fabricate Si photonic resonators on the corner of a die.

3.2 Silicon-on-Insulator Substrates

As discussed in Section 2.4, Si is a highly attractive material for nanophotonic devices [1]. High mode confinement and low propagation losses in a Si waveguide requires a high index contrast between the Si waveguide core and the surrounding environment, which consists of a substrate and a cladding. Simply etching photonic waveguides into a Si wafer effectively exposes three sides of the waveguide to a low refractive index material (air), but this is not sufficient confinement for waveguiding, and the modes will rapidly leak

into the substrate. In principle, a low refractive index environment could be created on the bottom interface of the waveguide by masking the waveguide and isotropically etching the underside of the waveguide, so that all four sides were exposed to air. However, this adds a high level of complexity to the fabrication and severely compromises the structural integrity of the waveguide.

Achieving high mode confinement is simplified immensely by considering a substrate comprised of three layers: (1) the handle wafer ($n = 3.48$); (2) the buried oxide (BOX) layer ($n = 1.45$); and (3) the device layer ($n = 3.48$). Several considerations must be made when choosing the different layer thicknesses. Firstly, the device layer thickness, t_{dev} , should be thick enough that a large fraction of the mode energy at $\lambda = 1550\text{nm}$ propagates inside the Si core, but thin enough to allow for single-mode operation. In addition, the BOX layer should be thick enough to prevent mode leakage into the handle wafer. In this thesis, a device layer thickness of $t_{dev} = 340\text{nm}$ and a BOX layer thickness of $t_{BOX} = 1\mu\text{m}$ were found to satisfy these requirements and all fabrication processes were performed on these substrates.

The simultaneous requirement of a thin device layer and thick BOX layer severely limits the possible wafer fabrication approaches and very few suppliers provide wafers with the required dimensions. The wafers used in this thesis were purchased from Soitec, and the wafers were produced using the "Smart Cut" process [2, 3]. Briefly, the process begins with two standard silicon wafers. A thermal oxide is grown on one wafer, and the thickness of this oxide determines the t_{BOX} dimension. Hydrogen implantation is then performed on the oxidized wafer, where the ions travel through the oxide and into the silicon, and the implantation energy is designed so that the stopping distance of the ions in the Si is slightly greater than the desired device layer thickness. The oxide side of the wafer is then bonded to the second silicon wafer and the wafers are subjected to a heat treatment that fractures the wafer along the plane in which the hydrogen ions stopped, leaving one wafer with the SOI structure and the other wafer as bare Si. The fractured interface is then polished using chemical-mechanical polishing.

A cross-sectional scanning electron micrograph (SEM) of a SOI wafer with an additional $t = 465\text{nm}$ oxide film deposited via plasma-enhanced chemical vapour deposition (PECVD) is shown in Fig. 3.1. Each layer thickness is confirmed to match the manufacturer's specifications. Imaging artifacts due to electron charging on the top insulating layer are observed in this image.

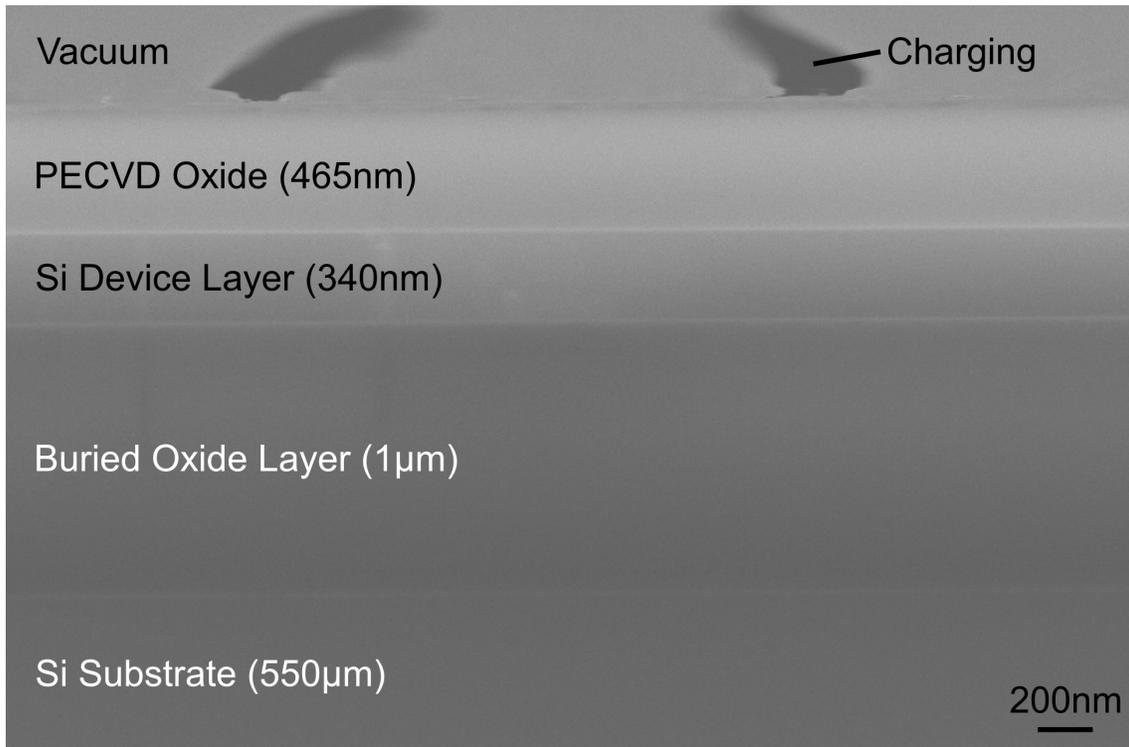


Figure 3.1: Cross-sectional SEM of a cleaved SOI wafer with a 465nm thick PECVD oxide.

3.3 Integration of Nanoplasmonic Waveguides onto a Micron-Scale Characterization Beam

3.3.1 Background

As discussed in the previous chapter, the strong electric field confinement of nanoplasmonic structures produces an enhancement in the amplitude of the electric field in the vicinity of the metal features. In the case of a nanoplasmonic waveguide, fields can be enhanced by several times their original strength [4]. The enhanced fields can be used to magnify light-matter interaction processes, such as chemical detection or nonlinear optical processes. While plasmonic waveguides have several attractive characteristics, strong confinement of electromagnetic energy to the metal interfaces results in Ohmic losses and propagation losses of plasmonic waveguides are considerably higher than photonic waveguides [5]. The characteristic propagation length of electromagnetic energy in a subwavelength plasmonic waveguide depends strongly on the materials and design of the waveguide and can range from several microns to several tens of microns [6].

Therefore, optimum coupling of radiation into nanoplasmonic waveguides with micron-scale lengths requires a carefully designed on-chip coupler, which can add additional fabrication complexities [7, 8]. Previous coupling schemes have incorporated prisms [9], photonic bus waveguides [10], gratings [11], plasmonic antennas [12], and nano-mirrors

[13] to direct the free-space radiation to the nanoplasmonic waveguide mode.

However, a key predicament in accessing ultrafast active operation in a nanoplasmonic waveguide device is the ability to couple the optical signal directly into the device. That is, when considering nonlinear optical interaction effects in Si-based nanoplasmonic waveguides, it is not possible to achieve efficient nonlinear interaction in the plasmonic region, as nonlinear interaction can occur within the on-chip coupler. End-fire coupling radiation directly into the nanoplasmonic waveguide would eliminate the need for a coupler, but requires separation of the substrate into micron-scale die, which would make subsequent handling and characterization very challenging. Moreover, these dimensional constraints push traditional die separation techniques, such as dicing or scribing beyond their prescribed abilities. Therefore, it would be beneficial to develop a means to integrate nanoplasmonic waveguides with lengths of several microns onto a die in a configuration that would allow their end-facets to be directly accessed with macroscopic objects such as microscope objectives and optical fibres. Realization of such a sample would enable efficient access to nonlinear optical effects and ultrafast modulation in Si-loaded nanoplasmonic waveguides.

3.3.2 Device Design

The waveguide geometry consists of a Si waveguide core with cross-sectional dimensions, $w \times h = 100\text{nm} \times 340\text{nm}$. The Si core is capped by a gold (Au) film with a thickness, $t = 60\text{nm}$. The cross-sectional geometry of the waveguide is depicted schematically in Fig. 2.15(a) and the corresponding nanoplasmonic mode profile was obtained with an electromagnetic mode solver¹ are formulated and is shown in Fig. 2.16(a). The theoretical distance propagated by the mode before attenuating to e^{-1} of its initial amplitude was found to be $L_{prop} = 3.1\mu\text{m}$. To ensure a strong signal is transmitted through the device, the waveguide lengths are designed such that $L \leq 10\mu\text{m}$.

In this process, Si-loaded nanoplasmonic waveguides are fabricated and integrated onto a micron-scale characterization beam. This characterization beam allows for direct coupling to the waveguides using microscope objectives and optical fibres and enables accurate measurements of ultrafast nonlinear interactions taking place in the waveguide. A schematic depiction of several nanoplasmonic waveguides integrated onto the characterization beam is shown in Fig. 3.2. In this design, the width of the characterization beam is increased in steps, allowing the waveguide length to be varied. As labeled in Fig. 3.2, one face of the characterization beam is defined by cleaving the SOI wafer and the other is defined by dry etching the Si. A deeply etched line is used to define the cleavage plane, ensuring that the cleave intersects the waveguide end-facet. The implementation of this technique is discussed further in Section 3.3.3. A cleaved facet allows for a bulky

¹The mode profiles were obtained using Lumerical FDTD Solutions, where the waveguide geometry is meshed using the finite difference algorithm, a matrix eigenvalue problem based on Maxwell's equations is setup, and the matrix is solved using sparse matrix techniques.

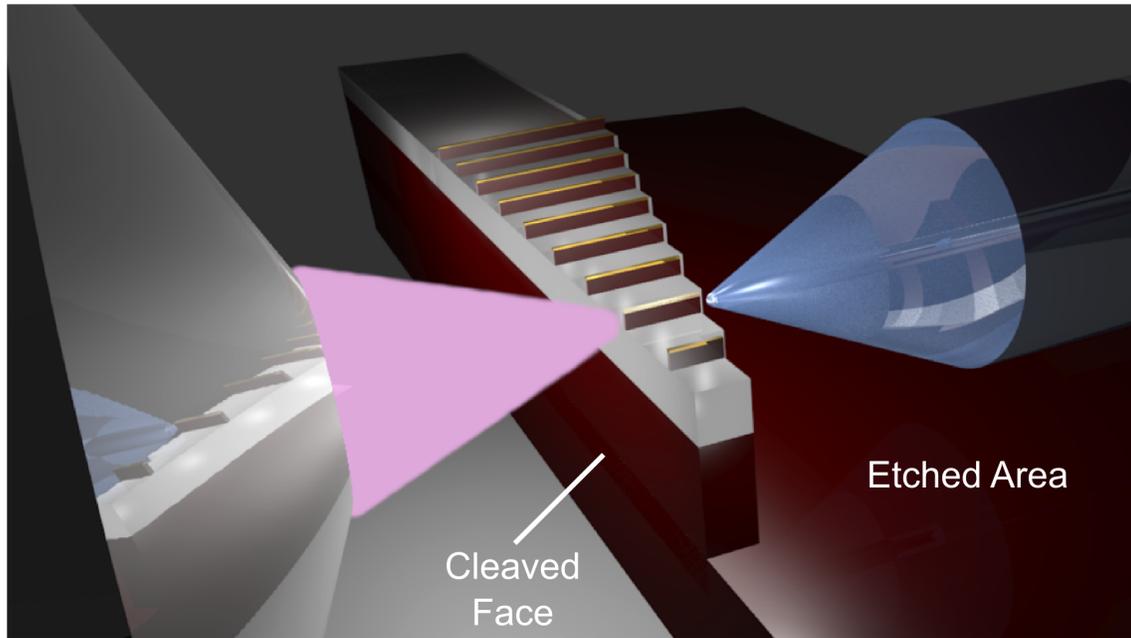


Figure 3.2: Rendering of the envisioned sample and excitation scheme. Integrating nanoplasmonic devices onto a large characterization beam would allow for direct probing and detection with macroscopic optical components, including microscope objectives and optical fibres.

microscope objective to be brought in close proximity to the sample for excitation. By considering an etching depth, $d_{etch} = 75\mu\text{m}$, an optical fibre can be brought over the substrate to the output facet of the waveguide. The proposed excitation and detection scheme is depicted schematically in Fig. 3.2. Notably, such a structure combines fabrication techniques designed for length-scales ranging from sub-100nm to $\sim 100\mu\text{m}$, which introduced significant challenges to the process development and largely shaped the final process flow.

3.3.3 Fabrication Process Flow

Nanoplasmonic Waveguide Definition

The purpose of the first layer of lithography was to define the nanoplasmonic waveguides and to protect them for subsequent deep etching. Organic contamination was removed from the sample by placing it in an ultrasonic bath of acetone for 10 minutes, followed by thorough rinsing in deionized water, followed by a 15 minute immersion in Piranha solution (3:1 mixture of 96% H_2SO_4 to 30% H_2O_2). The native oxide was removed by a 45 second bath in buffered hydrofluoric acid (1 HF:10 NH_4F). A schematic depiction of the bare SOI substrate is shown in Fig. 3.3(a).

In preparation for the first layer of electron beam lithography (EBL), a layer of 6% 495K poly(methyl methacrylate) in anisole (495k PMMA A6) was spun onto the substrate at 5000 revolutions per minute (RPM), producing a resist thickness of 275nm. The sample

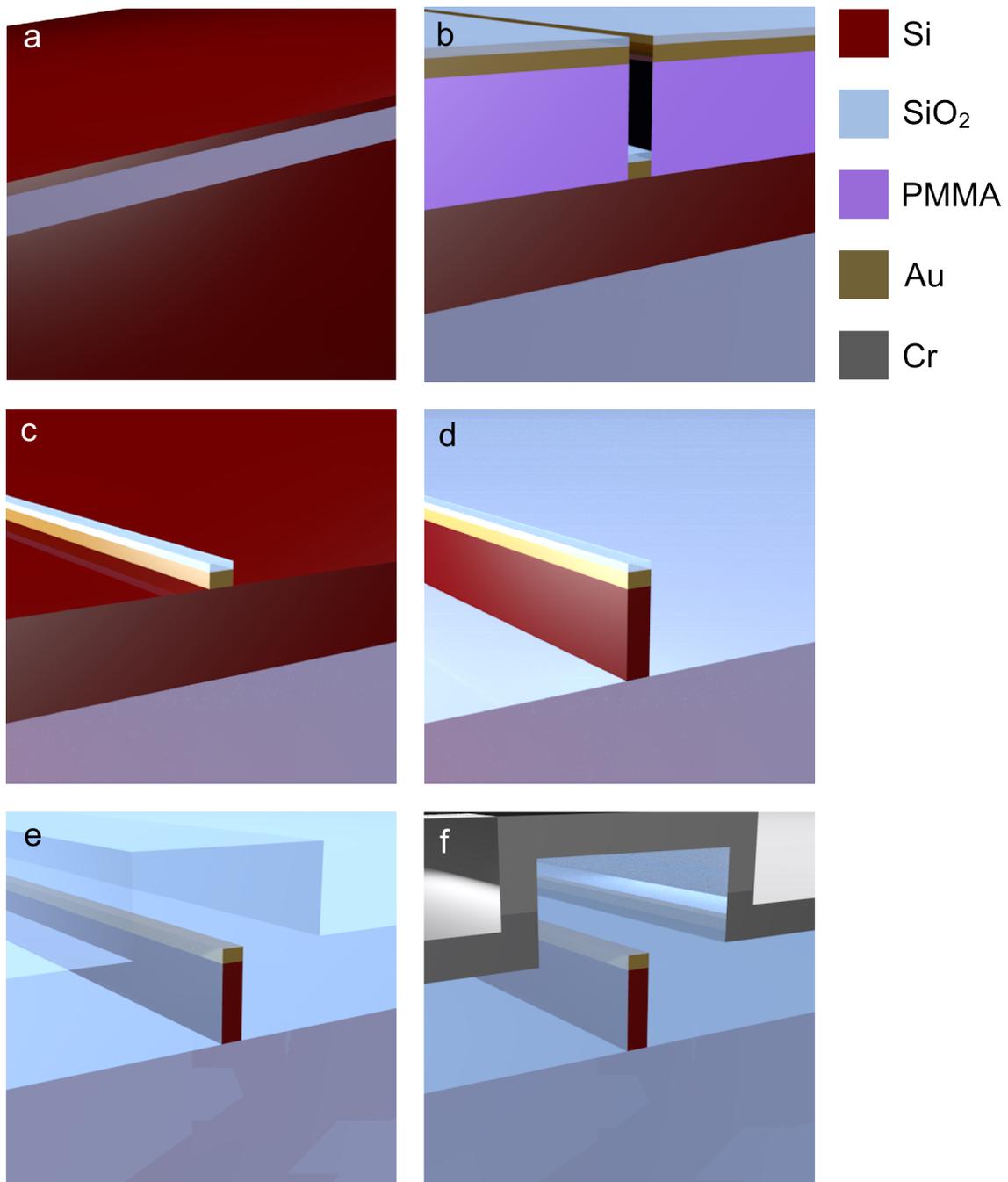


Figure 3.3: Nanofabrication process flow for the definition of nanoplasmonic waveguides with a Si core and Au cap. (a) Cleaned SOI wafer. (b) Electron-beam lithography exposure, followed by electron beam evaporation of Cr(5nm) + Au(60nm) + Cr(5nm) + SiO₂(40nm). (c) Evaporated films after lift-off. (d) A fully-defined nanoplasmonic waveguide after ICPRIE etching through the Si device layer. (e) A nanoplasmonic waveguide clad in a 465nm thick SiO₂. (f) Subsequent deposition of 150nm thick Cr etch mask.

was then baked at 180°C for 15 minutes.

The EBL pattern was exposed using an aperture diameter, $d_{EBL} = 20\mu\text{m}$, an acceleration voltage, $V = 20\text{keV}$, and an area dosage, $D = 450\mu\text{C}/\text{cm}^2$. The sample was developed in a methylisobutylketone (MIBK)-based developer (1 MIBK : 3 isopropyl alcohol (IPA)) for 45 seconds. Four subsequent thin films were deposited on the sample using an electron beam evaporation system. The first layer was a 5nm chromium (Cr) adhesion layer. The second was a 60nm Au layer, which gave the waveguide its ability to function as a subwavelength nanoplasmonic device. The third was an additional 5nm Cr adhesion layer, and the final layer was a 40nm SiO_2 layer, which acted as an etch mask to prevent sputtering of the soft Au features during subsequent plasma processing. The EBL pattern along with the deposited films is shown schematically in Fig. 3.3(b). After deposition, a standard lift-off in an ultrasonic bath of acetone was performed, and the sample was cleaned in this manner for 20 minutes. A nanoplasmonic waveguide after the lift-off step is shown in Fig. 3.3(c).

The pattern was transferred through the Si device layer using an unswitched recipe in an inductively coupled plasma reactive ion etching (ICPRIE) system, shown in Fig. 3.3(d). Next, a 465nm layer of SiO_2 was deposited via PECVD, as shown in Fig. 3.3(e). A Cr layer with a thickness of 150nm was deposited on the sample via magnetron sputtering, as shown in Fig. 3.3(f), and was used to define an etch mask.

Characterization Beam Definition

In preparation for the second layer of EBL, a 275nm layer of 495K PMMA A6 was spun onto the sample and baked for 15 minutes at 180°C. The exposure time is reduced by using a large electron beam aperture diameter, $d_{EBL} = 120\mu\text{m}$, an acceleration voltage, $V = 20\text{keV}$, and an area dose of $D = 200\mu\text{C}/\text{cm}^2$. The sample was developed in 1 MIBK : 3 IPA developer for 45 seconds.

The purpose of the second layer of lithography was to define the characterization beam, as shown in Fig. 3.2. One side of the beam was defined by deep etching, while the other side was defined by cleaving. The mask defined in this layer of lithography consisted of two main features: (1) a large area exposed adjacent to one end facet of the waveguides (labeled “Etched Area” in Fig. 3.2); and (2) a line intersecting the waveguide end-facets, which acts as a guide for cleaving the sample with a high accuracy.

The pattern was transferred through the Cr layer using a commercially available solution of ceric ammonium nitrate ($(\text{NH}_4)_2\text{Ce}(\text{NO}_3)_6$), nitric acid (HNO_3), and water (H_2O) and the sample was cleaned in an ultrasonic bath for 20 minutes. Figure 3.4(a) shows a schematic of a nanoplasmonic waveguide that is buried beneath an oxide cladding and a Cr etch mask. Figure 3.4(b) shows the Cr etching mask that has been defined via EBL and transferred through the Cr layer via wet etching.

The Cr mask was then used to etch through the oxide layer (465nm PECVD oxide and

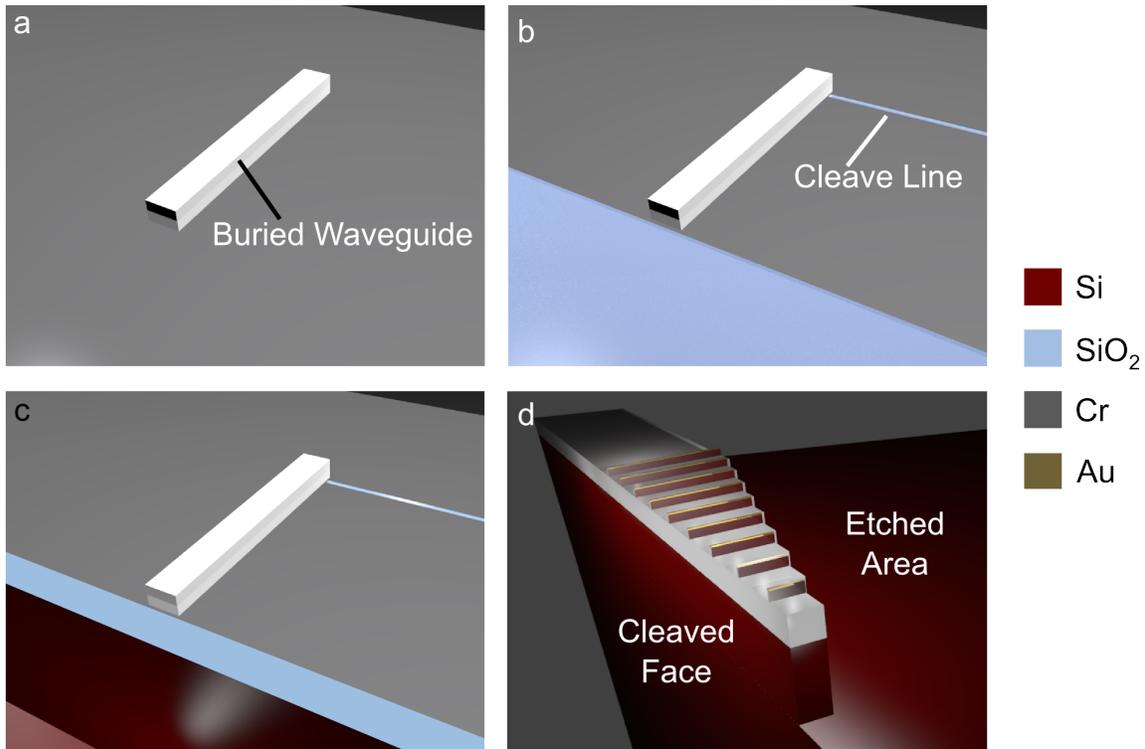


Figure 3.4: (a) Schematic of a nanoplasmonic waveguide buried beneath a 465nm SiO₂ cladding and a 150nm Cr etch mask. (b) The Cr etch mask consisting of a cleave guide adjacent to one waveguide end facet and a large area adjacent to the other end facet. (c) Etching through the BOX layer and 75 μ m into the Si substrate. (d) Schematic of the final structure consisting of nanoplasmonic waveguides integrated onto a characterization beam. The SiO₂ cladding has been omitted for clarity.

1000nm thermal oxide), via reactive ion etching (RIE), and to a depth of $d_{etch} = 75\mu\text{m}$ into the Si wafer using the Bosch process in an ICPRIE system [14]. A schematic view of the sample after Bosch etching is shown in Fig. 3.4(c). At this point, the Cr mask was removed and the sample was cleaved and cleaned in an ultrasonic bath for 20 minutes. A schematic of the final sample geometry is shown in Fig. 3.4(d).

Lift-off Development

The choice of deposition technique to use for metallization was an important factor to consider. Although magnetron sputtering is a very effective deposition technique for many purposes, it has one main draw-back for lift-off: its poor directionality. This results in a significant thickness of the deposition material being deposited on the resist sidewalls. If the film is thick enough, this can prevent large areas of the film from lifting off. Even for thinner films, it can result in lift-off artifacts, such as “rabbit ears” or “whiskers.” While this may not be a problem for large features or certain device applications, it is a serious concern for nanoscale features, where the dimensions of the lift-off artifacts can be comparable to the feature dimensions themselves. It should be noted that several advanced

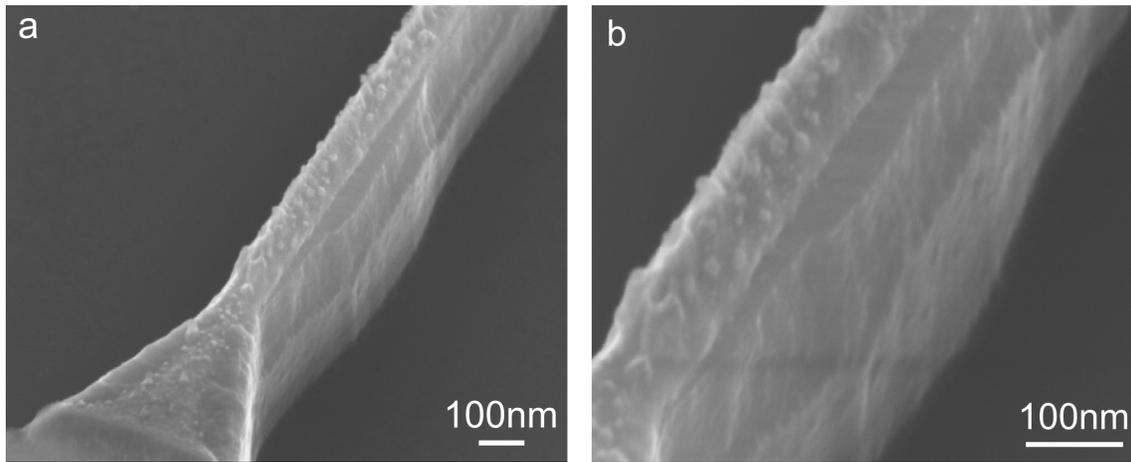


Figure 3.5: (a) Optimization of the Cr adhesion layer thickness between the Si device layer and the Au nanoplasmonic waveguide cap. (a) Marginal adhesion with a 2.5nm thick Cr adhesion layer, where most of the Au cap adheres to the Si core, but other regions are detached. (b) Close-up of a region of poor adhesion.

processing techniques have been developed to circumvent these artifacts [15]. In the interest of maintaining a simple process and minimizing any possible damage or contamination to the structures, a directional deposition technique was used, which enabled artifact-free features with sub-100nm line widths and aspect ratios of ~ 1 . Therefore, wherever lift-off feature dimensions of ~ 100 nm were required, electron beam evaporation was used as a deposition technique.

Due to the poor adhesion of Au to Si, it was necessary to include a thin adhesion layer to ensure a high-yield process. Typical materials used for this purpose are Cr and titanium. Both of these materials have very high optical losses at $\lambda = 1550$ nm, and therefore the optical interaction of the $\lambda = 1550$ nm radiation with the adhesion layer must be minimized by reducing the adhesion layer thickness as much as possible. Several adhesion layer thicknesses in the range of 0-5nm were tested to determine the thinnest possible Cr adhesion layer that would effectively adhere ~ 60 nm Au features to Si. Figure 3.5 shows an SEM from a structure that was fabricated with 2.5nm thick Cr adhesion layer, after subsequent Si etching was performed. While the entire Au structure is intact, careful examination of the sample revealed that certain areas of the Au film were deformed from the Si surface and were not adhered properly. For thinner adhesion layers, the lift-off yield was extremely low. The Cr thickness had to be increased to 4-5nm to obtain a $\sim 100\%$ yield in the lift-off process.

It was also crucial to optimize the EBL exposure dose for obtaining a high-yield lift-off process. Any remaining resist fragments could result in a complete failure of the lift-off, or compromise the device performance. Several EBL exposures were performed on the same sample at different nominal dosages of $100\mu\text{C}/\text{cm}^2$, $150\mu\text{C}/\text{cm}^2$, $200\mu\text{C}/\text{cm}^2$, and $300\mu\text{C}/\text{cm}^2$. At each nominal dose, 21 dose multipliers were applied to different individual structures, so that a precise value of the minimum dose required for a high-

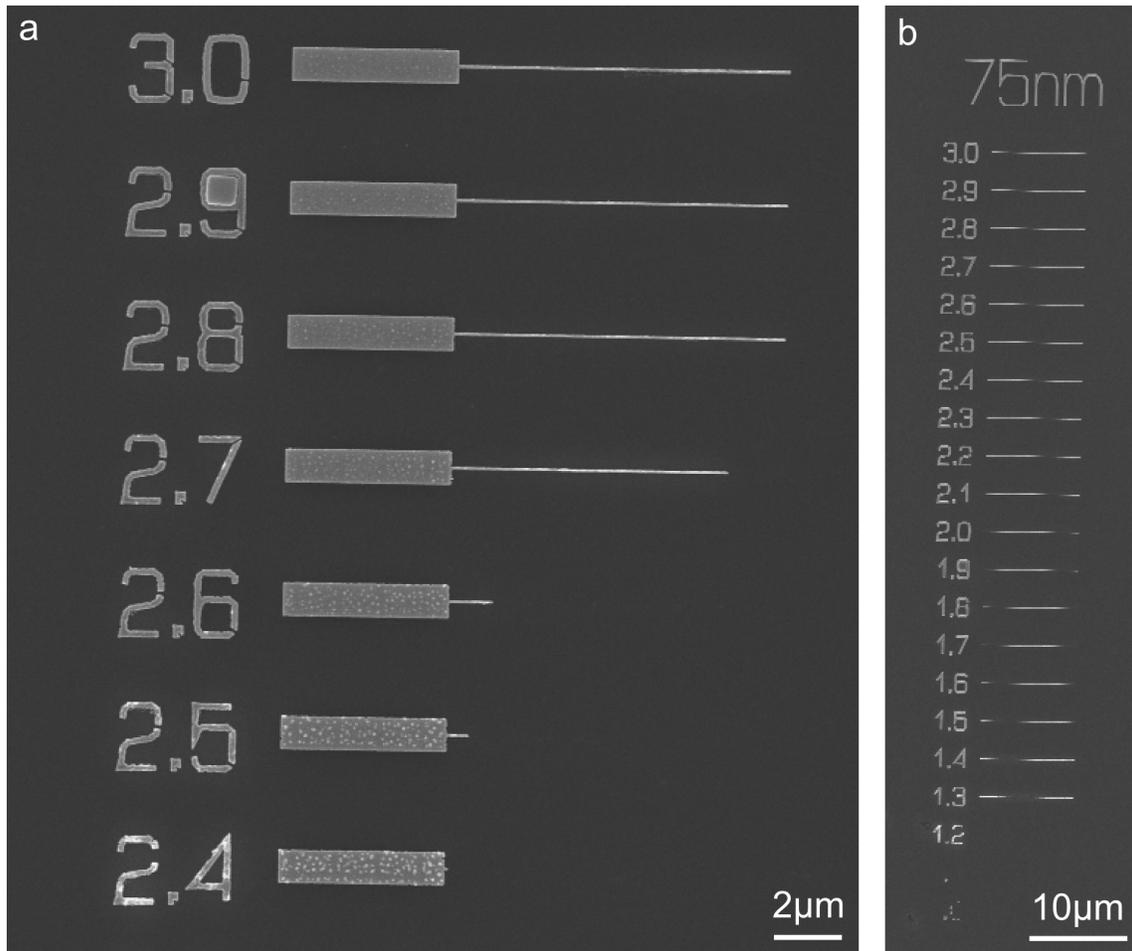


Figure 3.6: Optimization of the EBL exposure for a 495k PMMA A6 resist with a thickness of 275nm, exposed at $V = 20\text{keV}$ and $d_{ap} = 20\mu\text{m}$. (a) Dose test for a nominal dose of $150\mu\text{C}/\text{cm}^2$; and (b) $300\mu\text{C}/\text{cm}^2$. Electron beam evaporation was used to deposit Cr(5nm) + Au(60nm) + Cr(5nm) + SiO₂(40nm) layers and a standard lift-off was performed in acetone. A final dose of $D = 450\mu\text{C}/\text{cm}^2$ was used to obtain a very high $\approx 100\%$ yield lift-off.

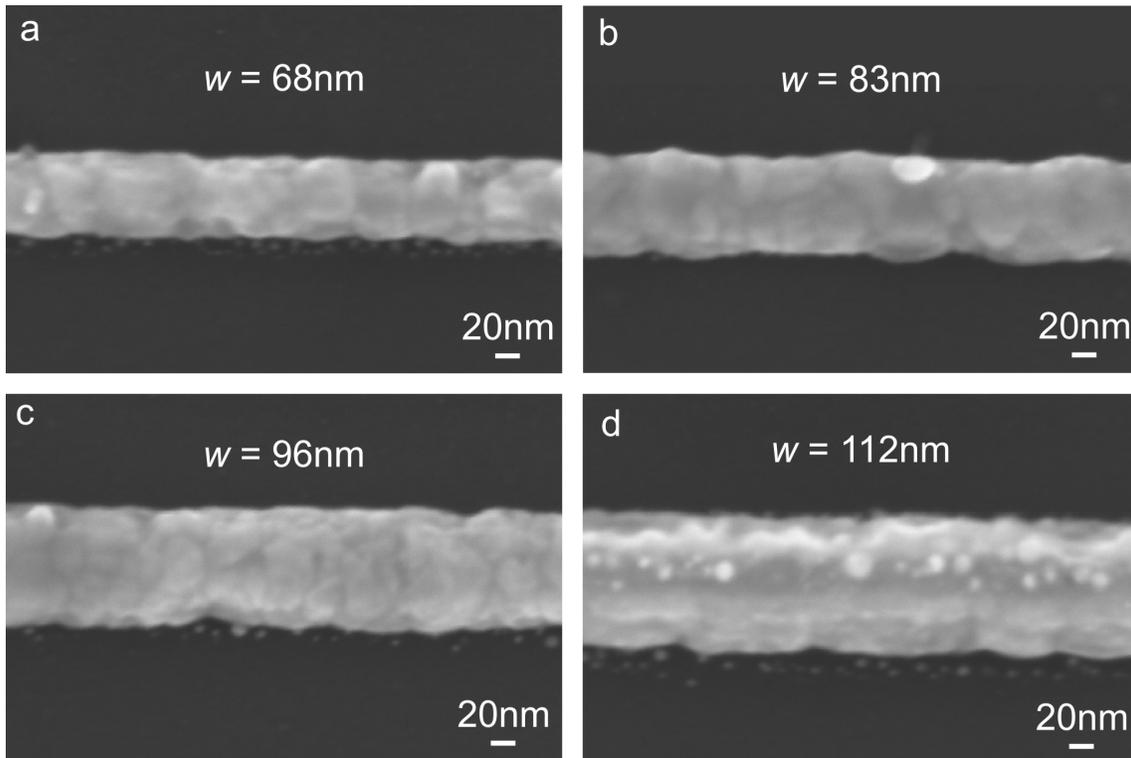


Figure 3.7: Optimization of the design pattern rectangle width required to obtain a lift-off feature with a width of 100nm, at an exposure dose, $Dose = 450\mu\text{C}/\text{cm}^2$. Several widths were obtained during this process, including: (a) $w = 68\text{nm}$, (b) $w = 83\text{nm}$, (c) $w = 96\text{nm}$, and (d) $w = 112\text{nm}$.

yield process could be obtained. Two different sets of exposure test results are shown in Figs. 3.6(a),(b). As the waveguide dimensions were decreased, a higher dose was required, and for the desired waveguide width of $w = 100\text{nm}$, an exposure dose of $450\mu\text{C}/\text{cm}^2$ was found to be optimal.

Due to the thickness of the resist (275nm), the dimensions of the fabricated structures did not always match the dimensions of the structures created in the exposure design software. Therefore, it was necessary to “calibrate” the width of the designed feature that would produce a fabricated structure width of 100nm. It was found that a width of $w = 75\text{nm}$ in the exposure design software yielded a fabricated structure width of $w = 100\text{nm}$ with a PMMA resist thickness of 275nm and an exposure dose of $D = 450\mu\text{C}/\text{cm}^2$. Figure 3.7 depicts several SEMs of different line widths obtained through this calibration process. High magnification SEMs of lift-off structures obtained with the final exposure settings are shown in Fig. 3.8.

Device Layer Etch Development

Another critical aspect of the fabrication process development was etching through the Si device layer. Due to the nanoscale feature dimensions and their aspect ratio of $\approx 3.4 : 1$, isotropic wet etching was immediately eliminated. Instead, dry etching in an ICPRIE

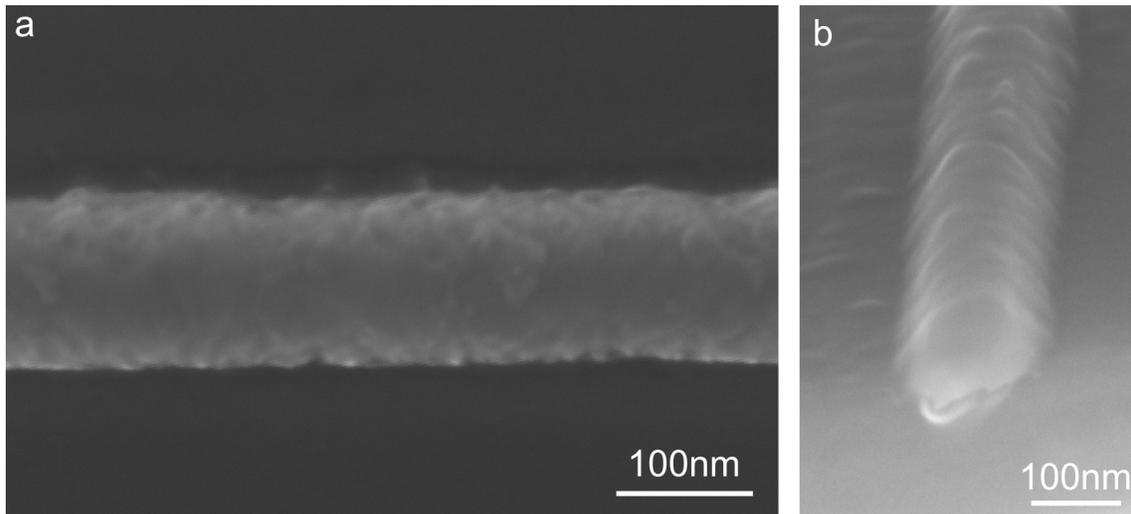


Figure 3.8: Lift-off features with $w \approx 100\text{nm}$ obtained using the final exposure parameters: (a) Top-view; and (b) 45° tilt-view of a cleaved lift-off feature, illustrating the topography.

system was pursued. Notably, a switched Bosch process produces sidewall scallops with a spatial period in the range of 100-500nm, depending on the etching conditions. Such large scallops are on the order of the desired device dimensions, and would result in very high optical scattering losses. In order to reduce the size of the scallops, one could simply reduce the temporal period of the gas switching in the Bosch process. The limiting case where the period approaches zero represents an unswitched recipe, and the dimensions of the sidewall scallops approach zero, which is ideal for waveguides. Therefore, an unswitched Bosch-like process was pursued, where the sample was exposed to C_4F_8 , SF_6 , and O_2 simultaneously.

The etching results were highly sensitive to the mask materials and the materials adjacent to the etching surfaces. The etching recipes were initially developed to be used with a PMMA mask. However, when the same recipes were applied to the SiO_2 mask with the underlying Au features, severe undercutting was observed and the structures were completely destroyed, as shown in Fig. 3.9. Subsequent investigations revealed that metallic masks can lead to substantial undercut in the etching profile, and that additional passivation was necessary [16]. Although the Au features were capped by a 40nm SiO_2 etch mask, it was found that even etching in close proximity to Au features produced substantial undercut, and it was necessary to increase the C_4F_8 passivation gas flow from 40sccm (which was found to be optimal for a PMMA mask) to 80sccm. The SF_6 and O_2 gas flow rates were kept constant at 85sccm and 5sccm, respectively, throughout the optimization. The redevelopment of this etching process is presented in Fig. 3.10 and typical results obtained with the final etching conditions are shown in Fig. 3.11. Virtually all sidewall roughness was eliminated and only small “dimples” with a diameter of $\sim 15\text{nm}$ and an estimated depth of $\sim 3-5\text{nm}$ were observed on the waveguide sidewalls.

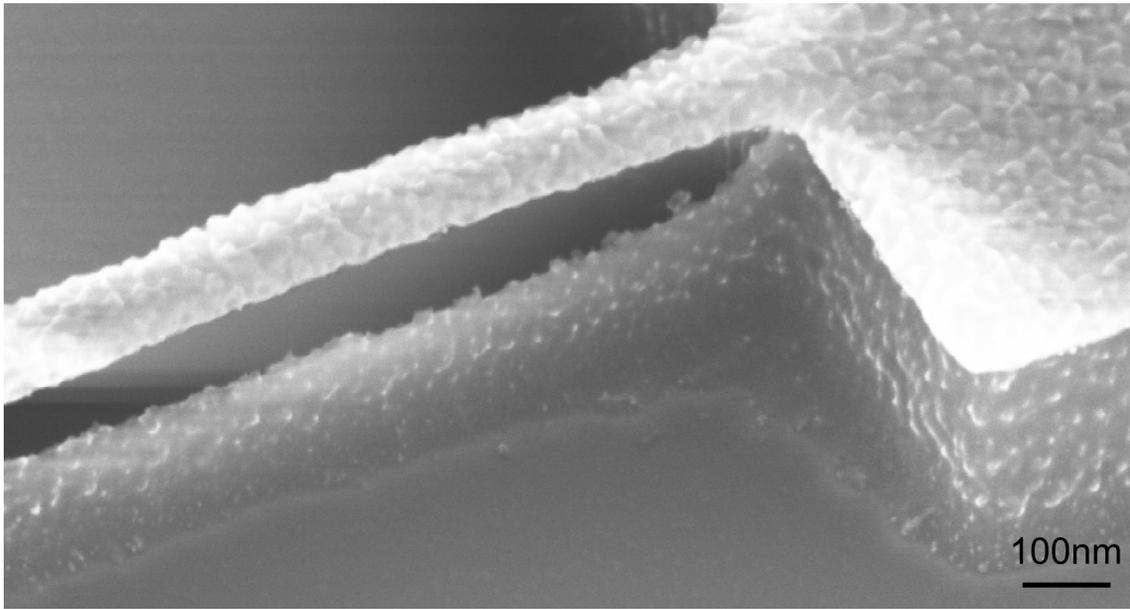


Figure 3.9: Example of severe under-cut obtained when using a $\text{SiO}_2 + \text{Au}$ etch mask for a process that was designed for a PMMA mask.

Further modification of the etching process did not eliminate these dimples, and additional passivation produced an etching profile that was flared outwards. Therefore, it is believed that these are the best possible etching results achievable with this combination of sample and etching system.

Cladding and Etch Mask Deposition

The nanoplasmonic waveguides were clad in SiO_2 for several reasons. The primary function of the oxide cladding was to protect the nanoplasmonic waveguides from the deep Si etching that was performed to define the characterization beam. By cladding the waveguides in a 465nm thick oxide and then using a 150nm thick Cr etching mask, the waveguides were completely isolated from the etching environment and showed no signs of damage. It was very important that both the oxide and Cr deposition films were conformal. Otherwise, the etching gases were able to interact with the nanoplasmonic structures and severely damage them over many etching cycles.

An additional function of the cladding was to prevent the nanoplasmonic waveguides from detaching from the characterization beam when the sample was cleaved. Notably, some of the waveguides had lengths as short as $\sim 2\mu\text{m}$, and it is anticipated that the strain created by the cleave would have easily detached or disintegrated the waveguides. Furthermore, the cladding isolated the nanoplasmonic waveguides from the surrounding environment during the experiments and prevented any particles from affecting the device performance. Finally, because etching had to be performed through oxide and silicon layers, a metal mask had to be deposited, and eventually removed from the sample with a wet etchant. Therefore, the oxide cladding also prevented chemical damage to the Cr

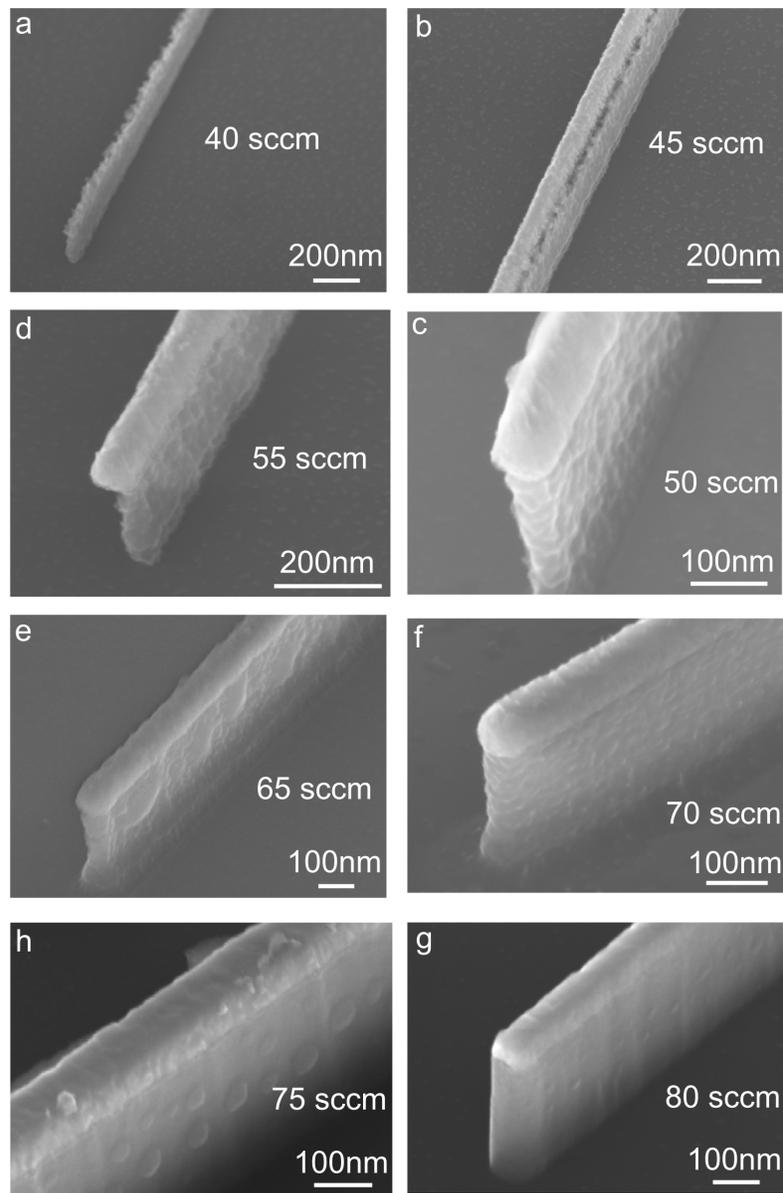


Figure 3.10: Redevelopment of the unswitched Si Bosch-like process used for etching through the Si device layer. All parameters were kept fixed, except for the C_4F_8 gas flow. Scanning-electron micrographs are shown for C_4F_8 flow rates of: (a) 40sccm, (b) 45sccm, (c) 55sccm, (d) 60sccm, (e) 65sccm, (f) 70sccm, (g) 75sccm, and (h) 80sccm. Subsequent increases in the C_4F_8 flow rate did not yield any improvement in the sidewall roughness.

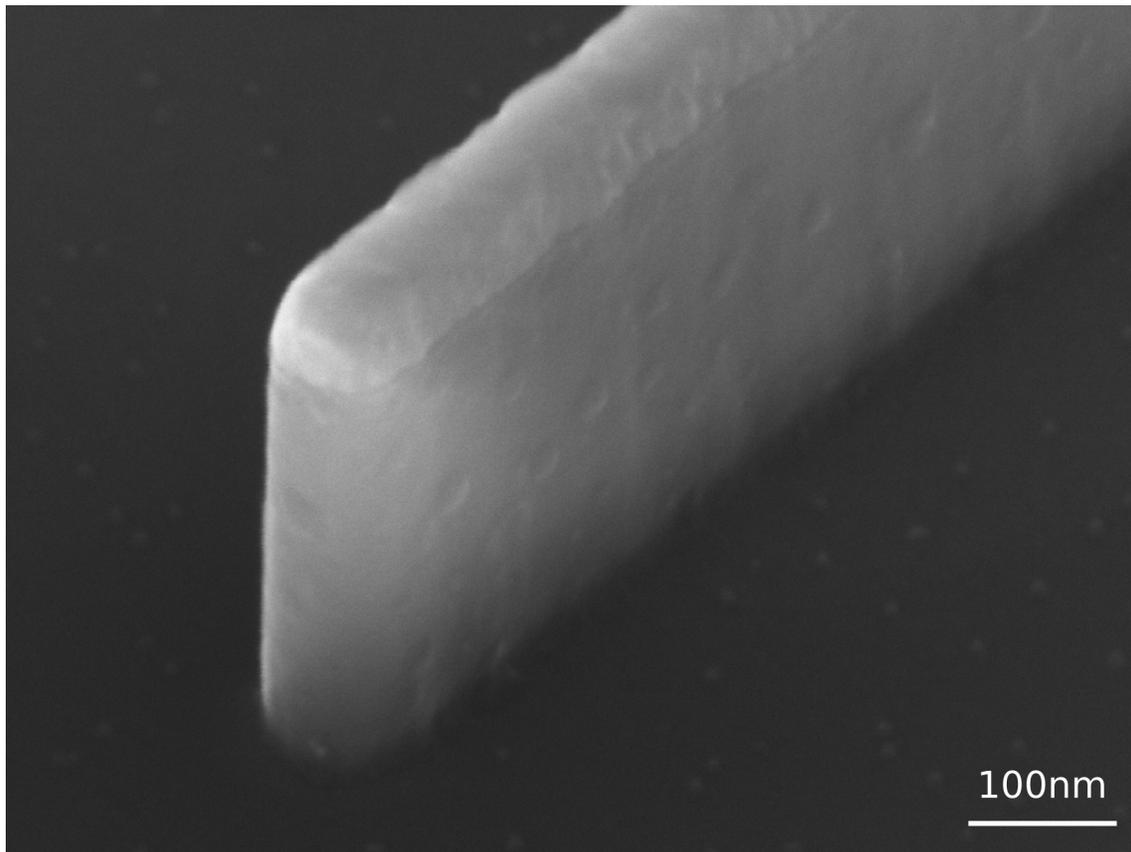


Figure 3.11: Typical etching results obtained with the redeveloped etching process, with a C_4F_8 flow rate of 80sccm.

adhesion layer or the Au cap during the Cr mask removal.

Figure 3.12(a) depicts a top-view SEM of several nanoplasmonic waveguides buried in a 465nm thick SiO_2 cladding and a close-up of a single clad waveguide is shown in Fig. 3.12(b). A single nanoplasmonic waveguide coated in both a 465nm thick SiO_2 film and 150nm Cr film is shown in Fig. 3.12(c). Clearly, the waveguide is completely isolated from the surrounding environment, and the film quality does not appear to be compromised in any way.

Characterization Beam Etching Steps

Prior to fabricating the characterization beam, significant work had to be done to adapt a Bosch process to be suitable for this application. Initial tests were performed with a radio-frequency (RF) power of 3000W, a SF_6 flow rate of 350sccm, an etching cycle duration of 4s, a C_4F_8 flow rate of 200sccm, and a passivation cycle duration of 1.3s. This recipe performed well for shallow etching on the order of $\sim 10\mu m$. However, for etching depths of 75-100 μm , the top $\sim 10\mu m$ of the sidewalls became disfigured due to repeated cycles of etching and passivation, and it became clear that the recipe was too aggressive for the current application. Typical results obtained with these parameters are shown

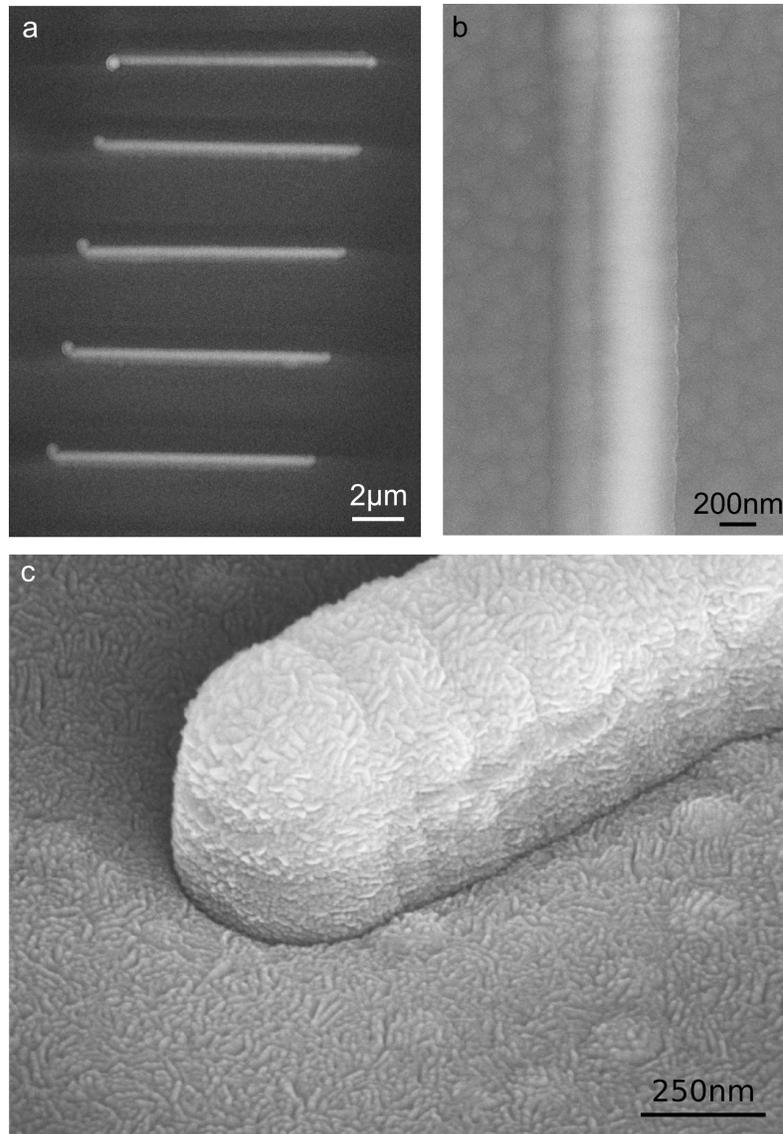


Figure 3.12: (a) Several nanoplasmonic waveguides clad in a 465nm thick PECVD oxide. (b) Close-up of a single nanoplasmonic waveguide clad in a 465nm thick PECVD oxide. (c) Close-up of a single nanoplasmonic waveguide clad in 465nm thick PECVD oxide and a 150nm thick Cr etch mask.

in Figs. 3.13(a),(b). The critical nanoplasmonic devices were located directly beneath the etch mask, and every possible precaution was taken to avoid damaging them or their surrounding environment.

Therefore, the process was systematically modified to improve the results and several parameters were investigated. First, the RF power was decreased from 3000W to 2200W to reduce damages from any plasma species that grazed the sidewall. In addition, the SF₆ gas flow was reduced from 350sccm to 300sccm, and the passivation cycle duration was increased from 1.3s to 1.8s. Both of these measures were designed to achieve a process that was less chemically aggressive and to add additional protection to the sidewalls. Through multiple iterations of development, virtually all of the damage near the etching mask was eliminated and the sidewall scallops were well-defined, as shown in Figs. 3.13(c),(d). By developing a gentler process, the total Bosch cycle duration became longer and etched a smaller distance per cycle, resulting in a greater overall etch time. However, this eliminated any damage to the nanoplasmonic waveguides and their environment.

As described earlier, a second layer of EBL was performed to define an etching mask for the characterization beam etching. Although the minimum feature size of this lithography mask was approximately 10 μ m, making it suitable for photolithography, the mask had to be aligned to the existing structures with an accuracy on the order of \sim 100nm. Such precise alignment was required to ensure that the cleave line was accurately aligned with one end of the waveguide end-facets and the large area adjacent to the other side of the waveguide was aligned close to the other waveguide end-facet. A slight misalignment would completely destroy the waveguides during the deep etching step. The EBL system is equipped with a laser interferometer stage positioning system, which allows for alignment accuracy on the order of \sim 30nm. By using a large beam aperture, and using the fixed-beam moving stage (FBMS) feature, each die could be processed in approximately 45 minutes [17].

A SEM of the Cr etch mask adjacent to a nanoplasmonic waveguide is shown in Fig. 3.14(a). The mask is designed so that nanoplasmonic waveguide is separated from the etching face by approximately 1 μ m, so that even if under-cut occurred during the etching, the waveguides would be protected. The final structure obtained after etching, cleaving, and removal of the Cr etch mask is shown in Figs. 3.14(b)-(d). By careful design of the exposure patterns, combinations of nanoplasmonic and nanophotonic waveguides with varying lengths in the range of $1.8\mu\text{m} \leq L \leq 65\mu\text{m}$ were obtained on the same sample, allowing for investigations into structures with different lengths, and side-by-side comparison of photonic and nanoplasmonic waveguides.

Cleaving the sample to intersect the waveguide end-facets also created an opportunity to observe the quality of the waveguide cross-sectional geometry. A cross-section of a nanoplasmonic waveguide with $w = 95\text{nm}$ is shown in Fig. 3.15(a). The etching profile is virtually vertical, and there is no observable sidewall roughness. The Au cap is adhered well to the Si core, and careful examination of the Au cap reveals that a thin remnant

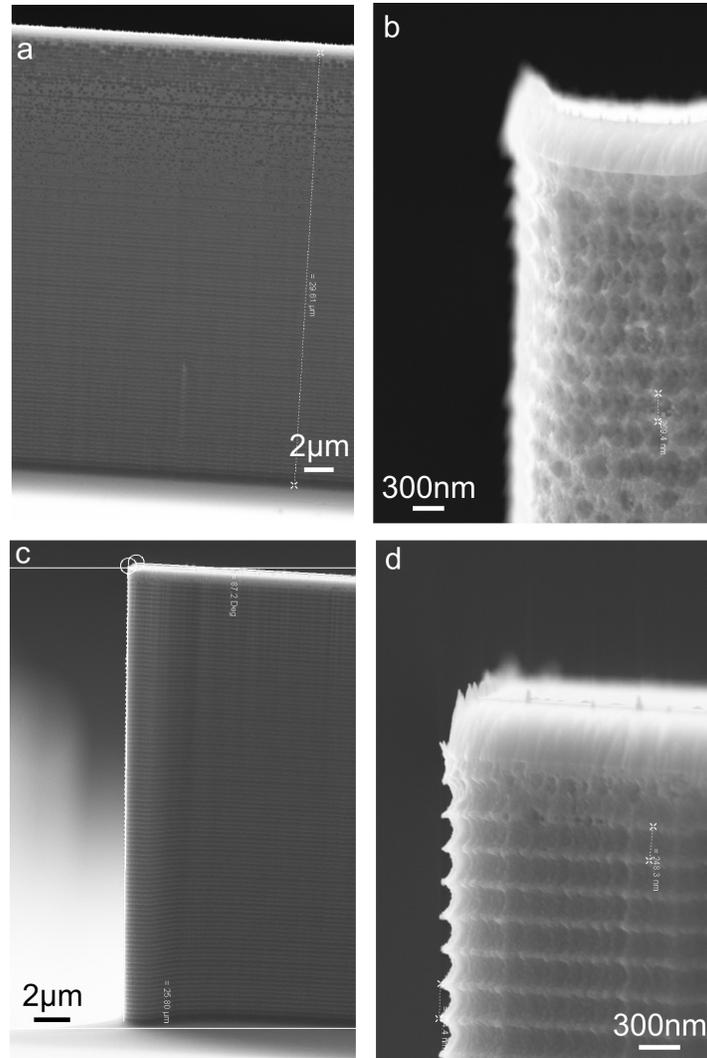


Figure 3.13: Optimization of a Bosch process used for deep silicon etching. (a) Sidewalls obtained using a standard process revealed significant roughness near the etch mask. (b) Close-up of the roughness near the etch mask, where the etching scallops were barely distinguishable. (c) Development of a less aggressive etching process minimized the roughness near the etch mask. (d) Close-up of the sidewall near the etch mask, revealing well-defined scallops.

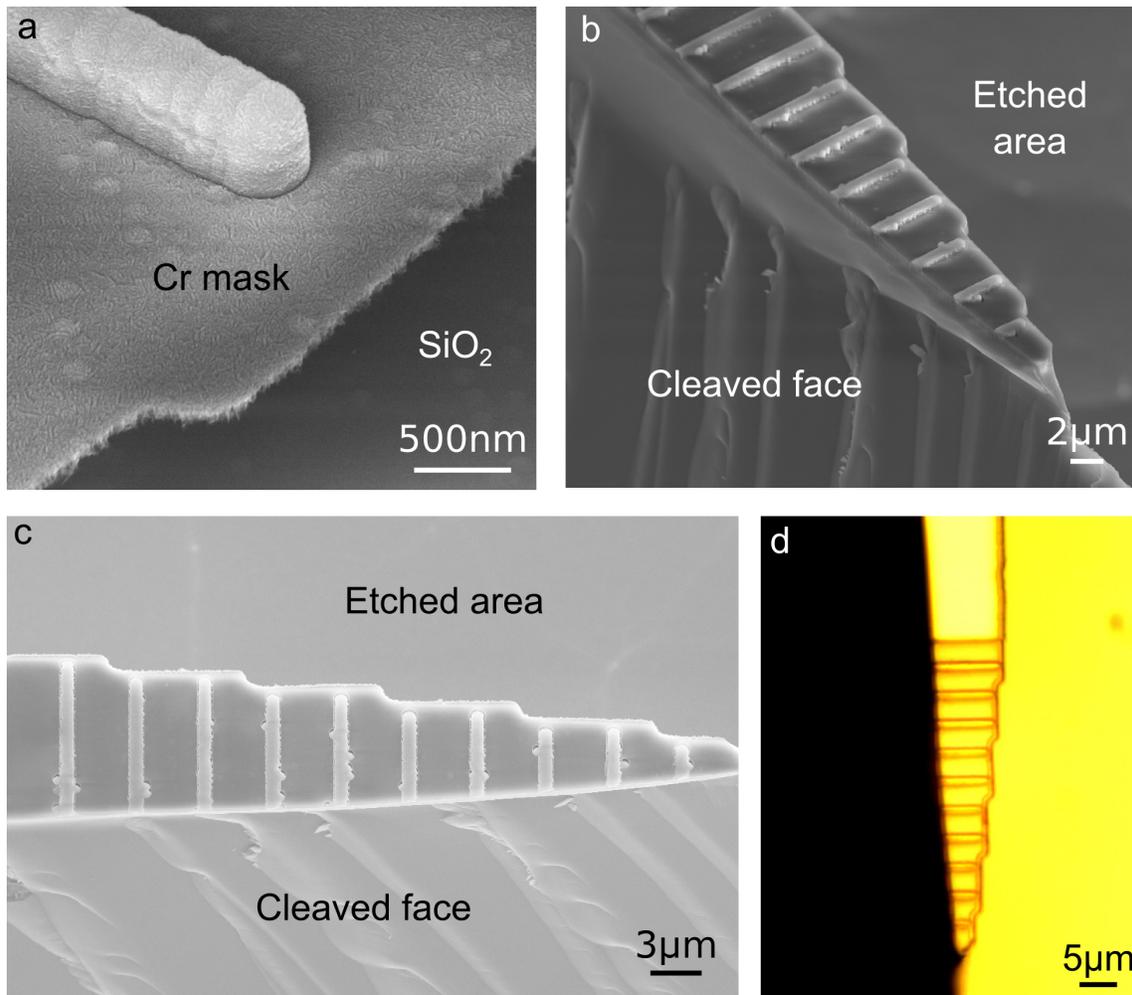


Figure 3.14: (a) Scanning electron micrograph of the Cr etch mask defined adjacent to a buried nanoplasmonic waveguide, prior to etching through the BOX layer or into the Si substrate. (b) Scanning electron micrograph of the final structure, imaged after deep etching, removal of the Cr etch mask, and cleaving. (c) Top-view SEM of the final structure. (d) Top-view optical microscope image of the final structure.

of the SiO_2 etching mask remains on its top surface. Similarly, the end-facet of a $w = 340\text{nm}$ nanophotonic waveguide is shown in Fig. 3.15(b), and end-facets of side-by-side nanophotonic and nanoplasmonic waveguides, each with $w = 340\text{nm}$ and the same length, is shown in Fig. 3.15(c), allowing for simple side-by-side comparison of the performance of the two types of waveguides.

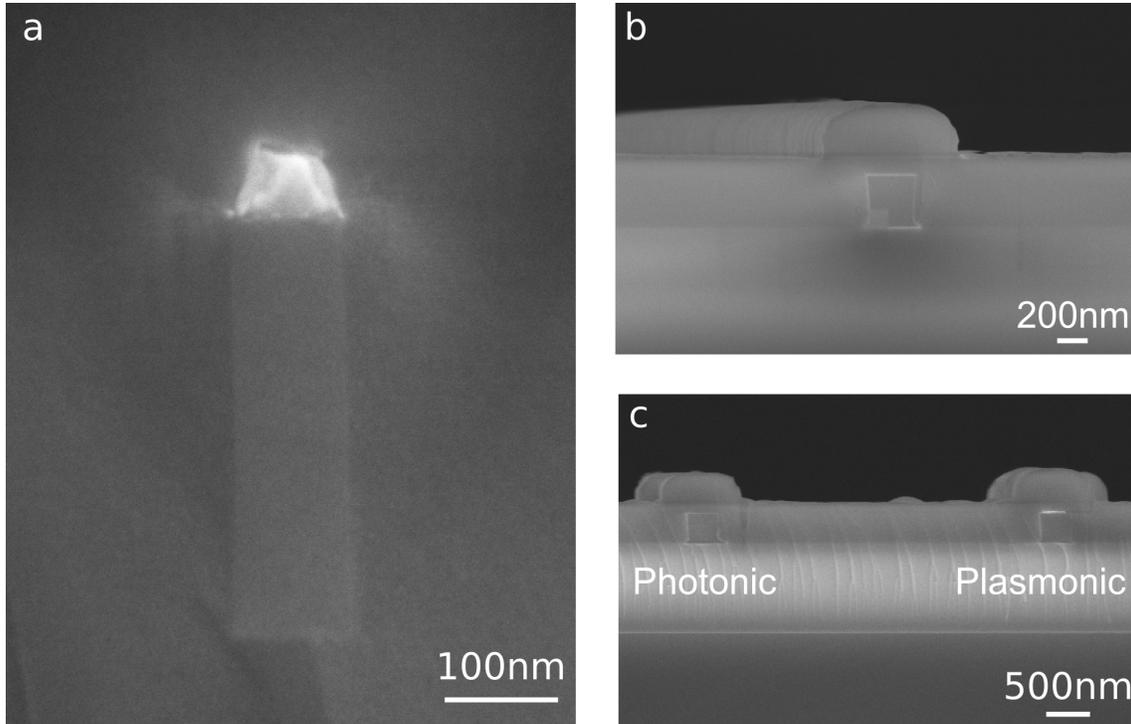


Figure 3.15: Cross-sectional scanning electron micrographs of end facets of waveguides integrated onto the characterization beam. (a) Nanoplasmonic waveguide with $w \approx 100\text{nm}$, (b) $w = 340\text{nm}$ photonic waveguide, and (c) $w = 340\text{nm}$ photonic and plasmonic waveguides integrated onto the same characterization beam.

3.4 Fabrication of Silicon-on-Insulator Nanophotonic Waveguides

The fabrication of SOI nanophotonic waveguides was significantly simpler than the previous process, and it consisted of only a single layer of lithography. The fabrication process was as follows:

1. Removal of organics via piranha cleaning.
2. Removal of native oxide layer via buffered oxide etching.
3. Spin 950 PMMA A2 resist.
4. Exposure in Raith 150-TWO EBL system.

5. Development in MIBK/IPA 1:3 solution for 25 seconds, rinse in IPA for 20 seconds.
6. Dry etch through Si device layer and strip resist.
7. Cleave to define waveguide end-facets.

Due to the conceptual simplicity of this process, schematics of the process flow are omitted, and a top-view SEM of a $w = 2.0\mu\text{m}$ silicon photonic waveguide is shown in Fig. 3.16.

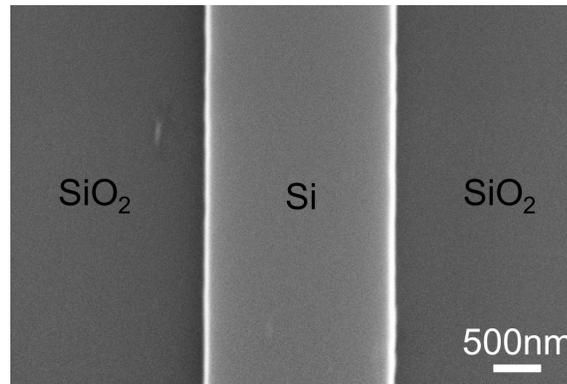


Figure 3.16: Top-view SEM of a $w = 2.0\mu\text{m}$ silicon photonic waveguide.

3.5 Integration of Silicon-based Nanoplasmonic Waveguides with Silicon Photonic Waveguides

To investigate both passive routing in silicon-based nanoplasmonic waveguides and interfacing with silicon photonic waveguides, a variety of lengths of silicon-based nanoplasmonic waveguides were fabricated. Both the input and output of each nanoplasmonic waveguide was interfaced to a silicon photonic waveguide. In this way, optical signals could be coupled to the input silicon photonic waveguide, which in turn excited the nanoplasmonic waveguide, and the radiation transmitted through the nanoplasmonic waveguide would then be coupled into the output silicon photonic waveguide.

Therefore, it was necessary to fabricate the two types of waveguide on the same die. The presence of two different types of waveguide with different materials and geometries made it necessary to perform two layers of EBL. As before, the small waveguide dimensions and the critical overlay alignment made EBL the most suitable lithography technique available in the Nanofab. The gold caps for the nanoplasmonic waveguides were defined in the first layer of EBL, as shown in Figs. 3.17(a),(b), and the silicon photonic input and output waveguides were defined in the second layer of EBL, as shown in Figs. 3.17(c),(d). The fabrication process flow is summarized as follows:

1. Removal of organics via piranha cleaning.

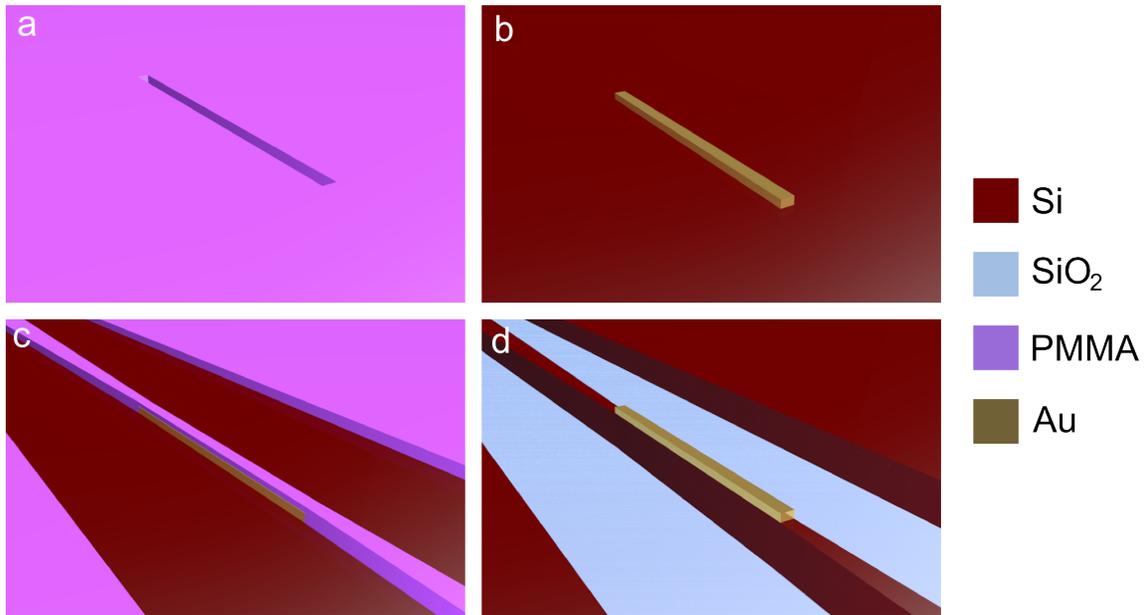


Figure 3.17: Fabrication steps for interfacing Si photonic waveguides with nanoplasmonic waveguides. (a) Exposure of nanoplasmonic features for lift-off metallization. (b) Nanoplasmonic waveguide cap after lift-off. (c) Exposure of Si photonic waveguide etch mask. (d) Final structures after silicon device layer etching and resist removal.

2. Removal of native oxide layer via buffered oxide etching.
3. Spin bilayer 495 PMMA A2 / 950 PMMA A2 resist.
4. Exposure in Raith 150-TWO EBL system.
5. Development in MIBK/IPA 1:3 solution for 45 seconds, rinse in IPA for 20 seconds.
6. Sputter 5nm Cr adhesion layer and 50nm gold.
7. Soak sample in acetone for one hour. Use ultrasonic bath to remove unwanted resist and gold, leaving behind the plasmonic waveguides.
8. Spin 950 PMMA A2 resist.
9. Exposure in Raith 150-TWO EBL system.
10. Development in MIBK/IPA 1:3 solution for 25 seconds, rinse in IPA for 20 seconds.
11. Etch through Si device layer, defining the silicon waveguides.

Figure 3.18(a)-(c) depict silicon-based nanoplasmonic waveguides with lengths of $L = 1\mu\text{m}$, $L = 2\mu\text{m}$, and $L = 3\mu\text{m}$, respectively, that have been interfaced to tapered silicon photonic waveguides. Figure 3.18(d) depicts a nanoplasmonic structure with multiple input and output ports, demonstrating the flexibility of this fabrication process. Several other interfaced structures are imaged with an optical microscope and shown in Fig. 3.19.

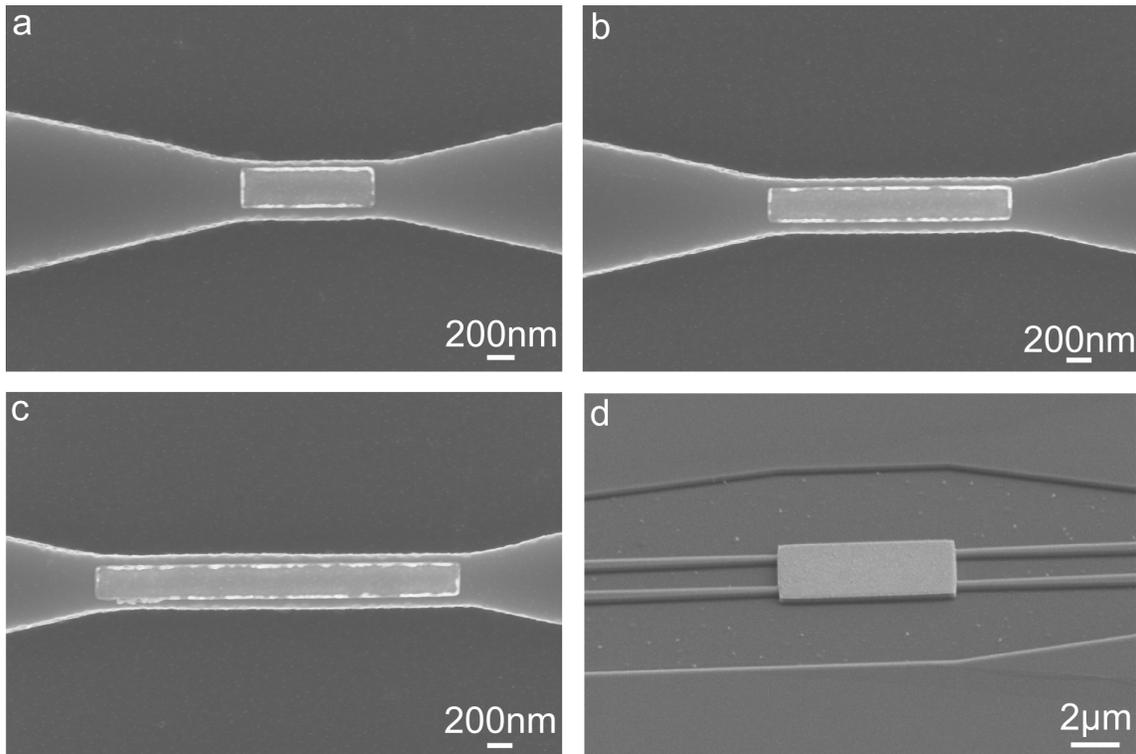


Figure 3.18: Top-view SEM of $w = 300\text{nm}$ nanoplasmonic waveguides interfaced to tapered silicon photonic waveguides. (a) $L = 1\mu\text{m}$, (b) $L = 2\mu\text{m}$, (c) $L = 3\mu\text{m}$, and (d) a multiple input, multiple output nanoplasmonic device.

3.6 Fabrication of Trapezoid Fabry-Pérot Resonator Waveguides

Section 3.3 described a technique to integrate nanoplasmonic and nanophotonic waveguides onto a characterization beam that would allow for direction collinear probing with microscope objectives and optical fibres. A simpler means to obtain very short waveguides is to fabricate them on the corner of a die. In this scheme, the two waveguide end-facets are oriented at 90° relative to one another, and the structure has a trapezoidal shape. Section 4.5 discusses the operation of these structures and demonstrates that they behave as Fabry-Pérot resonators. Therefore, they will be referred to as trapezoid Fabry-Pérot resonator (TFPR) waveguides.

The fabrication of TFPR waveguides was performed with the same process flow as in Section 3.4. To situate the waveguides on the corner of a die with a high process yield, a $400\mu\text{m} \times 400\mu\text{m}$ array of Si photonic waveguides was fabricated at a 45° angle relative to the cleavage planes of the SOI wafer, as shown in Fig. 3.20(a). Due to the large dimensions of the waveguide array, a cleave could easily be aligned to the waveguide array with the naked eye. The sample was cleaved in half, as shown in Fig. 3.20(b), and then into quarters, as shown in Fig. 3.20(c). Provided that the cleaves all intersected the waveguide array, this would yield two samples with usable TFPRs. Moreover, by careful

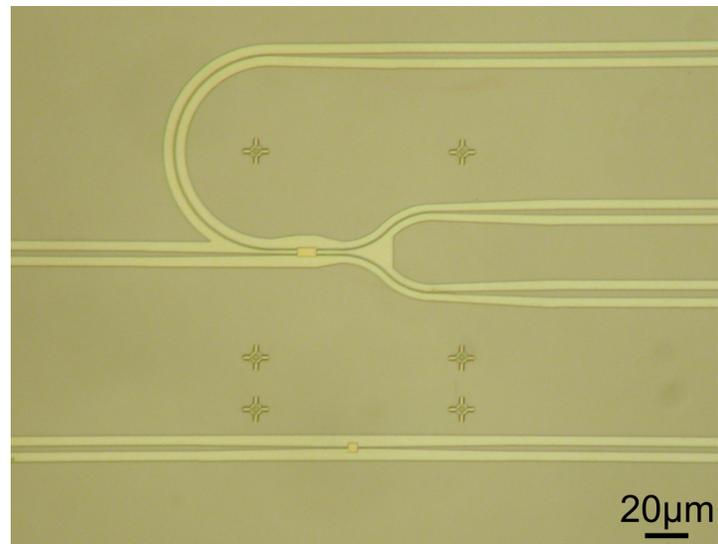


Figure 3.19: Optical microscope image of silicon photonic waveguides interfaced to nanoplasmonic structures.

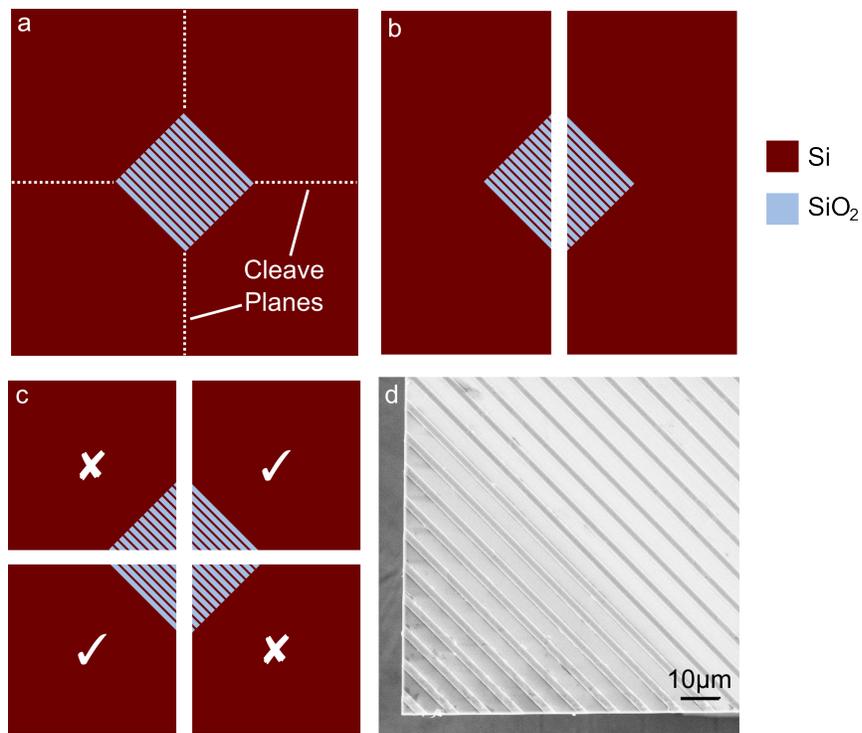


Figure 3.20: Fabrication steps for obtaining TFPRs on the corner of a SOI wafer. (a) A $400\mu\text{m} \times 400\mu\text{m}$ array of Si photonic waveguides is defined at a 45° angle to the cleave planes of the sample. (b) The sample was cleaved in half, where the cleave was aligned to intersect the waveguide array. (c) Each half was cleaved again to intersect the waveguide arrays, yielding two samples with TFPRs (denoted by ✓) and two unusable samples (denoted by ✗). (d) Top-view SEM of TFPR waveguides fabricated on a SOI substrate.

design of the waveguide array geometry, it was possible to obtain TFPRs with lengths less than $L \leq 10\mu\text{m}$ from every sample. A SEM of an array of TFPR waveguides is shown in Fig. 3.20(d).

3.7 Summary

Nanofabrication processes for realizing a variety of nanoplasmonic and nanophotonic waveguide devices have been developed. These processes provide a means to fabricate a wide range of photonic waveguides, to fabricate ultracompact photonic resonators on the corner of a die, and to interface nanoplasmonic and nanophotonic device platforms. The most involved fabrication process involved integrating nanoplasmonic structures onto a large characterization beam, requiring the combination of sub-100nm processing with $\sim 100\mu\text{m}$ processing. It is envisioned that this fabrication approach will be a valuable tool for accurate characterization of nonlinear optical devices, regardless of what their design or geometry can be.

The development of these processes required countless iterations of process optimization, including every aspect of lift-off metallization, silicon device layer etching, device protection, deep etching, and precision cleaving. However, the resulting devices enabled several very unique sets of experiments to be performed, which are outlined in the subsequent chapters. Virtually all of the processing steps used were compatible with complementary metal-oxide-semiconductor processing techniques, demonstrating the potential for monolithic integration of electronic, silicon photonic, and silicon-based nanoplasmonic device platforms.

References

- [1] R. Soref. The past, present, and future of silicon photonics. *IEEE J. Sel. Topics*, 12:1678–1687, 2006.
- [2] M. Bruel, B. Aspar, and A. J Auberton-Hervé. Smart-cut: A new silicon on insulator material technology based on hydrogen implantation and wafer bonding. *Jpn. J. Appl. Phys.*, 36:1636–1641, 1997.
- [3] G. K. Celler and S. Cristoloveanu. Frontiers in silicon-on-insulator. *J. Appl. Phys.*, 93:4955, 2002.
- [4] S. A. Maier. *Plasmonics: Fundamentals and Applications*. Springer, New York, 2007.
- [5] M. I. Stockman. Nanoplasmonics: past, present, and glimpse into future. *Opt. Express*, 19:22029–22106, 2011.
- [6] Z. Han and S. I. Bozhevolnyi. Radiation guiding with surface plasmon polaritons. *Reports on Progress in Physics*, 76:016402, 2013.
- [7] C. Chiu, E. Lisicka-Skrzek, R. N. Tait, and P. Berini. Fabrication of surface plasmon waveguides and devices in cytop with integrated microfluidic channels. *J. Vac. Sci. Technol. B*, 28:729, 2010.
- [8] C. Chiu, E. Lisicka-Skrzek, R. N. Tait, and P. Berini. Surface plasmon waveguide devices with tg-bonded cyop claddings. *J. Vac. Sci. Technol. B*, 29:062601, 2011.
- [9] E. Kretschmann. Decay of non radiative surface plasmons into light on rough silver films. comparison of experimental and theoretical results. *Opt. Commun.*, 6:185–187, 1972.
- [10] S. Sederberg, V. Van, and A. Y. Elezzabi. Monolithic integration of plasmonic waveguides into a complimentary metal-oxide-semiconductor- and photonic-compatible platform. *Appl. Phys. Lett.*, 96:121101, 2010.
- [11] R. H. Ritchie, E. T. Arakawa, J. J. Cowan, and R. N. Hamm. Surface-plasmon resonance effect in grating diffraction. *Phys. Rev. Lett.*, 21:1530–1533, 1968.

- [12] J. S. Huang, T. Feichtner, P. Biagioni, and B. Hecht. Impedance matching and emission properties of nanoantennas in an optical nanocircuit. *Nano Letters*, 9:1897–1902, 2009.
- [13] I. P. Radko, T. Holmgaard, Z. Han, K. Pedersen, and S. I. Bozhevolnyi. Efficient channel-plasmon excitation by nano-mirrors. *Appl. Phys. Lett.*, 99:213109, 2011.
- [14] J. Kiihamaki and S. Franssila. Deep silicon etching in inductively coupled plasma reactor for mems. *Physica Scripta*, 1999:250, 1999.
- [15] P. Maraghechi, K. Cadien, and A. Y. Elezzabi. A novel nanofabrication damascene lift-off technique. *Nanotechnology, IEEE Transactions on*, 10:822–826, 2011.
- [16] M. D. Henry. Icp etching of silicon for micro and nanoscale devices. *Ph. D. Thesis*, pages 51–52, 2010.
- [17] J. E. Sanabia, K. E. Burcham, J. Klingfus, G. Piaszenski, M. Kahl, and R. Jede. Fixed beam moving stage electron beam lithography of waveguide coupling device structures. In *Conference on Lasers and Electro-Optics 2012*, page CM4L.3. Optical Society of America, 2012.

Chapter 4

Ultrafast Nonlinear Effects in Silicon-on-Insulator Waveguides

4.1 Introduction

The most highly developed and intensively researched platform for integrated optical circuitry is currently silicon photonic waveguides fabricated on silicon-on-insulator (SOI) substrates. Silicon has several attractive features that have facilitated the rapid development of this field of research. Perhaps the most significant of these has been the well-established growth and processing techniques of silicon. State-of-the-art complementary metal-oxide-semiconductor (CMOS) processes have easily been adapted to SOI waveguides, enabling low-cost manufacture and device feature sizes on the order of $\sim 10\text{nm}$. The transparency of silicon at telecommunication wavelengths ($1260\text{nm} \leq \lambda \leq 1675\text{nm}$) enables the development of low-loss waveguides with propagation losses on the order of $\sim 0.1\text{dB/cm}$, permitting across-chip signal transmission in complex optical circuits and high quality factor filters. The high index contrast between Si ($n = 3.48$) and SiO_2 ($n = 1.45$) at telecommunication wavelengths facilitates high confinement waveguide modes and waveguide core dimensions on the order of a few hundred nanometres are standard in silicon photonics. The high index contrast also allows for circuitry with low losses around sharp bends, compact ring resonators, and slow-light photonic crystals. As a result of these features, silicon photonics is undoubtedly the most practical and accessible platform for integrated photonics, and has undergone rapid development over the last 20 years.

These developments have given silicon photonics a promising future in two main fields in industry: (1) hybrid electronic optical integrated circuitry; and (2) telecommunications. In the case of electronic circuitry, the main constraint on information processing bandwidth arises from the resistive-capacitive (R-C) time-delays in copper interconnects between electronic components. The compatibility of silicon photonic waveguides with CMOS fabrication processes has been a key motivator for replacing copper electrical inter-

connects with silicon photonic interconnects, alleviating the R-C time-delays and reducing power consumption. This has fuelled intensive efforts to develop the necessary infrastructure for hybrid electronic optical circuitry, including the integration of III-V semiconductor sources, Si-Ge photo detectors, and micro-optical components for alignment-insensitive coupling of radiation from fibre optical components to the integrated optical circuitry.

In the case of the telecommunications industry, the optical properties of silicon and its well-established processing techniques make it ideal for developing dense all-optical circuitry. All-optical chips would enable sophisticated, stand-alone information processing modules that could be used in fibre-optical networks, eliminating the need to convert energy from optical to electrical and back. In such a system, it is crucial to be able to modulate or switch signals, and therefore, nonlinear functionality is a key aspect of all-optical circuitry.

The linear loss of a photonic system places a limitation on the maximum length scale a signal can propagate, and therefore, the maximum optical circuit length that can be designed before the signal-to-noise ratio (SNR) deteriorates. A natural way to extend this interaction length is to increase the signal amplitude, extending the length that the signal can propagate before the SNR becomes unacceptable. However, increasing the input power can also lead to nonlinear interactions, which can distort the signal. When the signal power is increased into the nonlinear regime in silicon, nonlinear loss (i.e. TPA and FCA) can saturate the signal, placing an upper limit on the maximum signal amplitude that can be transmitted through the circuit, and a maximum total length for the circuit. Furthermore, nonlinear frequency mixing can occur, generating new frequency components and distorting the signal. Therefore, in applications where only linear routing is desirable, nonlinear effects such as TPA and FCA are detrimental, and it would be desirable to operate at wavelengths beyond the TPA band edge of silicon (i.e. $\lambda \geq 2.2\mu\text{m}$). However, nonlinear loss can also be used to amplitude-modulate a signal, making it a useful for scheme for all-optical signal modulation.

As a nonlinear medium, silicon has both advantages and disadvantages. The nonlinear refractive index of silicon ($n_2 = 3.2 \times 10^{-14}\text{cm}^2/\text{W}$ [1]) is more than two orders of magnitude greater than in silica optical fibres ($n_2 = 2.36 \times 10^{-16}\text{cm}^2/\text{W}$ [2]). Therefore, nonlinear interactions can be observed at much lower intensities and with much shorter interaction lengths than in optical fibres, making it practical for chip-scale nonlinear circuitry. Alternatively, the nonlinear susceptibility can be written as: $\text{Re}\{\chi^{(3)}\} = 8n_0n_2/2$, and the values for silicon and silica are calculated to be $\chi^{(3)} = 2.97 \times 10^{-13}\text{cm}^2/\text{V}$ and $\chi^{(3)} = 9.13 \times 10^{-16}\text{cm}^2/\text{V}$, respectively. For comparison, the value for a 50nm thick gold film has been measured to be $\chi^{(3)} \approx 2 \times 10^{-15}\text{cm}^2/\text{V}$ [3]. A number of nonlinear optical effects have been studied in SOI waveguides, including TPA [4–8], FCA, free-carrier plasma dispersion (FCD), and optical Kerr-based effects, including four-wave mixing [9–14], cross-phase modulation (XPM) [15–17], and self-phase modulation (SPM), [4, 18–23]. Significant progress has also been made in Raman amplification in SOI waveguides [24–

28]. Devices based on these effects have been used for frequency mixing, modulation, and all-silicon Raman lasers. Investigating the regimes in which these effects become apparent reveals both the limitations they have on passive light propagation in SOI waveguides, and enable nonlinear functionality. Along with its high nonlinear coefficient, silicon has a high thermo-optic coefficient ($\Delta n/\Delta T = 1.86 \times 10^{-4}/\text{K}$, where Δn is the change in refractive index due to a change in temperature, ΔT), which has been used to demonstrate thermal switching devices [29–36].

Several challenges related to the nonlinear properties of silicon should also be noted. The presence of TPA and FCA limit the potential of optical Kerr-based effects at high excitation intensities and over long interaction lengths, as the pump radiation can be rapidly depleted by nonlinear loss, preventing optical Kerr-based effects from building up. The centrosymmetry of crystalline silicon makes the second-order nonlinear interaction negligible and makes the third-order nonlinear interaction dominant. Despite having a much higher $\chi^{(3)}$ coefficient than SiO_2 , the $\chi^{(3)}$ nonlinear interactions in silicon are still typically weak compared to those observed in exotic $\chi^{(2)}$ nonlinear materials, such as LiNbO_3 and LiTaO_3 . Therefore, it can be challenging to increase the efficiency of the nonlinear processes in integrated SOI circuitry to those attainable with exotic bulk samples. Furthermore, the indirect bandgap of silicon prevents efficient light emission from silicon, and therefore, the development of on-chip, silicon-based light emitters. Finally, the indirect silicon bandgap also limits the timescale of free-carrier relaxation in silicon photonic waveguides to the order of $100\text{ps} \leq \tau_c \leq 10\text{ns}$. This limits the maximum modulation bandwidth of devices that require free-carrier injection and sophisticated processing techniques such as ion implantation or integration with an electronic diode must be pursued to decrease the recovery time.

Despite the multitude of investigations into nonlinear effects in SOI waveguides, there is very little fundamental understanding of the interplay between the different nonlinear interactions. Several of the nonlinear effects, including TPA, FCA, and SPM do not need to be phase-matched, and as a result, there is no way to suppress unwanted nonlinear effects when a single specific effect is desired. Therefore, when operating in the nonlinear regime, all of these effects occur simultaneously, and their interaction can be highly dependent on the pulse spectrum and the device geometry. Furthermore, the nonlinear loss arising from TPA and FCA can generate significant heat in a device, depending on its geometry and the thermal dissipation of its surroundings. Hence, there can also be an interplay between the nonlinear optical and thermo-optical properties of certain devices.

4.2 Passive Routing

4.2.1 The Cut-Back Method

Before nonlinear measurements were pursued, it was important to understand the linear waveguide loss, so that a comparison of the nonlinear interaction length (which can be limited by the linear propagation losses of the waveguide) and the waveguide length could be made. A schematic depiction of a SOI waveguide is shown in Fig. 4.1, which has a width, w , height, h , and length, L . The cut-back method was used to measure the coupling efficiency to the waveguides and the propagation loss. In this technique, several waveguides with the same width and height, but different lengths were characterized at low input powers, where no nonlinear effects were significant. As a result of fabrication-induced roughness, the signal power will decay exponentially as it propagates along the length of the waveguide. Therefore, by coupling the same power into each waveguide, a plot of output power versus waveguide length can be generated, revealing the length scale of the exponential decay, L_{prop} . By fabricating all of the waveguides on the same chip, they will be processed and aligned in precisely the same manner, and fabrication-induced inaccuracies in the coupling efficiencies of each waveguide will be minimized. The y -intercept of this plot represents a waveguide with an infinitesimal length, where there is no propagation loss and the only losses that occur between the free-space beam and the detector arise from coupling radiation into and out of the waveguide. In reality, incident radiation would essentially scatter from an infinitesimally short waveguide (i.e. $L \ll \lambda$) and a proper waveguide mode would not be excited. Therefore, this serves primarily as an analytical technique that can be used to extract performance merits from the waveguides, and any physical interpretation is not valid for $L \ll \lambda$.

4.2.2 Measurements

In many of the measurements taken in this chapter, a microscope objective was used to couple laser pulses into SOI waveguides, and a lensed optical fibre was used to out-couple the transmitted radiation and deliver it to the detector. Typically, the coupling efficiencies of these two schemes (i.e. the microscope objective and the lensed optical fibre) are different. Therefore, the cut-back measurements were first performed with two identical lensed optical fibres¹, from which the coupling efficiency between the waveguide and the lensed optical fibre was obtained. To perform these measurements, two identical lensed optical fibres were used to couple continuous-wave $\lambda = 1550\text{nm}$ radiation into and out of the waveguides with several different lengths. An exemplary plot of output power versus waveguide length for a waveguide with $w \times h = 340\text{nm} \times 340\text{nm}$ and lengths of $L = \{200, 615, 825, 1120\}\mu\text{m}$, is shown in Fig. 4.2. In this case, a propagation length

¹OZ Optics, Model #: TSMJ-3A-1550-9/125-0.25-7-2.5-14-2, Corning SMF-28 fibre, spot diameter: $2.5\mu\text{m}$, working distance: $14\mu\text{m}$.

of $L_{prop} = 790\mu\text{m}$ was measured, corresponding to a propagation loss of $\alpha = 55\text{dB/cm}$. These losses are relatively large compared to the state-of-the-art silicon photonic devices ($\alpha = 3.6\text{dB/cm}$ or $L_{prop} = 1.2\text{cm}$ [37]), signifying significant sidewall roughness. However, because the length of these structures is on the order of the propagation length, they were still suitable for generating strong nonlinear interactions and taking measurements with high accuracy. The well-defined exponential decay trend demonstrates that inaccuracies related to nanofabrication inconsistencies or optical alignment have been minimized.

Subsequently, the same measurements were performed using a microscope objective and a lensed optical fibre. With the coupling efficiency between the waveguide and the lensed optical fibre already known, the coupling efficiency between the microscope objective and the waveguide could be extracted. In general, the coupling efficiencies depend strongly on the cross-sectional dimensions of the waveguide. In the case of a $NA = 0.65$ microscope objective, coupling efficiencies of 1 – 5% could be achieved. A $NA = 0.85$ microscope objective provided a tighter focus and enabled coupling efficiencies of 2 – 7%. For a lensed optical fibre with a focal spot diameter, $d = 2.5\mu\text{m}$, coupling efficiencies of 10 – 25% were measured.

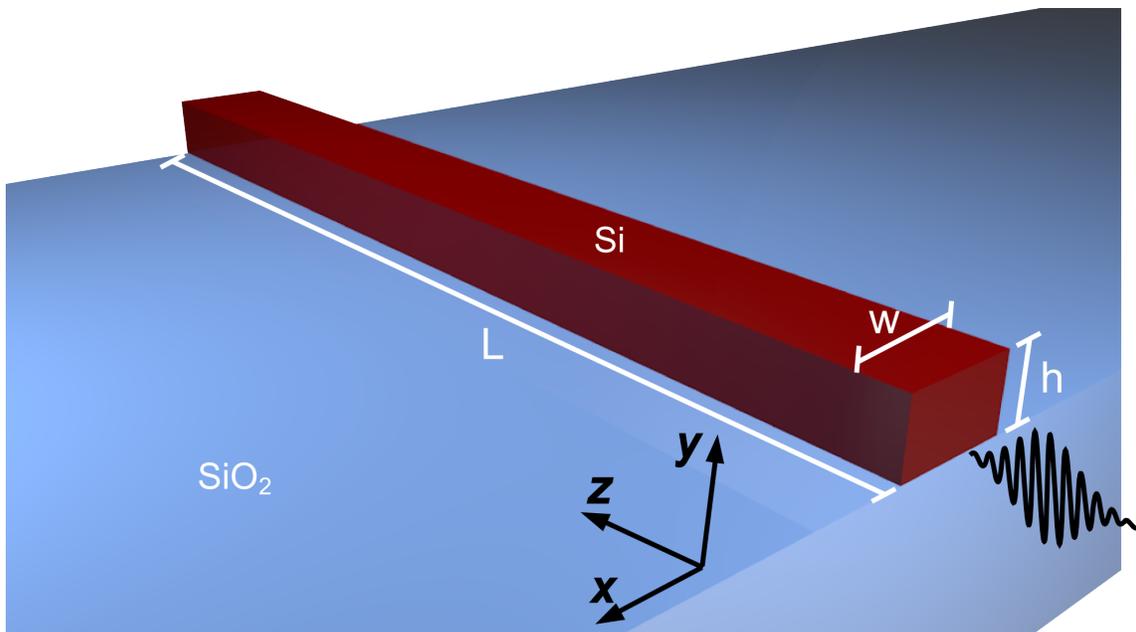


Figure 4.1: Schematic depiction of a SOI waveguide.

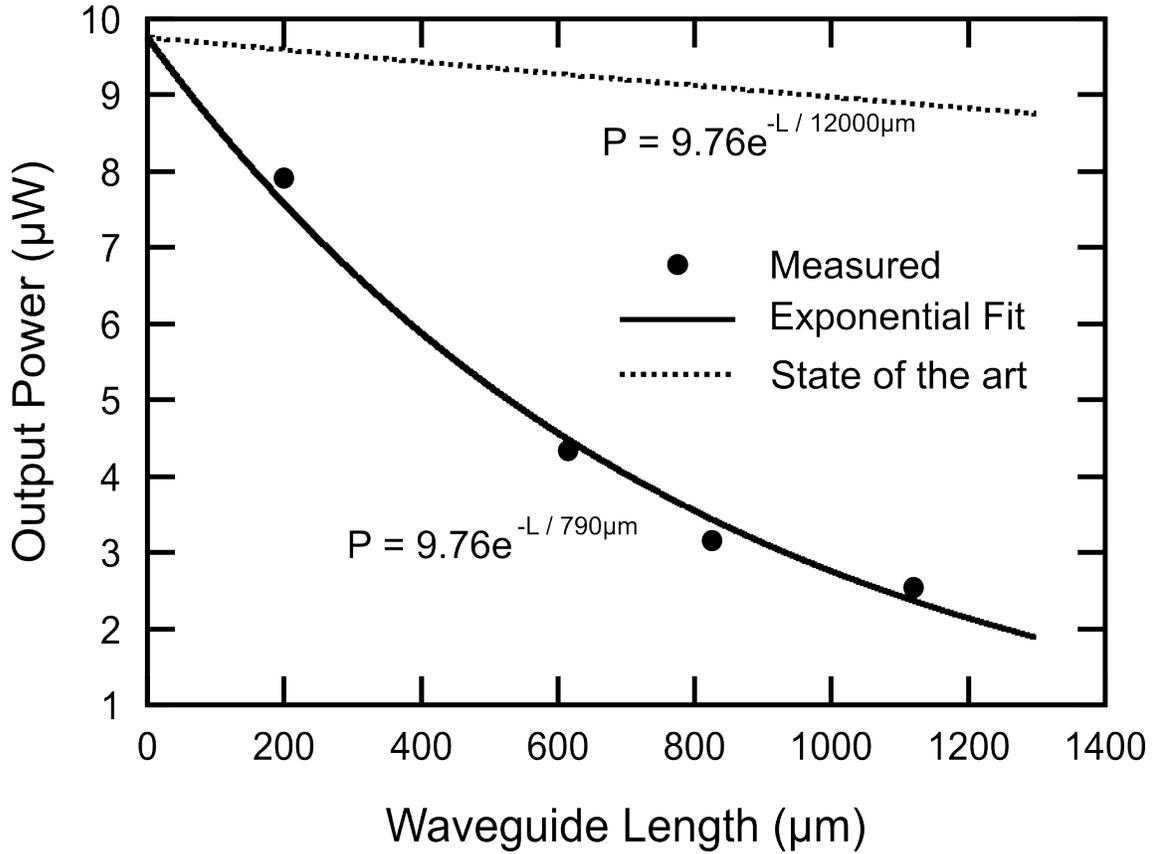


Figure 4.2: Plot of power transmitted through waveguides with fixed cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and lengths of $200\mu\text{m}$, $615\mu\text{m}$, $825\mu\text{m}$, and $1120\mu\text{m}$. An exponential decay function is fit to these data points and reveals a propagation length of $790\mu\text{m}$, corresponding to a linear propagation loss of $\alpha = 55\text{dB/cm}$. The y -intercept represents the power transmitted through a waveguide of infinitesimal length and provides the input and output coupling loss product.

4.3 Two-photon Absorption and Free-carrier Absorption

4.3.1 Power Transmission Saturation

Two-photon absorption and FCA are investigated by coupling femtosecond pulses with a duration, $\tau_p = 84\text{fs}$, centred at $\lambda = 1550\text{nm}$ at a repetition rate of $F_{rep} = 90\text{MHz}$ into SOI waveguides with various dimensions. The pulses were end-fire coupled to the SOI waveguides from free-space with a microscope objective ($NA = 0.65$), and transmitted radiation was coupled into a lensed single mode optical fibre with a focal spot diameter, $d = 2.5\mu\text{m}$. By orienting the polarization of the electric field along the y -axis (as shown in Fig. 4.1), the quasi-TM waveguide mode is the primary mode that is excited. Due to the high confinement of the SOI waveguide, the quasi-TE mode will have an E_y field component and will also be excited, but with a much lower efficiency. The quasi-TE will not be sufficiently intense to excite significant nonlinear interactions, and therefore, the

observed nonlinear interactions will arise from the quasi-TM mode.

The free-space power incident on the microscope objective was adjusted using a continuously-variable neutral density filter wheel and was set to the desired amplitude with a power meter. A power meter with a fibre connector was used to measure the transmitted power, eliminating background light contributions. By measuring the transmitted power for varying input power, a transition from a linear transmission regime to a nonlinear saturation regime was observed, where the saturation was produced by TPA and FCA. A typical power transmission curve for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 1120\mu\text{m}$ is shown in Fig. 4.3. For peak input powers less than $P_{peak}^\omega \leq 18\text{W}$, the output power scales approximately linearly with P_{peak}^ω . A linear fit was performed on the data points in this regime, and was extrapolated to the maximum peak power that was considered, $P_{peak}^\omega = 165\text{W}$, representing the output power that would be observed in the absence of nonlinear loss. As P_{peak}^ω was increased above 20W, the transmission made a distinct deviation from the linear trend, and began to saturate. For $P_{peak}^\omega \geq 75\text{W}$, the nonlinear loss became substantial enough that further increases in input power yielded very little increases in the output power. The total nonlinear loss due to TPA and FCA was calculated by subtracting the measured output power from the linear fit, which is shown by the triangles. For $P_{peak}^\omega = 95\text{W}$ (corresponding to the point where the nonlinear loss curve intersects the output power curve in Fig. 4.3), approximately half of the input power was lost to TPA and FCA. As the peak power was increased further, more than half of the input power was absorbed.

More insight can be gained by studying power transmission saturation in waveguides with the same cross-sectional dimensions, $w \times h = 340\text{nm} \times 340\text{nm}$, and varying lengths. Power saturation curves for 4 different waveguides with $L = \{200, 615, 825, 1120\}\mu\text{m}$ are shown in Fig. 4.4. As expected, the total power transmission for each of the waveguides decreases as the waveguide length increases, regardless of the input power. In the linear regime, this is due to the linear waveguide losses. For higher powers in the nonlinear regime, this is due to a combination of linear waveguide losses and nonlinear losses that accumulate over a longer interaction length. The $L = 200\mu\text{m}$ waveguide is significantly shorter than $L_{prop} = 780\mu\text{m}$, and therefore, it can be expected to have a significantly lower linear loss, yielding a higher signal transmission. At the same time, the nonlinear interaction length is limited to the waveguide length, and the nonlinear interactions are less pronounced. For this reason, the linear regime of this transmission curve is extended to approximately $P_{peak}^\omega \approx 40\text{W}$, whereas the other three waveguides begin to saturate at input powers in the range of $20\text{W} \leq P_{peak}^\omega \leq 30\text{W}$. In the case of the three longer waveguides with $L = \{615, 825, 1120\}\mu\text{m}$, the waveguide length is comparable to, or greater than L_{prop} . Therefore, it can be expected that the nonlinear interaction length will be longer and will be approximately equal to L_{prop} . Because these waveguides each have a length that is greater than or approximately equal to the propagation length of the waveguide, they exhibit similar saturation trends.

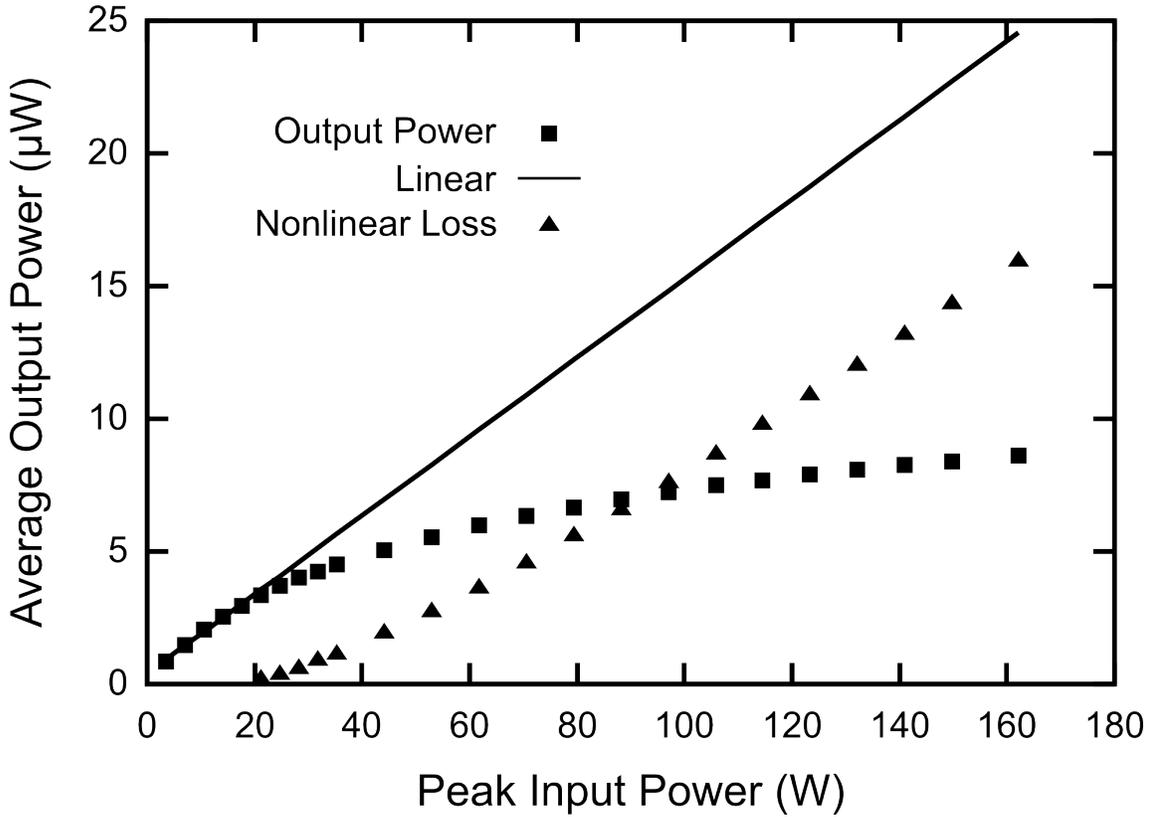


Figure 4.3: Plot demonstrating the saturation of power transmission through a waveguide with cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and a length of $L = 1120\mu\text{m}$. The squares represent the average output power transmitted through the waveguide as a function of the peak power coupled to the waveguide. The solid line is an extrapolation of the low-power linear region of the measured data points and represents the average output power that would be observed in the absence of nonlinear loss mechanisms in the silicon core of the waveguide. The nonlinear loss is calculated by subtracting the measured data points from the linear extrapolation, and is represented by the triangles.

It is also important to study the influence of the cross-sectional dimensions of the waveguide on the power saturation behaviour. In general, for a fixed input power, a reduction in the waveguide core dimensions will produce a greater optical intensity in the nonlinear silicon core, and stronger nonlinear interactions. Mode profile intensity distributions are calculated for waveguides with a fixed height, $h = 340\text{nm}$, and varying widths, $w = \{260, 340, 470, 540, 760, 1000\}$ nm for a fixed input peak power, $P_{peak}^\omega = 75\text{W}$, and are shown in Fig. 4.5. The maximum intensity in the silicon core of each mode profile is determined and is shown in Table 4.1. In general, as w is decreased, the maximum intensity in the Si core of the waveguide increases due to the decreasing effective area, A_{eff} , of the waveguide mode. For the $w = 260\text{nm}$ waveguide, 54.3% of the mode energy propagates outside of the silicon core, whereas only 43.8% propagates outside of the silicon core for the $w = 340\text{nm}$ waveguide. As a result, the maximum intensity in the silicon core of the $w = 260\text{nm}$ waveguide is approximately 4% less than for the $w = 340\text{nm}$ waveguide.

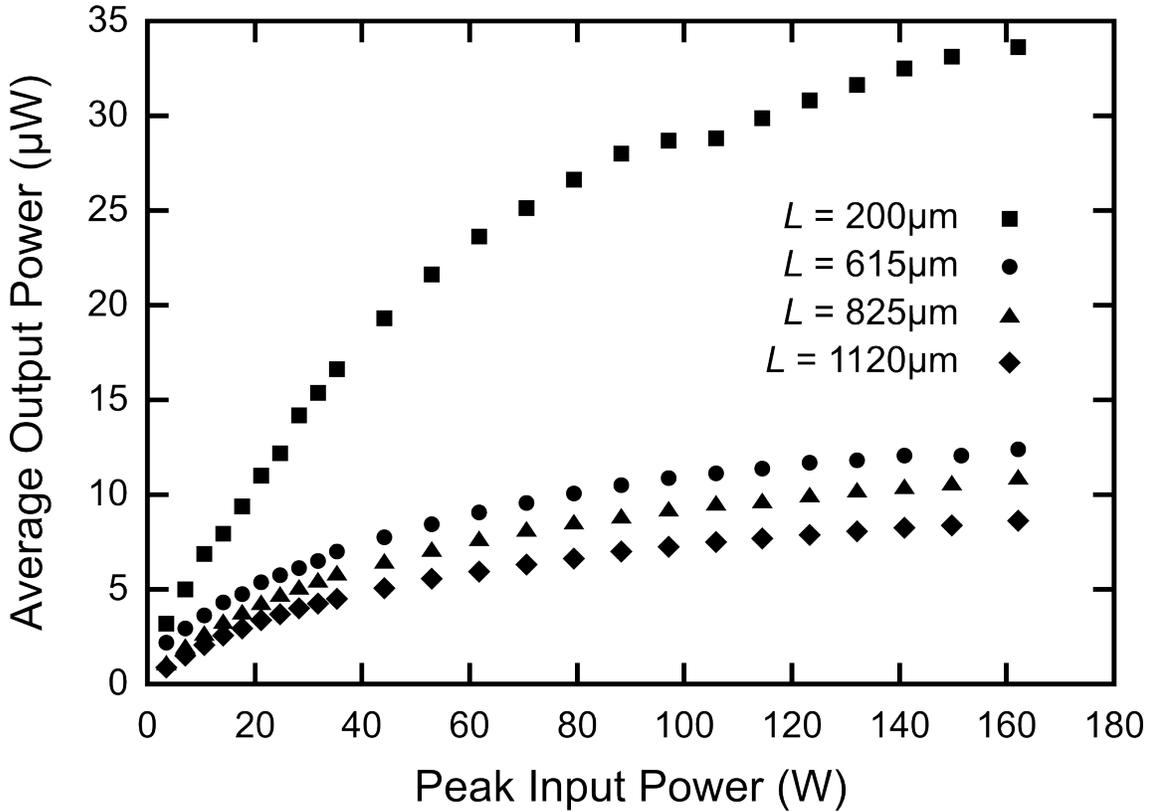


Figure 4.4: Power transmission saturation plots for four waveguides with fixed cross-sectional dimensions of $w \times h = 340\text{nm} \times 340\text{nm}$ and varying lengths of $L = 200\mu\text{m}$, $615\mu\text{m}$, $825\mu\text{m}$, $1120\mu\text{m}$.

Based on these calculations, it can be expected that the nonlinear interactions present in the waveguides will become greater as w is decreased to $w \approx 300\text{nm}$, and that stronger nonlinear interactions will not be observed in waveguides with $w \leq 300\text{nm}$ due to poorer waveguide mode confinement. Power saturation curves for waveguides with a fixed length, $L = 1120\mu\text{m}$, a fixed height, $h = 340\text{nm}$, and widths of $w = \{260, 340, 470, 540, 760\}$ nm are shown in Fig. 4.6. For widths of $w = 260\text{nm}$ and $w = 340\text{nm}$, linear signal transmission is observed for $P_{peak}^\omega \leq 20\text{W}$. As the width is increased to $w = 760\text{nm}$, the linear regime is extended to $P_{peak}^\omega = 40\text{W}$, signifying that weaker nonlinear interaction is taking place as w is increased. In addition, it is observed that the saturation becomes stronger (i.e. an incremental increase in input power yields a smaller increase in output power) and the transmitted power in the saturation regime decreases as w is decreased. It should also be noted that the waveguides with $w = 260\text{nm}$ and $w = 340\text{nm}$ follow almost the same shape, which is consistent with their very similar maximum intensities, as shown in Table 4.1.

Based on these measurements, it is clear that a strong signal modulation can be obtained via TPA and FCA in these waveguides. In the interest of using these effects to modulate a signal, it is important to understand the time-scale of the excitation and relaxation processes, as this determines the maximum modulation bandwidth that can be

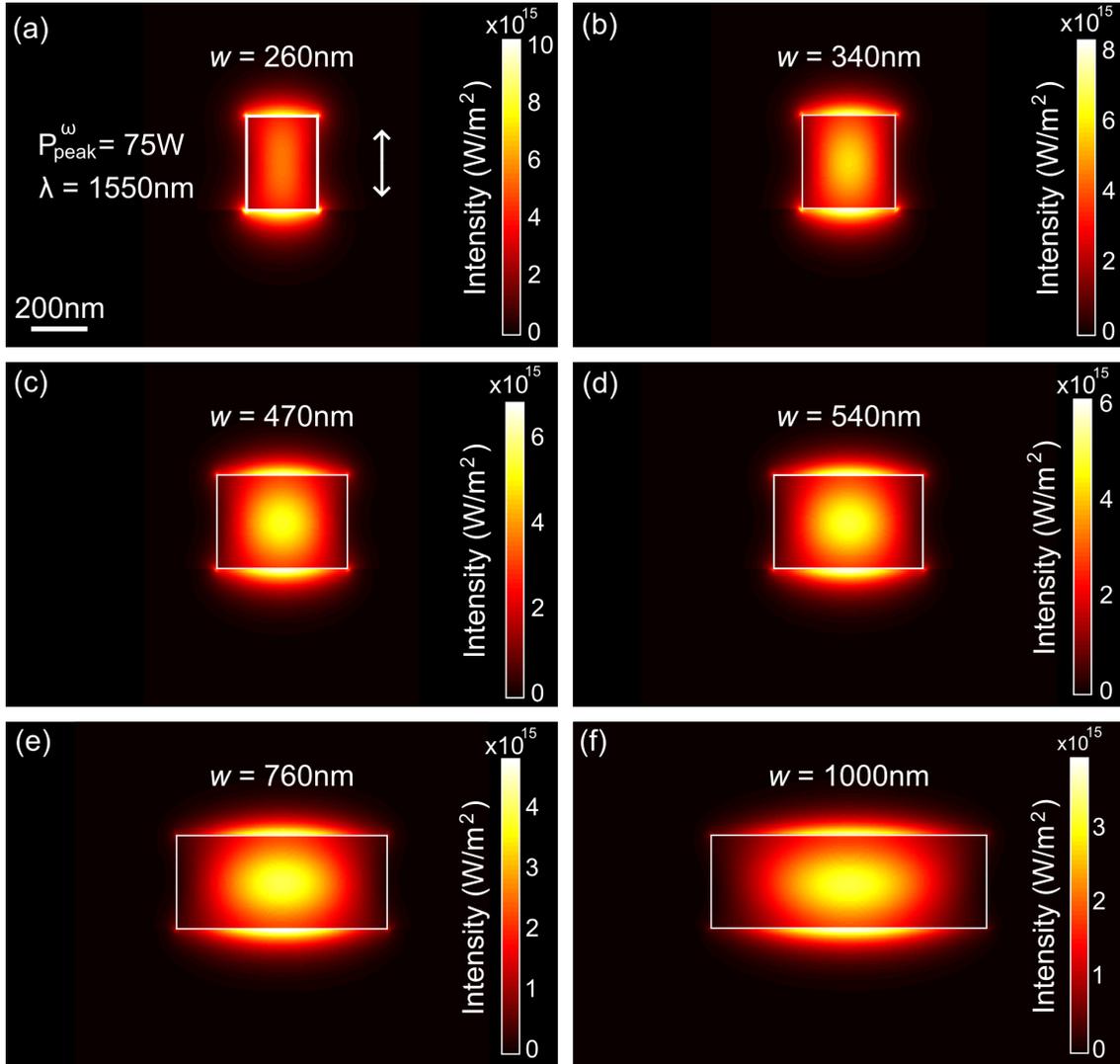


Figure 4.5: Quasi-transverse-magnetic mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$, for a fixed peak input power, $P_{\text{peak}}^{\omega} = 75\text{W}$, for waveguides with $h = 340\text{nm}$ and $w =$ (a) 260nm , (b) 340nm , (c) 470nm , (d) 540nm , (e) 760nm , and (f) 1000nm .

Width (nm)	$I_{max}(\times 10^{15}\text{W}/\text{m}^2)$
260	5.53
340	5.76
470	5.31
540	4.94
760	3.97
1000	3.21
1510	2.25
2050	1.70
4150	0.86

Table 4.1: Maximum intensity in Si core of waveguides with different widths at a peak input power, $P_{peak}^\omega = 75\text{W}$.

achieved.

4.3.2 Time Dynamics

Pump-Probe Measurements

To measure the timescale of TPA and FCA, cross-polarized pump-probe experiments are performed on the SOI waveguides (described in Appendix D). In these measurements, the presence of a strong pump pulse attenuates a weak probe pulse via TPA and FCA. The attenuation of one signal due to the presence of another is often referred to as cross-amplitude modulation (XAM). A microscope objective ($NA = 0.65$) is used to end-fire couple both the pump and probe into the SOI waveguide and the transmitted radiation is out-coupled from the waveguide using a $2.5\mu\text{m}$ focal spot lensed optical fibre. Performing pump-probe measurements on a waveguide with sub-micron cross-sectional dimensions is very challenging, as both beams must be focused with a precision of $\sim 10\text{nm}$, and this precision must be maintained as the retroreflector is repositioned. This becomes particularly challenging when time delays greater than $\Delta t \geq 100\text{ps}$ are desired. For example, a time delay of $\Delta t = 200\text{ps}$ requires the retroreflector to be moved approximately $\Delta x = 3\text{cm}$. The alignment must be perfect for such long scans to be performed with virtually no beam drift, which would change the coupling efficiency of the probe to the waveguide and would introduce inaccuracies to the pump-probe trace.

Pump-probe transmission experiments are performed on a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$ and a typical pump-probe time-domain trace for a peak pump power, $P_{peak}^\omega = 28\text{W}$, is shown in Fig. 4.7. It should be noted that this waveguide is short enough that several pump reflections from the waveguide end-facets are present in the trace. However, the reflected power is not high enough to excite its own nonlinear interaction, and the contribution of these reflections can be ignored. For time delays, Δt , where the probe arrives before the pump ($\Delta t \leq 0\text{ps}$), no nonlinear loss is mapped onto the probe and the probe power transmission is constant. As the time delay is decreased and the pulse peaks overlap, a sharp drop in the transmission due to TPA is observed,

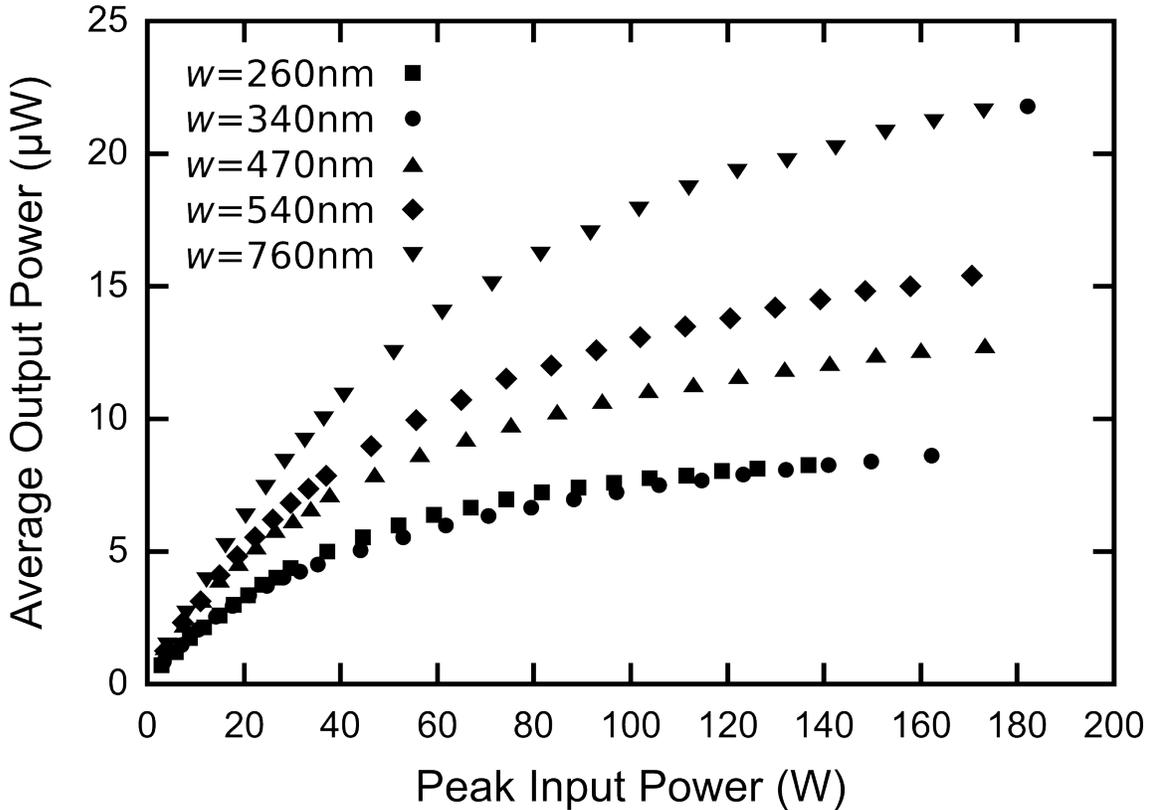


Figure 4.6: Power transmission saturation plots for five waveguides with a fixed length, $L = 1120\mu\text{m}$, a fixed height, $h = 340\text{nm}$, and varying widths of $w = 260\text{nm}$, 340nm , 470nm , 540nm , 760nm .

which follows the timescale of the overlap between the two pulses ($\sim 100\text{fs}$). As the time-delay is increased further so that the probe arrives after the pump, the free-carrier relaxation process is observed. This presents itself as a long recovery tail that extends beyond the temporal window that was scanned. The presence of free-carriers attenuates the probe via FCA, and the amplitude of the probe attenuation is proportional to the free-carrier density that it interacts with. Therefore, further loss due to FCA is observed. The amplitude of this loss decreases as the time delay is increased and the free-carrier population decreases. The free-carrier relaxation time is determined to be $\tau_r = 265\text{ps}$ by fitting an exponential curve to the recovery tail. Therefore, this limits the modulation bandwidth of this waveguide to approximately $1/\tau_r = 3.8\text{GHz}$.

Still using the same waveguide, the modulation depth was controlled by adjusting the pump power, which is shown in Fig. 4.8. Here, the black trace was taken at a low pump power, $P_{peak}^\omega = 2.8\text{W}$, which was below the threshold for nonlinear interaction. Therefore, no signature of nonlinear interaction was mapped onto the probe. As the pump peak power was increased to $P_{peak}^\omega = 11.0\text{W}$, $P_{peak}^\omega = 22.1\text{W}$, and $P_{peak}^\omega = 33.1\text{W}$, the depth of modulation due to TPA became greater, and therefore, so did the FCA-induced loss of the recovery tail. Investigating a shorter temporal window also revealed more details of the interaction. As the time delay was decreased so that the pump and probe pulses

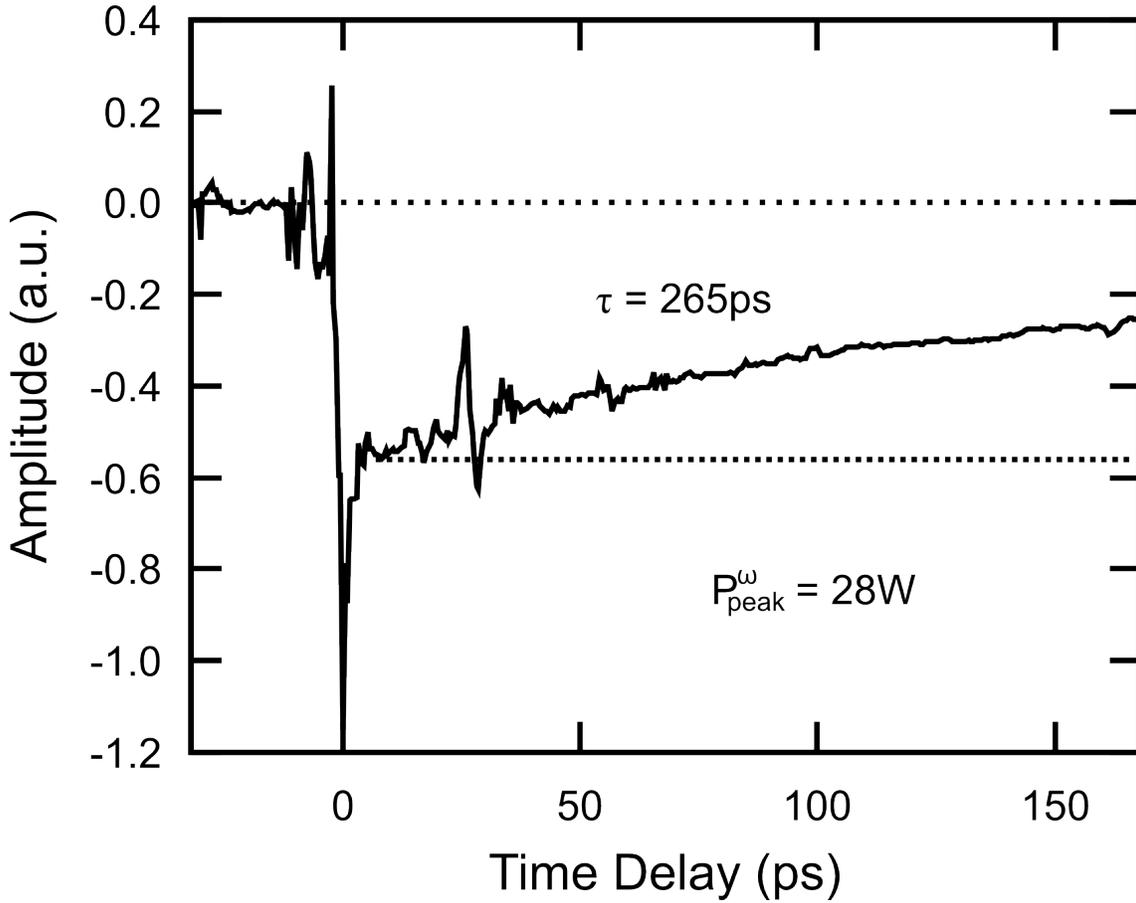


Figure 4.7: Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at $P_{peak}^{\omega} = 28\text{W}$. At a $t = 0\text{ps}$ time-delay, TPA produces a sharp drop in transmission. For a $t > 0\text{ps}$ time delay, FCA attenuates the probe beam by an amplitude that is proportional to the instantaneous free-carrier density. An exponential fit to the recovery tail yields a free-carrier recombination time of $\tau_r = 265\text{ps}$.

began to overlap, fringes appeared due to linear interference between the quasi-TE and quasi-TM modes (which do not have perfectly linear polarization, and therefore, some interference was present). Then, when the pulse peaks overlapped, there was a sharp drop due to TPA. The reflections of the pump from the waveguide end-facets also became more clearly distinguishable.

In principle, these measurements demonstrate the potential to use all-optical XAM to achieve amplitude modulation. Alternatively, this can be used as an all-optical switch, where the pump switches the probe from the “on” state to the “off” state on the timescale of the pulse duration, and the probe can then recover back to the “on” state on the timescale of τ_r .

Further analysis of these curves reveals interesting details related to the free-carrier recombination dynamics. Figure 4.9 shows the fitted free-carrier recovery time as a function of the input peak power. For $P_{peak}^{\omega} \leq 16.6\text{W}$, the recovery time is relatively constant

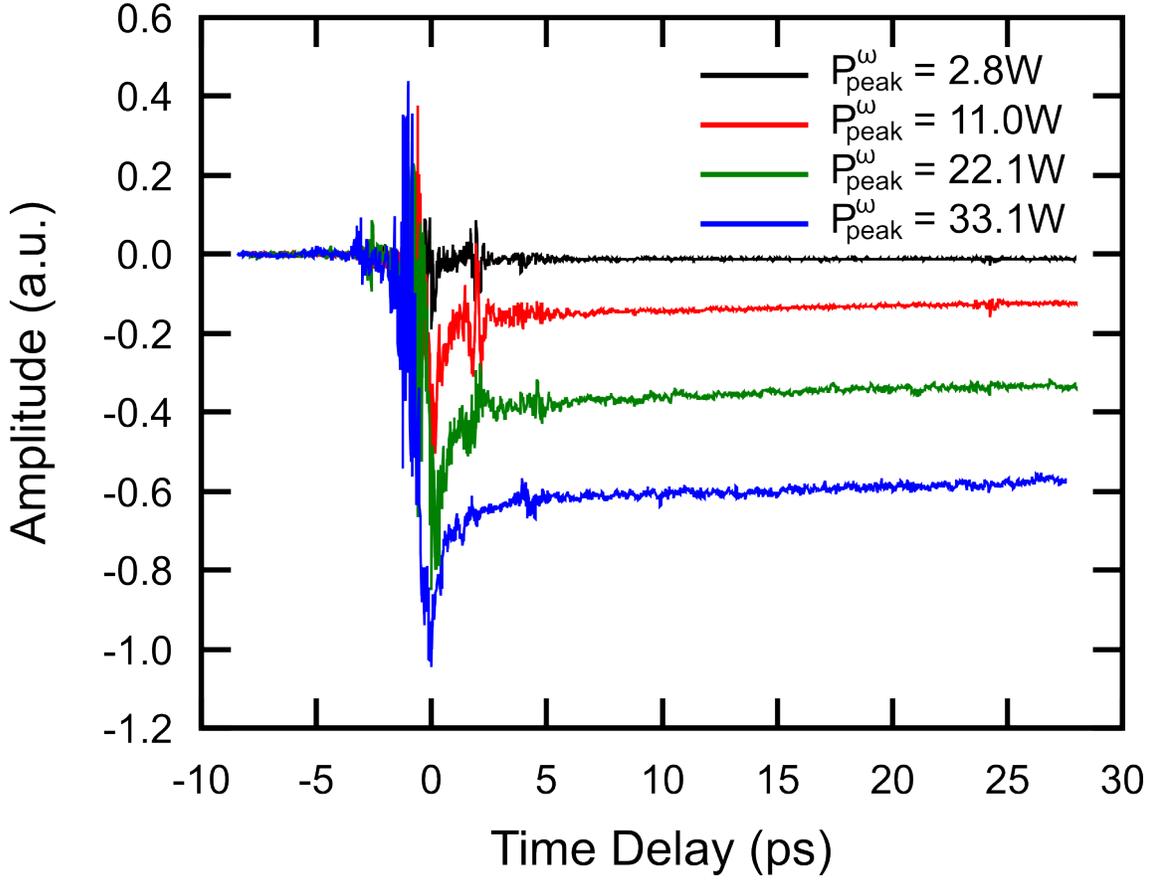


Figure 4.8: Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at peak pump powers of $P_{peak}^\omega = 2.8\text{W}$ (black), $P_{peak}^\omega = 11.0\text{W}$ (red), $P_{peak}^\omega = 22.1\text{W}$ (green), and $P_{peak}^\omega = 33.1\text{W}$ (blue). The nonlinear loss imparted to the probe due to TPA and FCA from the pump increases with the pump power, enabling control of the modulation depth.

at $\tau_r \approx 125\text{ps}$. However, a linear increase in the recombination time is observed as the pump power is increased above $P_{peak}^\omega = 16.6\text{W}$, and for $P_{peak}^\omega = 33.1\text{W}$, the recovery time increased by a factor of 2.5 to $\tau_r = 312\text{ps}$. There are two dominant electron scattering interactions that take place on an ultrafast timescale [38]: (1) electron-electron scattering, which redistributes the energy of the carriers; and (2) electron-phonon scattering, which transfers energy from the hot electrons to the lattice (i.e. carrier relaxation). The energy redistribution, or thermalization, occurs within less than 120fs [39]. Following this, the electrons relax to the conduction band minimum through emission of longitudinal optical phonons [40]. The electron-phonon relaxation rate depends strongly on the density of the excited carriers and for high-carrier densities, an optical phonon population may buildup, introducing a bottleneck to the relaxation process. The increase of the free-carrier relaxation time for $P_{peak}^\omega \geq 16.6\text{W}$ comes as a direct consequence of the optical phonon bottleneck.

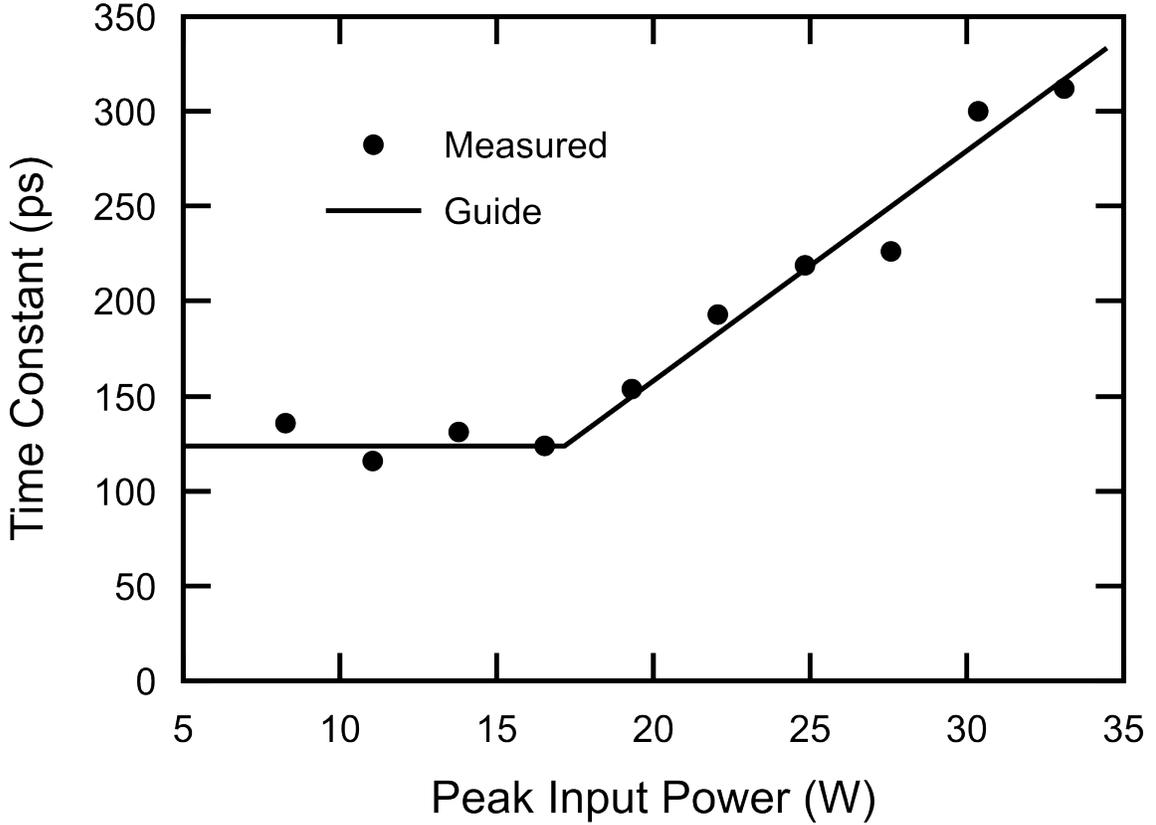


Figure 4.9: Free-carrier recovery times extracted from the recovery tail of pump-probe time domain traces for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, taken at different peak pump powers.

Modulation Over a Very Short Lengthscale

Modulators working on the principle of TPA and FCA typically take the form of either a very long photonic waveguide (typically millimetres in length) or a high quality-factor resonator. Both of these structures are characterized by a very long nonlinear interaction length, allowing the nonlinear loss to accumulate over a long length scale, and generating a strong nonlinear signature. In the interest of developing non-resonant modulators (i.e. modulators that can be used at a variety of wavelengths) that occupy a very small footprint area, it is necessary to determine what interaction length the nonlinear loss must accumulate over to generate a strong nonlinear signature that would be suitable for XAM. In the case of TPA, the probe intensity, I_{probe} is depleted by the pump, I_{pump} , according to:

$$\frac{dI_{probe}}{dz} \approx -\beta_{TPA} I_{pump}^2. \quad (4.1)$$

Similarly, for FCA, the probe intensity is depleted according to:

$$\frac{dI_{probe}}{dz} \approx -\sigma_{FCA} N I_{pump}. \quad (4.2)$$

Therefore, it can be expected that the modulation depth of the recovery tail will reduce significantly as the device length is reduced.

Figure 4.10 presents a typical pump-probe trace for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 25\mu\text{m}$ when excited by a peak pump power, $P_{peak}^\omega = 27.6\text{W}$. Comparing this trace to the traces in Fig. 4.8, it is clear that there is already a diminishing modulation depth. Decreasing the device length further to $L = 6.5\mu\text{m}$ and using an even higher peak pump power of P_{peak}^ω yields a modulation depth that is even less pronounced. Such weak modulation would not permit a high on-off contrast ratio and, therefore, would not be suitable for a modulation device.

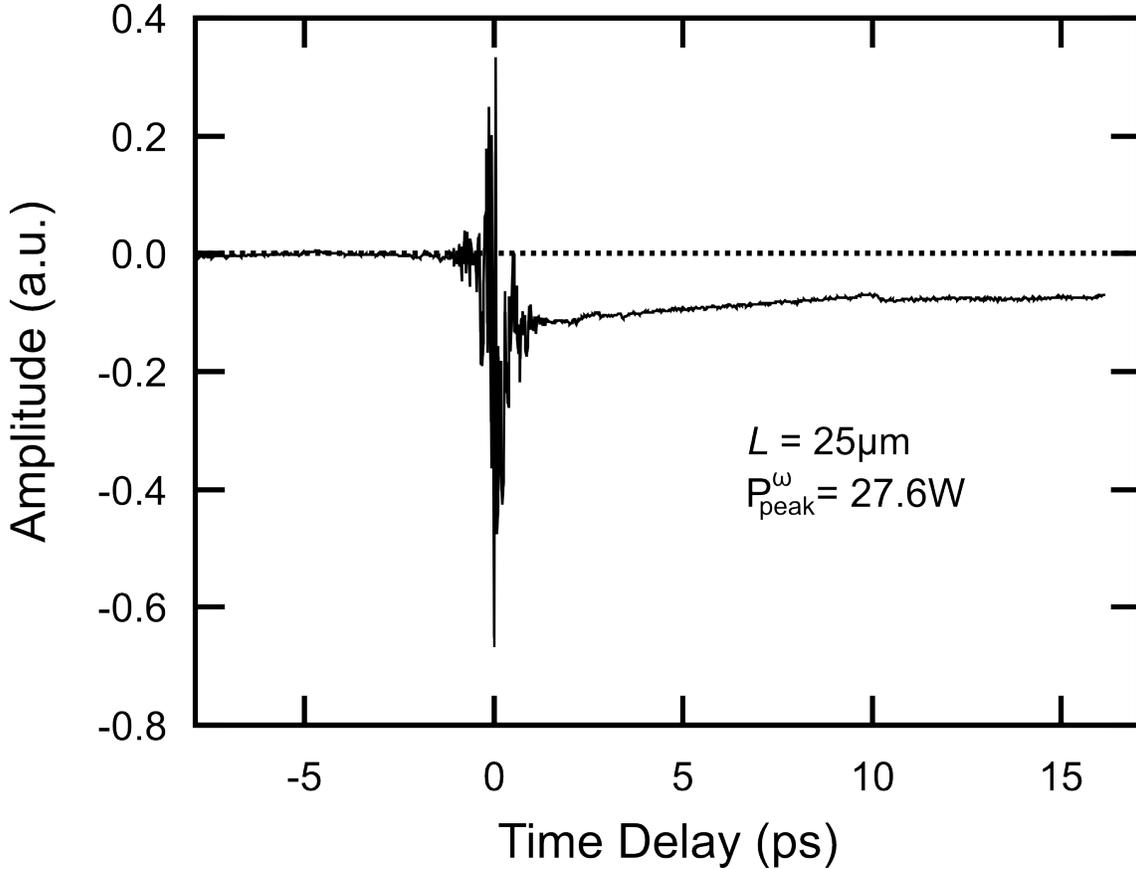


Figure 4.10: Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 25\mu\text{m}$.

4.4 Self-phase Modulation

Besides nonlinear loss, it is also interesting to consider the nonlinear refractive index, i.e. the optical Kerr effect. In the case of femtosecond pulses propagating through SOI waveguides, substantial pulse broadening can result from self-phase modulation (SPM). As described in Section 2.4.2, self-phase modulation develops at a high laser intensity where the optical Kerr effect (i.e. $\text{Re}\{\chi^{(3)}(-\omega; \omega, -\omega, \omega)\}$) produces an intensity-dependent re-

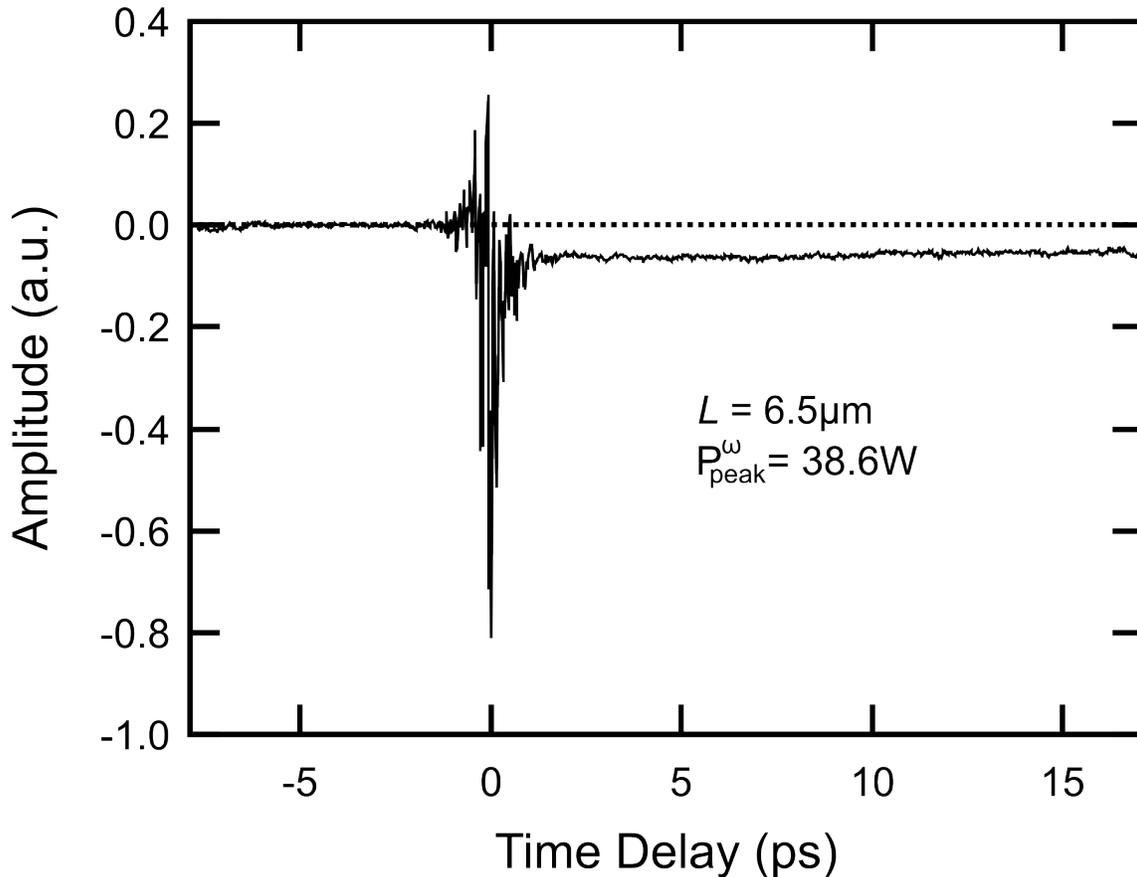


Figure 4.11: Cross-polarized pump-probe time domain trace for a waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 6.5\mu\text{m}$.

fractive index. As a femtosecond pulse propagates through a waveguide, it induces a time-varying effective mode index, which modifies the local phase of the electric field. The nonstationary time-dependent phase appears as an instantaneous frequency of the pulse. The temporal front of the pulse experiences an increasing refractive index, which corresponds to a red-shift, whereas the temporal back of the pulse experiences a decreasing refractive index, which corresponds to a blue-shift. In the absence of TPA-generated free-carriers in silicon, the pulse broadening would be symmetric. However, free-carriers generated at the pulse peak can interact with the temporal back of the pulse, yielding asymmetric temporal broadening. Free-carriers excited by the pulse peak produce a decreasing refractive index for the temporal back of the pulse, leading to additional blue-shifting. Furthermore, the free-carriers can attenuate the temporal back of the pulse via FCA.

Although useful for signal modulation, TPA and FCA are highly detrimental to SPM. In particular, they limit the maximum pulse broadening that can be achieved, regardless of the input peak power or the interaction length with the waveguide. The extent to which nonlinear loss inhibits the broadening is expressed as a nonlinear figure of merit

(FOM),

$$FOM = \frac{n_2}{\beta_{TPA}\lambda} \quad (4.3)$$

where n_2 is the second order nonlinear refractive index and β_{TPA} is the TPA coefficient. For silicon, the typical measured values at $\lambda = 1550\text{nm}$ are: $\beta_{TPA} = 0.68\text{cm/GW}$ and $n_2 = 3.2 \times 10^{-14}\text{cm}^2/\text{W}$, and hence, $FOM = 0.302$ [1].

Previous investigations into pulse broadening in SOI waveguides have focused on the input power-dependence on pulse broadening [16, 18–22], the dependence of the pulse broadening on the central wavelength of the input pulse [21], and the use of SPM in nonlinear all-optical functionality [20]. However, despite being very important parameters in Si nonlinear device operation, there have been no investigations into the waveguide geometry dependence, or equivalently, the mode-profile dependence of SPM in SOI waveguides. Furthermore, there has been very little knowledge of the physics behind the interplay between mode profile, SPM, TPA, FCA, and plasma dispersion, and how all these effects contribute to the power scaling of various frequency components of the pulse.

Using the same set of waveguides as Section 4.3, it is possible to investigate SPM-induced spectral broadening of $\lambda = 1550\text{nm}$ femtosecond pulses for SOI waveguides with cross-sectional dimensions spanning a factor of 16, permitting careful investigation into the influence of mode confinement on the pulse-broadening. The $\lambda = 1550\text{nm}$ femtosecond pulses were focused on the input facet of the waveguides using a $NA = 0.65$ microscope objective and the transmitted pulses are out-coupled with a lensed optical fibre and delivered to an optical spectrum analyzer (OSA), allowing for careful examination of SPM-induced pulse broadening. The polarization of the laser beam was oriented to excite the quasi-TM mode of the SOI waveguides.

Typical measured transmitted spectra from a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 610\mu\text{m}$ are normalized and shown on a logarithmic y -axis scale in Fig. 4.12. The spectrum obtained at $P_{peak}^\omega = 7\text{W}$ (dotted line) is taken below the threshold for nonlinear interaction and represents the linear transmission of the $\lambda = 1550\text{nm}$ laser pulses through the SOI waveguide. As depicted in the figure, insignificant power is detected at wavelengths below $\lambda \leq 1480\text{nm}$ or above $\lambda \geq 1625\text{nm}$. The spectrum taken at $P_{peak}^\omega = 152\text{W}$ (solid line) is much broader due to SPM, indicating a strong nonlinear interaction within the Si core. The power of the spectral components in the range, $1450\text{nm} \leq \lambda \leq 1480\text{nm}$, increases by a factor of ≈ 225 , whereas the spectral components in the range, $1625\text{nm} \leq \lambda \leq 1650\text{nm}$, are enhanced only by a factor of ≈ 15 . Interestingly, there are some spectral components whose power decreases as the peak power is increased. This suggests a more complex interaction than SPM alone, and is evidence of an interplay between SPM, TPA, FCA, and plasma dispersion.

To investigate the complex interplay between these effects, it is insightful to examine the evolution of various spectral components as the peak input power is increased. On a linear scale, a 2D surface plot depicting the spectral power components of the transmitted

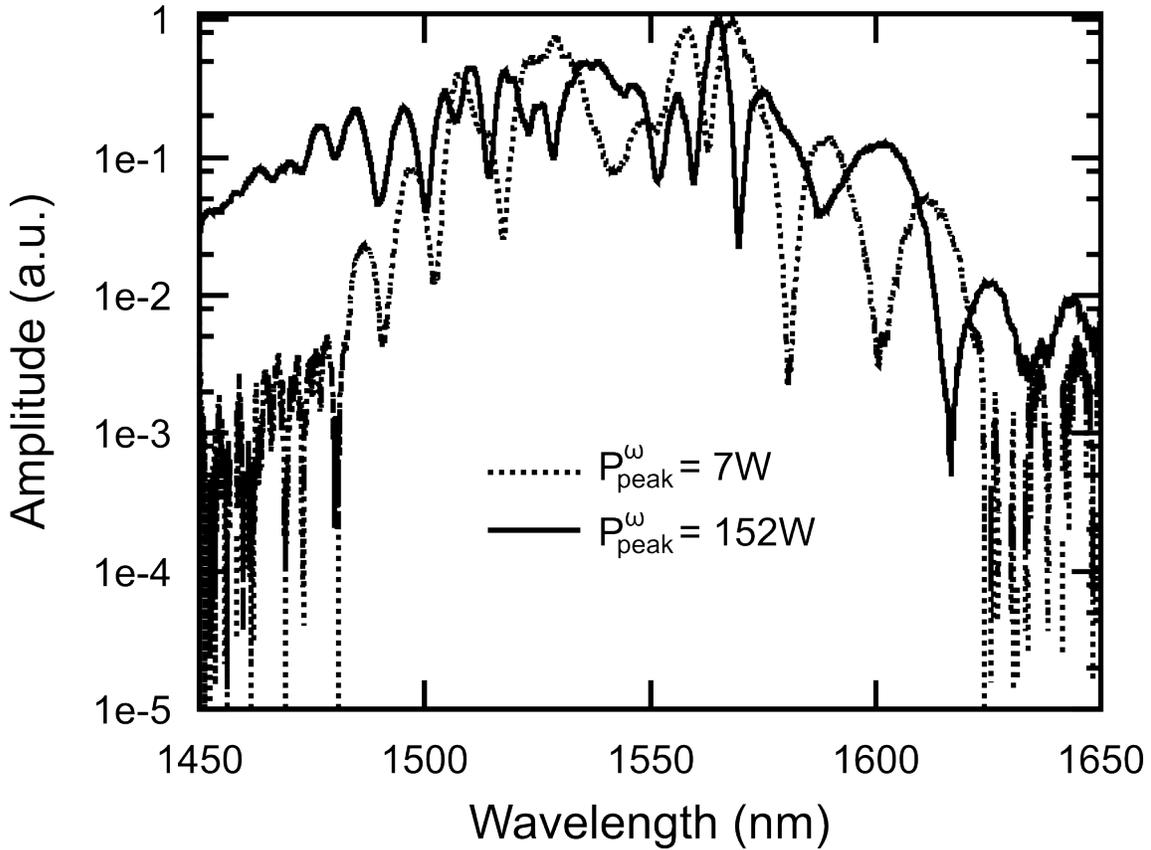


Figure 4.12: Transmitted $\lambda = 1550\text{nm}$ pulse spectrum through a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 610\mu\text{m}$, taken at peak input powers of $P_{peak}^\omega = 7\text{W}$ (dotted line) and $P_{peak}^\omega = 152\text{W}$ (solid line). Both spectra are normalized and shown on a logarithmic y -axis scale.

pulse as a function of the input peak power for a waveguide with $w = 340\text{nm}$ is shown in Fig. 4.13(a). Significant spectral reshaping and broadening is evident as the peak power is increased above $P_{peak}^\omega = 20\text{W}$. The most noticeable transformations are the appearance of spectral components in the range, $1535\text{nm} \leq \lambda \leq 1550\text{nm}$, and distinct multiple sidebands at $\lambda \leq 1480\text{nm}$ and $\lambda \geq 1580\text{nm}$. Figure 4.13(b) displays the entire integrated transmitted spectrum for each input peak power, P_{peak}^ω . Notably, for $P_{peak}^\omega \leq 20\text{W}$, the trend is approximately linear, but nonlinear loss (i.e. TPA and FCA) gradually saturates the transmitted power and for $P_{peak}^\omega \geq 110\text{W}$, subsequent increases in the input power yield virtually no increase in the transmitted power.

While the general power dependence depicted in Fig. 4.13(b) is well-known and was studied in Section 4.3.1, it is erroneous to generalize across the whole spectrum of the pulse as strong spectral reshaping takes place and the growth/saturation of specific wavelengths with P_{peak}^ω may not exhibit such behaviour. It is more insightful to investigate this scaling for various wavelength regions within the pulse. This is particularly important for ultrafast all-optical wavelength-dependent modulation and switching devices. To illustrate this peculiar property of the SOI waveguide, several interesting spectral slices depicting

different power scaling trends are shown in Figs. 4.13(c)-(f). Figure 4.13(c) represents the integrated power transmitted in the wavelength range, $1517\text{nm} \leq \lambda \leq 1519\text{nm}$, a spectral region where there is very little power emitted by the laser pulses. For input peak powers below $P_{peak}^\omega = 50\text{W}$, virtually no power is transmitted within this spectral region, as no SPM pulse broadening has occurred to generate new wavelengths in this region. As the peak power is increased above $P_{peak}^\omega = 50\text{W}$, a substantial free-carrier population ($N \sim 10^{19} - 10^{20}\text{cm}^{-3}$) is excited via TPA and gives rise to blue-shifting of the spectral components (see Fig. 4.13(a)) due to the plasma effect. In this case, an adjacent spectral peak is blue-shifted into this spectral range, giving rise to superlinear growth in the spectral power.

Figure 4.13(d) depicts the power scaling in the wavelength range, $1535.7\text{nm} \leq \lambda \leq 1542\text{nm}$. Similar to the power scaling integrated over the entire pulse spectrum, the relation is linear for $P_{peak}^\omega \leq 20\text{W}$. Further increases in power produce a slight increase in the growth trend due to SPM, which continues until $P_{peak}^\omega = 90\text{W}$. Subsequent increases in P_{peak}^ω produce saturation and even negative differential transmission due to TPA and FCA. A far different power-scaling trend for $1552.6\text{nm} \leq \lambda \leq 1559.6\text{nm}$ is shown in Fig. 4.13(e), where again there is linear growth for peak powers less than $P_{peak}^\omega \leq 20\text{W}$. However, in this spectral region it is followed by strong saturation at a much lower $P_{peak}^\omega = 44.8\text{W}$ and large negative differential transmission due to a combination of SPM, TPA, and FCA. Notably, the power transmitted in this spectral range at $P_{peak}^\omega = 137\text{W}$ is reduced by a factor of ≈ 2.12 compared to the maximum transmitted power at $P_{peak}^\omega = 44.8\text{W}$.

Peculiar power scaling is observed in the wavelength range, $1561\text{nm} \leq \lambda \leq 1570\text{nm}$, as shown in Figure 4.13(f). Here, evidence of transmission saturation occurs at much lower powers ($12\text{W} \leq P_{peak}^\omega \leq 40\text{W}$) and further increases in P_{peak}^ω increase the transmitted power without signs of any noticeable saturation. In this spectral region and at $P_{peak}^\omega \geq 40\text{W}$, SPM, TPA, and FCA do not affect the transmitted power as greatly as the spectral ranges depicted in Figs. 4.13(d)-(e). These distinct wavelength-dependent power scaling trends allow for a multitude of functionalities at discriminate wavelength regions on a single SOI platform device.

The observed wavelength-dependent power scaling trends are not unique to the $w = 340\text{nm}$ SOI waveguide. The various waveguide widths studied are observed to exhibit essentially the same wavelength-dependent power trend in each of these spectral windows. However, due to different A_{eff} , I_{max} inside the Si cores, and n_{eff} , some of the spectral features occur at different peak powers and have different amplitudes relative to adjacent features. In particular, the waveguides with $w = \{340, 470, 540\}\text{nm}$ scale with virtually identical trends, whereas the waveguide with $w = 260\text{nm}$ exhibited the greatest deviation in power scaling trends, which will be discussed subsequently.

Two-dimensional surfaces depicting the transmitted spectrum as a function of P_{peak}^ω for each of the nine waveguides studied are presented in Fig. 4.14(a)-(i). The strongest signatures of SPM are observed in Figs. 4.14(b)-(e), which correspond to $w = \{340, 470,$

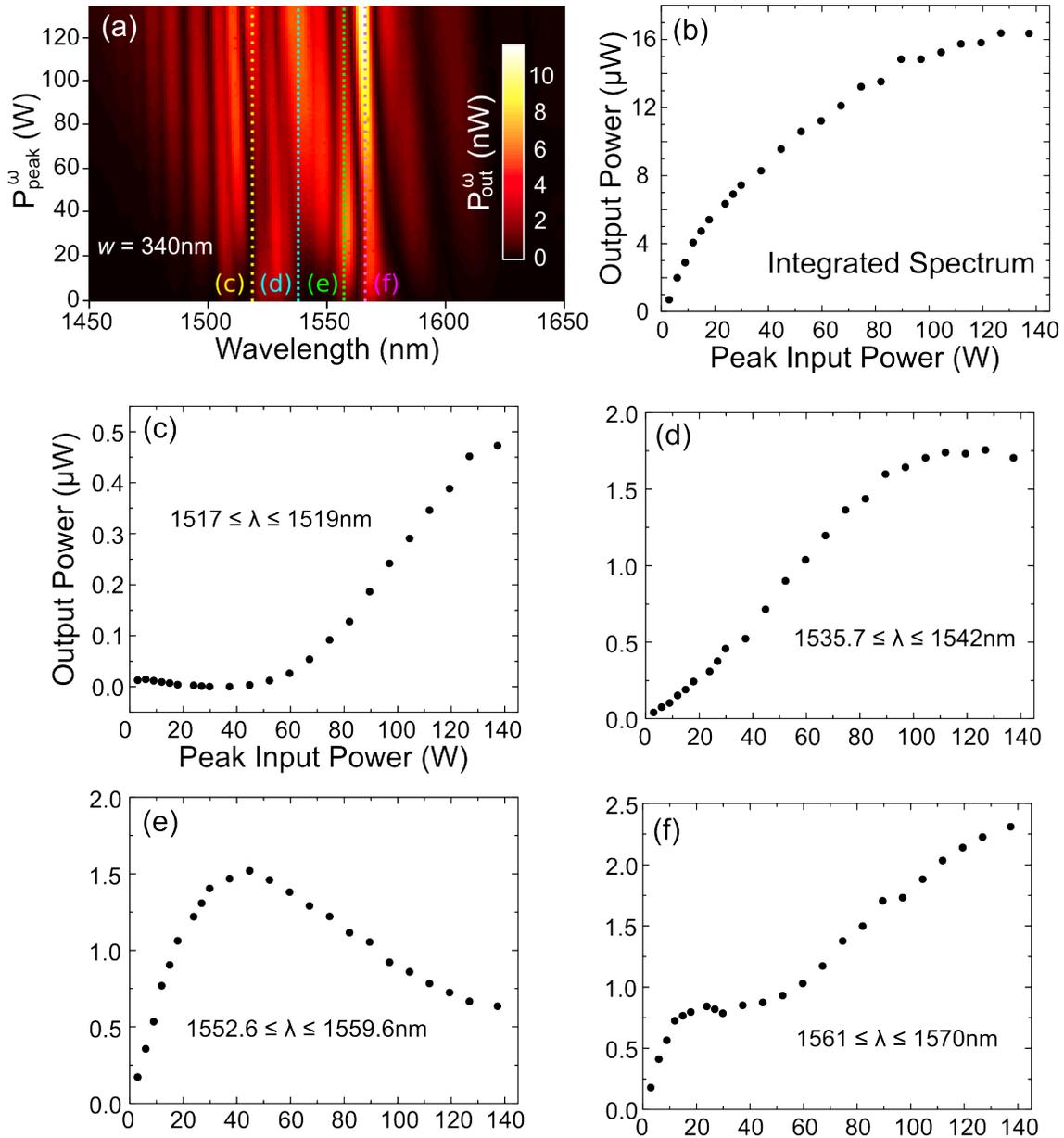


Figure 4.13: (a) Surface depicting the transmitted pulse spectrum versus input peak power for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 610\mu\text{m}$. (b) Total transmitted spectral power versus input peak power, depicting transmission saturation due to TPA and FCA. Transmitted power is also calculated for the spectral ranges: (c) $1517\text{nm} \leq \lambda \leq 1519\text{nm}$, (d) $1535.7\text{nm} \leq \lambda \leq 1542\text{nm}$, (e) $1552.6\text{nm} \leq \lambda \leq 1559.6\text{nm}$, and (f) $1561\text{nm} \leq \lambda \leq 1570\text{nm}$.

540, 760}nm. In these waveguides, the small Si core gives rise to a small A_{eff} and a high intensity. Interestingly, there is virtually no signature of SPM in Fig. 4.14(a), with $w = 260$ nm. Instead, the maximum of this spectrum (which appears as yellow in the surface plot colormap) undergoes a significant blue shift from $\lambda = 1522.3$ nm to $\lambda = 1514.9$ nm, corresponding to $\Delta\lambda = -7.4$ nm. Such an extreme blue shift is strong evidence of a large TPA-induced free-carrier population ($N \sim 10^{19} - 10^{20}$ cm $^{-3}$), which severely limits any SPM from accumulating within the Si core. In other words, the peak power of the pulse is reduced via TPA and FCA very early on in the waveguide, before a significant signature of SPM can accumulate. The waveguides depicted in Figs. 4.14(f)-(i) (i.e. 1000 nm $\leq w \leq 4150$ nm) are much larger waveguides with significantly lower confinement, and can additionally be operating with multiple modes. Therefore, as the width is increased for these waveguides, SPM-induced pulse broadening becomes less pronounced, particularly in the spectral range, 1535 nm $\leq \lambda \leq 1550$ nm, and for $\lambda \leq 1480$ nm and $\lambda \geq 1580$ nm. For the largest waveguide with $w = 4150$ nm, the SPM-induced spectral features are very weak compared to the main spectral components of the pulse.

To confirm the nonlinear scaling of the sidebands, the output power of the $w = 340$ nm waveguide at a representative wavelength of $\lambda = 1465$ nm is plotted as a function of the input peak power on a log-log scale, as shown in Fig. 4.15. The slope of a linear fit to the data points is found to be 3.03 ± 0.01 , confirming the presence of a $\chi^{(3)}$ SPM process.

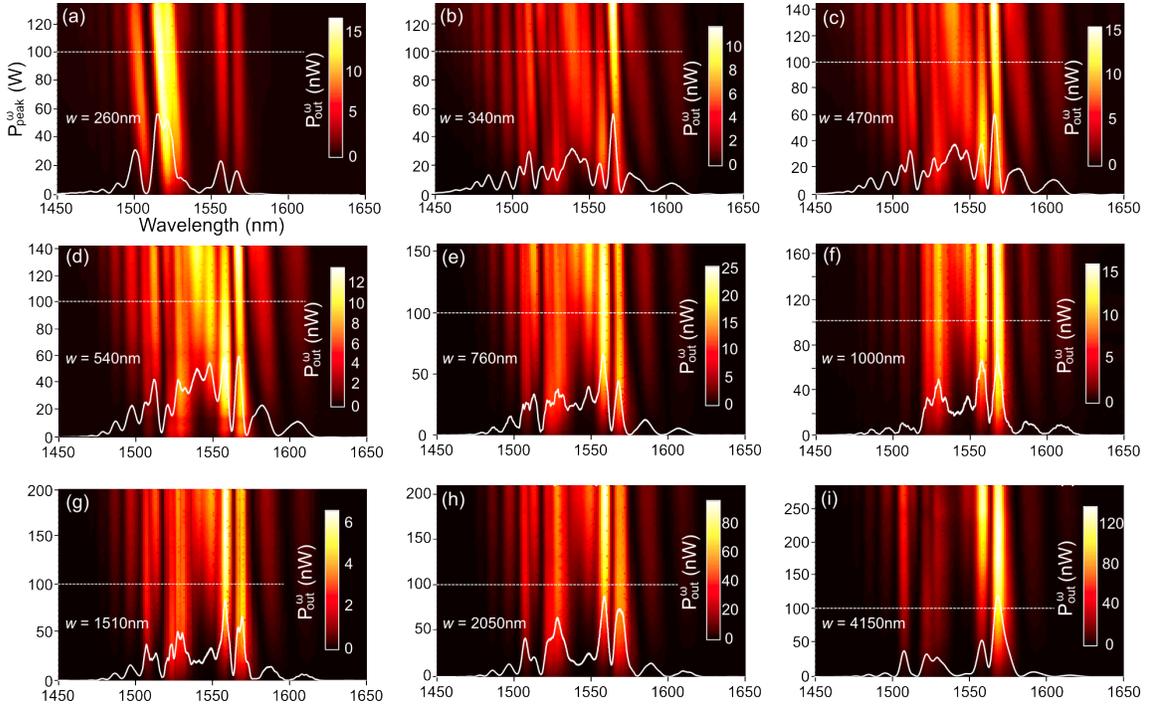


Figure 4.14: Surfaces depicting the transmitted pulse spectrum versus input peak power for waveguides with $h = 340$ nm, $L = 610\mu\text{m}$, and $w =$ (a) 260nm, (b) 340nm, (c) 470nm, (d), 540nm, (e) 760nm, (f) 1000nm, (g) 1510nm, (h) 2050nm, and (i) 4150nm. Sections of the transmitted spectrum taken at $P_{peak}^{\omega} = 100$ W are plotted in white.

Two-dimensional surface plots for waveguides with $h = 340\text{nm}$ and $L = 200\mu\text{m}$ are presented in Fig. 4.16. In this case, the pulses propagate a shorter distance by approximately $\Delta L = 410\mu\text{m}$ than in the $L = 610\mu\text{m}$ waveguide, and therefore the interaction length for the SPM to accumulate is greatly reduced. In this case, there is moderate SPM observed in waveguides with $w = \{340, 470, 540\}$ nm; however, the spectral power of frequency components in the range, $1535\text{nm} \leq \lambda \leq 1550\text{nm}$, and at the sidebands $\lambda \leq 1480\text{nm}$ and $\lambda \geq 1580\text{nm}$ generated by SPM are much weaker than what was observed for $L = 610\mu\text{m}$. There is no noticeable pulse broadening for the $w = 260\text{nm}$ waveguide, signifying that the nonlinear loss due to TPA and FCA is still significant enough to prevent SPM from building up. As the waveguide width is increased to above 540nm , only very weak side lobes are observed, and the spectral power of these side lobes is significantly less than the main spectral components of the pulse.

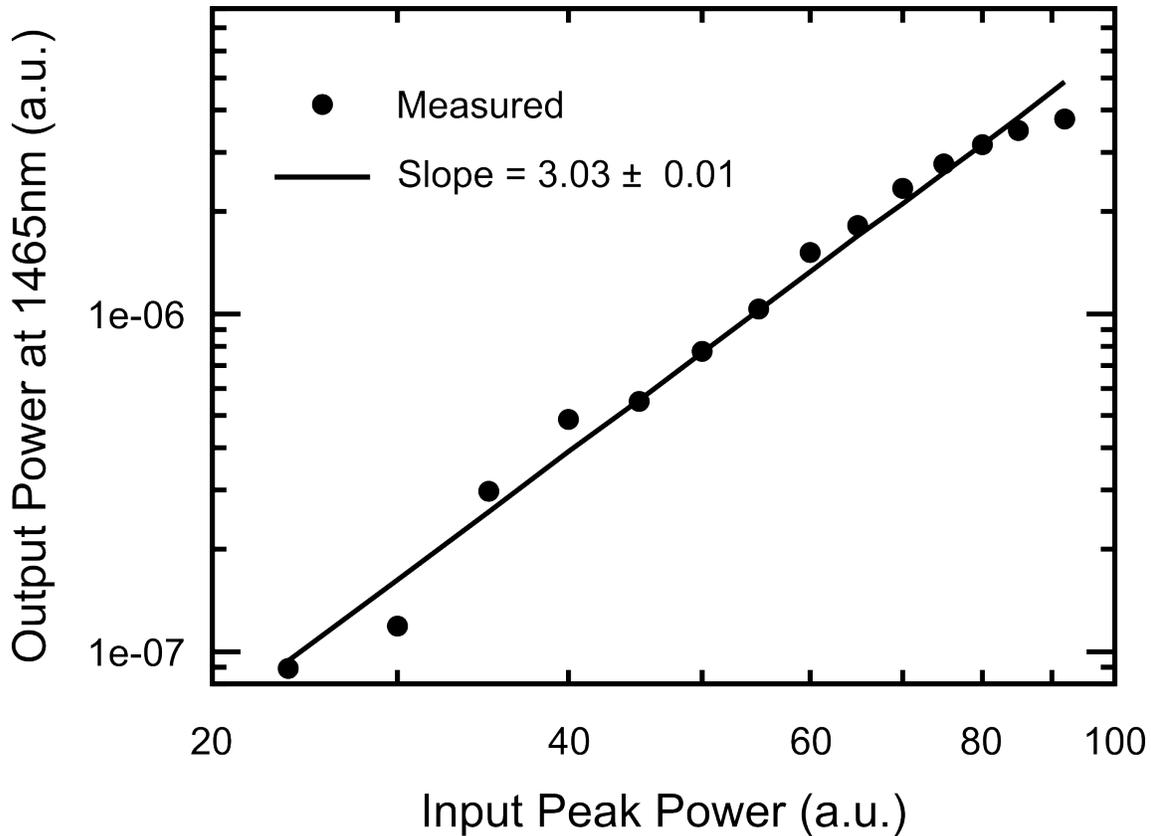


Figure 4.15: Plot of detected power at $\lambda = 1465\text{nm}$ versus peak input power, shown on a logarithmic x - and y -axis, for a waveguide with $h = 340\text{nm}$, $w = 340\text{nm}$, and $L = 200\mu\text{m}$.

4.5 Thermo-optic Tuning in a Fabry-Pérot Resonator via Nonlinear Loss

In the interest of developing high integration density nonlinear circuitry, it is essential that the footprint of the photonic modulators be reduced to as small an area as possible.

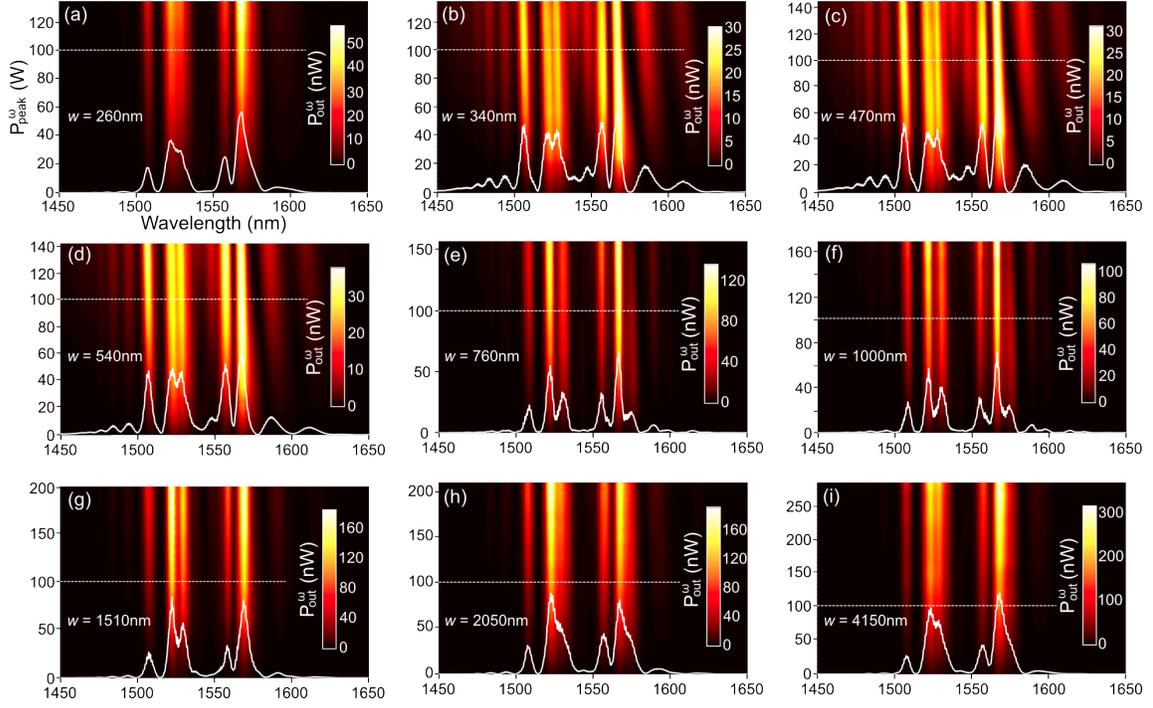


Figure 4.16: Surfaces depicting the transmitted pulse spectrum versus input peak power for waveguides with $h = 340\text{nm}$, $L = 200\mu\text{m}$, and $w =$ (a) 260nm, (b) 340nm, (c) 470nm, (d), 540nm, (e) 760nm, (f) 1000nm, (g) 1510nm, (h) 2050nm, and (i) 4150nm. Section of the transmitted spectrum taken at $P_{peak}^{\omega} = 100\text{W}$ are plotted in white.

In Section 4.3, it was demonstrated that for high input peak powers, a significant fraction (more than half) of the pulse power can be depleted by TPA and FCA. Therefore, to obtain strong signal modulation in a compact device footprint, a high optical power must be absorbed in a very small device volume. Due to silicon's indirect bandgap, nonradiative recombination is the dominant free-carrier relaxation mechanism, which was measured to occur on timescale of $125\text{ps} \leq \tau_r \leq 312\text{ps}$. As the free-carriers relax, thermal energy is deposited directly in the device's waveguide core, modifying the silicon's refractive index. In the case of SOI waveguides, the BOX layer has a very high thermal resistivity ($R_{\lambda} = 0.725\text{m} \cdot \text{K/W}$), and the temperature of the device can increase substantially [41]. While this can have minimal influence on straight waveguide modulators operating on the principle of nonlinear loss, the thermo-optic effect becomes more pronounced for nanoscale Si-based resonant or interference devices where a slight refractive index change shifts the device's resonant wavelength.

Although there have been extensive efforts to demonstrate ultrafast modulation devices based on TPA and FCA, and to demonstrate slower tuning devices based on the thermo-optic effect, these investigations have not been interlinked. Most ultrafast modulation experiments are performed in long device lengths ranging from $100\mu\text{m} \leq L \leq 10\text{cm}$, where the large volume of the device moderates the temperature of the device and thermo-optic effects are minimal. However, for compact, isolated SOI photonic devices, the interplay

between both TPA and FCA and the thermo-optic effect is significant and influences the operational characteristics of modulators, especially in resonant structures, where a small change in the effective refractive index of the mode can shift its resonant frequency.

To investigate this interplay, a resonator with a small quality-factor was studied and is depicted schematically in Fig. 4.17(a). The resonator consists of a trapezoid-shaped waveguide, which is referred to as a trapezoid Fabry-Pérot resonator (TFPR). Silicon-on-insulator TFPR waveguides with relatively large cross-sectional dimensions of $w \times h = 2.0\mu\text{m} \times 0.34\mu\text{m}$ were fabricated at the corner of a SOI wafer as described in Section 3.6 and a SEM of the sample is shown in Fig. 4.17(b). The most extreme interplay between nonlinear loss and thermal effects is studied by performing the measurements on the shortest TFPR, with a length of $L = 7.80\mu\text{m}$ (this is the length measured from the centre of the input facet to the centre of the output facet). This structure had a volume of only $V = 5.31\mu\text{m}^3$ and has very poor heat-sinking, as it is surrounded by air and a thermal oxide. The combination of small volume and poor heat-sinking make this structure excellent for studying thermal effects with high sensitivity.

Both a tunable CW diode laser ($1509\text{nm} \leq \lambda \leq 1581\text{nm}$) and the $\lambda = 1550\text{nm}$ femtosecond laser were used simultaneously to characterize the TFPR. The femtosecond pulses were propagated co-linearly with the CW radiation and both beams were end-fire coupled to the TFPR using a $NA = 0.85$ microscope objective. The polarization of both lasers was oriented to excite the quasi-TM waveguide modes. The transmitted radiation was collected with a lensed single mode optical fibre and was delivered to an OSA. The angle of incidence of the excitation was chosen to be perpendicular to the input facet of the waveguide, as shown in Fig. 4.17(a).

Interestingly, the width of the TFPR was large enough that waveguiding occurred in a similar manner as a fibre optical waveguide, where total internal reflection confines the signal to the fibre core. When high intensity femtosecond pulses were coupled to the TFPR, strong third-harmonic generation was excited (discussed in the next chapter), which appeared as bright green light. Imaging of the third-harmonic emission that was scattered from the waveguide sidewalls enabled accurate visualization of the path of optical power flow. A top-view image of scattered green light from a $L = 31\mu\text{m}$ long TFPR is shown in Fig. 4.18(a). The optical mode followed a zig-zag pattern as it underwent multiple total internal reflections at the waveguide sidewalls since the 45° angle of incidence for each reflection exceeded the 20.2° critical angle for total internal reflection. In other words, the radiation was not transmitted by a conventional waveguide mode, but as radiation modes that underwent total internal reflection at the sidewalls. As shown in Fig. 4.18(b), a distribution of the time-averaged intensity distribution at the centre of the TFPR obtained via finite-difference time-domain (FDTD) simulations exhibits a similar zig-zag trend. Choosing the angle of incidence of the excitation to be normal to the input waveguide end facet resulted in the propagation direction of the waveguide mode to be normally incident on the output facet of the resonator. A reflection coefficient of $R = 0.29$

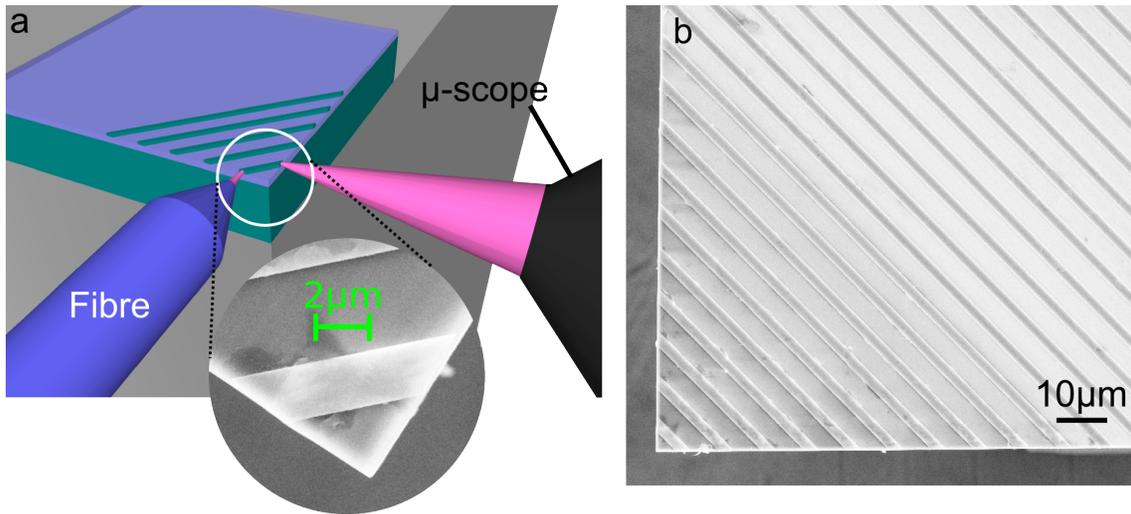


Figure 4.17: (a) Schematic depiction of the sample under test and the excitation and detection scheme. A microscope objective (shown on right) couples light into the TFPR and a lensed single mode optical fibre (shown on left), oriented perpendicularly to the angle of excitation, collects the transmitted radiation. A scanning electron micrograph of the $L = 7.80\mu\text{m}$ TFPR is shown in the inset. (b) Scanning electron micrograph of the fabricated sample, with multiple TFPR lengths.

was calculated from the broadband transmission data from the simulation.

It was essential to first characterize the TFPR wavelength response in the absence of the femtosecond laser pulse excitation (i.e. the passive response). Shown in Fig. 4.19 is an experimental scan of a resonance of the $L = 7.80\mu\text{m}$ TFPR combined with a calculated wavelength response using the FDTD technique. The Q of this ultracompact TFPR was measured to be 81, while the theoretical value was calculated to be 91. The slight disagreement between the experimental and theoretical Q values is attributed to loss arising from roughness-induced scattering at the TFPR sidewalls. The magnitude of the Q is limited by the short optical path length of the resonator ($n_{eff}L = 42.5\mu\text{m}$) and the low reflection coefficient of the waveguide end facets ($R = 0.29$).

Next, the femtosecond laser pulses were used to obtain the power-dependent transmission curve of the $L = 7.80\mu\text{m}$ TFPR, which is shown in Fig. 4.20(a). Despite the short interaction length within the structure, a transition from the linear transmission regime to the saturation regime was clearly evident. The nonlinear loss arises from both TPA and FCA acting in unison to attenuate the radiation and was calculated by extrapolating the linear region of the curve and subtracting the measured values from the linear trend. For an input power, $P_{ave}^\omega = 5.87\text{mW}$, approximately, $P_{loss} = 2.86\text{mW}$, was lost to TPA and FCA. Input powers less than $P_{ave}^\omega = 1.0\text{mW}$ were insufficient to induce TPA and FCA, and the radiation coupled into the TFPR only experienced linear losses due to scattering at the waveguide sidewalls.

Active tuning of the TFPR was demonstrated by collinearly coupling both the femtosecond laser pulses and the CW radiation into the device, where the polarization of

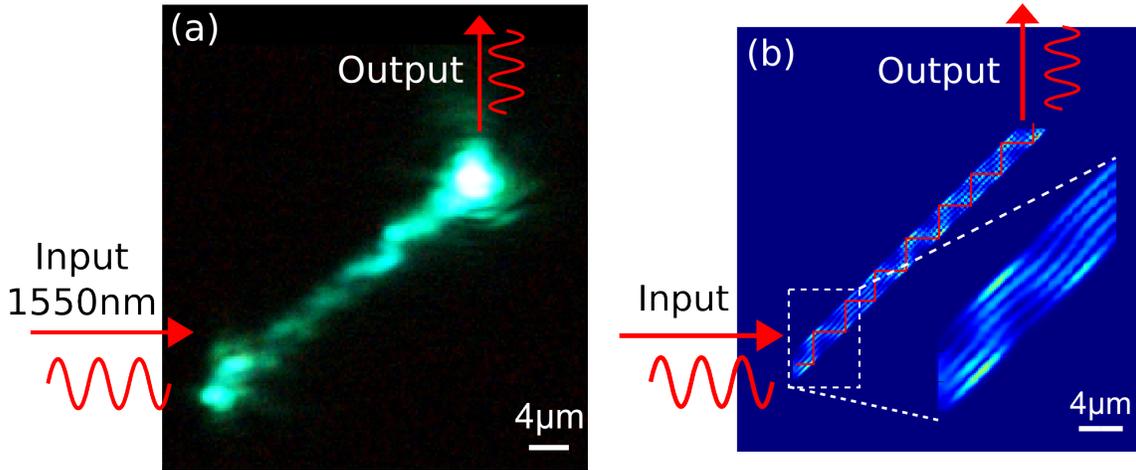


Figure 4.18: (a) Top-view camera image of scattered third-harmonic radiation from the waveguide interfaces following a zig-zag pattern as the mode undergoes total internal reflection at the waveguide interfaces. (b) Time-averaged intensity (I) distribution in the waveguide obtained by finite-difference time-domain simulation, verifying the multiple total internal reflections as the mode propagates from the input facet to the output. The length of the depicted TFPR is $L = 31.1\mu\text{m}$.

both beams was oriented to excite the quasi-TM waveguide mode. Here, the ultrafast pulses were used to generate free-carriers, which modulated the transmission of the CW radiation. Since there was an overlap between the spectra of the femtosecond pulses and the CW radiation, it was necessary to isolate only the induced modulation on the CW radiation. For a single femtosecond laser pulse, the coupled peak power was varied up to a maximum of $P_{peak}^\omega = 700\text{W}$ and this power was distributed over a wide spectrum ($1510\text{nm} \leq \lambda \leq 1610\text{nm}$). Conversely, the CW laser diode emits a single mode with a line width, $\Delta\lambda = 20\text{pm}$. An OSA was used as a narrow bandpass filter with $\Delta\lambda = 20\text{pm}$ centred at the wavelength of the CW radiation. This scheme effectively filtered out 99.99% of the power from the femtosecond laser pulses, and the collected radiation power from the CW laser diode exceeded that from the femtosecond laser pulse, in this same spectral range, by more than an order of magnitude. This technique allows for high resolution spectral sweeps to be performed with the tunable diode with minimal power interference from the femtosecond laser pulse.

By gradually increasing the input femtosecond pulse power and scanning the TFPR spectral response with the tunable CW radiation, it was possible to probe both the loss resulting from TPA and FCA and the change in refractive index, Δn , due to the thermo-optic effect. Free-carriers were produced when high power femtosecond pulses excited the resonator, and thermal energy was deposited in the TFPR as the free-carriers underwent nonradiative recombination. As shown for the $L = 7.80\mu\text{m}$ TFPR in Fig. 4.21, as the power from the femtosecond laser pulse train was increased, the resonant peak response of the TFPR underwent a red-shift and its amplitude was attenuated. For a low input power, $P_{ave}^\omega = 1.16\text{mW}$, which was close to the linear regime of the transmission curve in

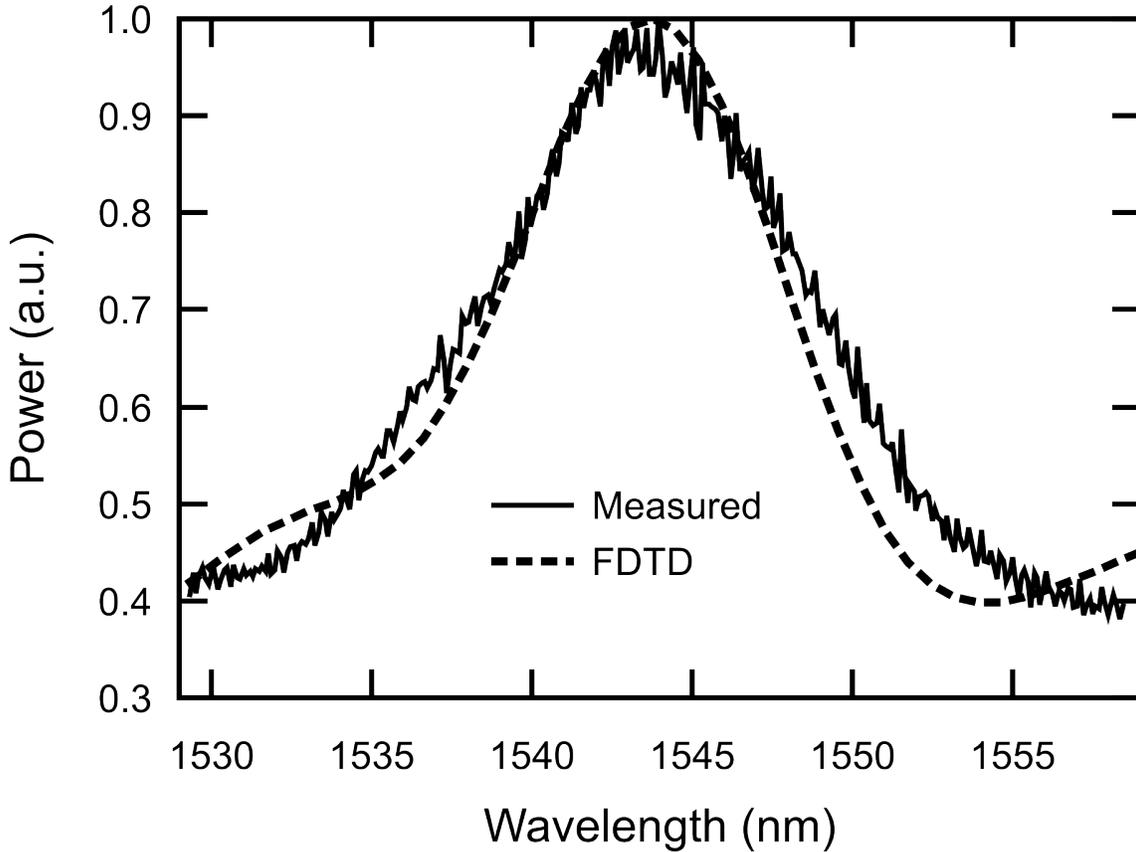


Figure 4.19: Measured broadband transmission spectrum from a TFPR with $L = 7.80\mu\text{m}$ (black) and the calculated spectrum obtained using finite-difference time-domain simulations.

Fig. 4.20(a), the nonlinear optical interaction was weak, and the resonance was virtually unchanged from the case when only the CW radiation was coupled into the TFPR. For a high coupled power of $P_{ave}^\omega = 11.3\text{mW}$, the resonant peak was red-shifted by $\Delta\lambda=7.57\text{nm}$ and was attenuated by 66%.

Notably, a plasma effect resulting from the excited free-carriers would give rise to a blue-shifted resonance as the refractive index is reduced. While the optical Kerr effect could in principle give rise to a red-shifted resonance peak, it is only present during femtosecond pulse excitation. This effect can be ruled out since the CW radiation probes the time-averaged response of the TFPR. The thermo-optic effect is the only possible effect that would give rise to a red-shift in the resonance that would be detectable with a CW probe. Previous investigations into the thermal dynamics of SOI waveguide devices determined a response time of $\tau_{thermal} \sim 1\mu\text{s}$ [42]. This is much longer than the period between femtosecond pulses ($T_{pulse} = 11.1\text{ns}$), and it is concluded that the temperature of the device is time-invariant.

Log-log plots relating the input power, P_{ave}^ω , to the wavelength shift, $\Delta\lambda$, and amplitude attenuation of the resonance are shown in Fig. 4.21(b). The $\Delta\lambda$ versus P_{ave}^ω dependence follows a line of slope of 2.01, demonstrating that $\Delta n \propto P_{in}^2$. The $\Delta\lambda$ (where

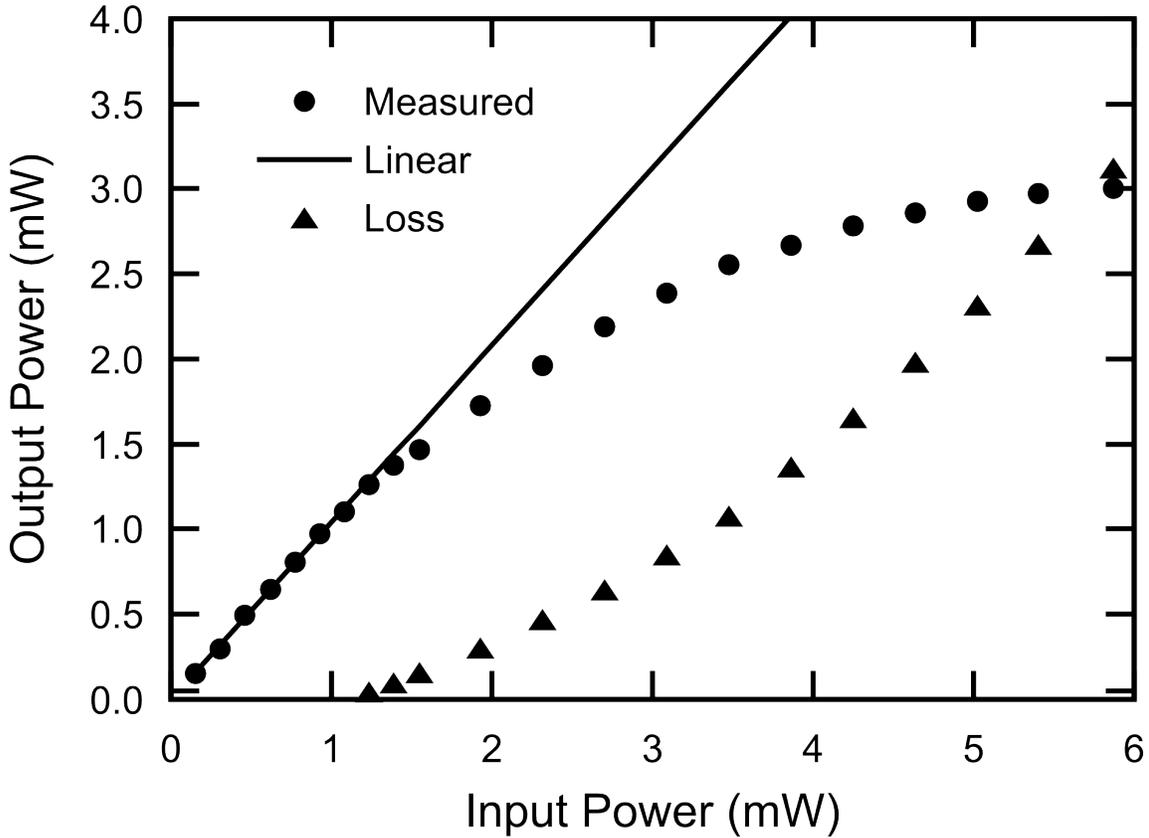


Figure 4.20: Power transmission saturation plots for a TFPR with $L = 7.80\mu\text{m}$, displaying a linear region and a nonlinear loss region. The linear region is extrapolated and the experimental values are subtracted from this line to obtain the nonlinear loss.

$\Delta\lambda \propto \Delta n$) is therefore proportional to the number density of free-carriers excited in the device. This demonstrates that the dominant source of heating is nonradiative recombination of free-carriers and that other factors, such as thermal expansion are negligible. This was confirmed by performing a thermal stress analysis on the structure using COMSOL, and it was determined that the maximum thermal expansion of any point on the TFPR was $\delta x = 0.2\text{nm}$, which was insignificant in altering the resonant frequency of the TFPR. However, the amplitude attenuation does not follow the same dependence, and it is concluded that a combination of the different power-dependencies of TPA and FCA, along with the varying pulse distortion as the TFPR resonance is shifted are responsible for these results.

Finite-difference time-domain simulations were performed to gain an understanding of the influence of operating temperature on the resonant characteristics of the TFPR. These simulations neglected nonlinear loss mechanisms, and therefore, were only used to study the influence of operating temperature on the broadband response of the FPTR. In specific, an increase in temperature produces an increase in the refractive index via the thermo-optic coefficient, $\Delta n/\Delta T = 1.86 \times 10^{-4}/\text{K}$. By varying ΔT systematically, as depicted in Fig. 4.22(a), the temperature change required to shift the resonant wavelength

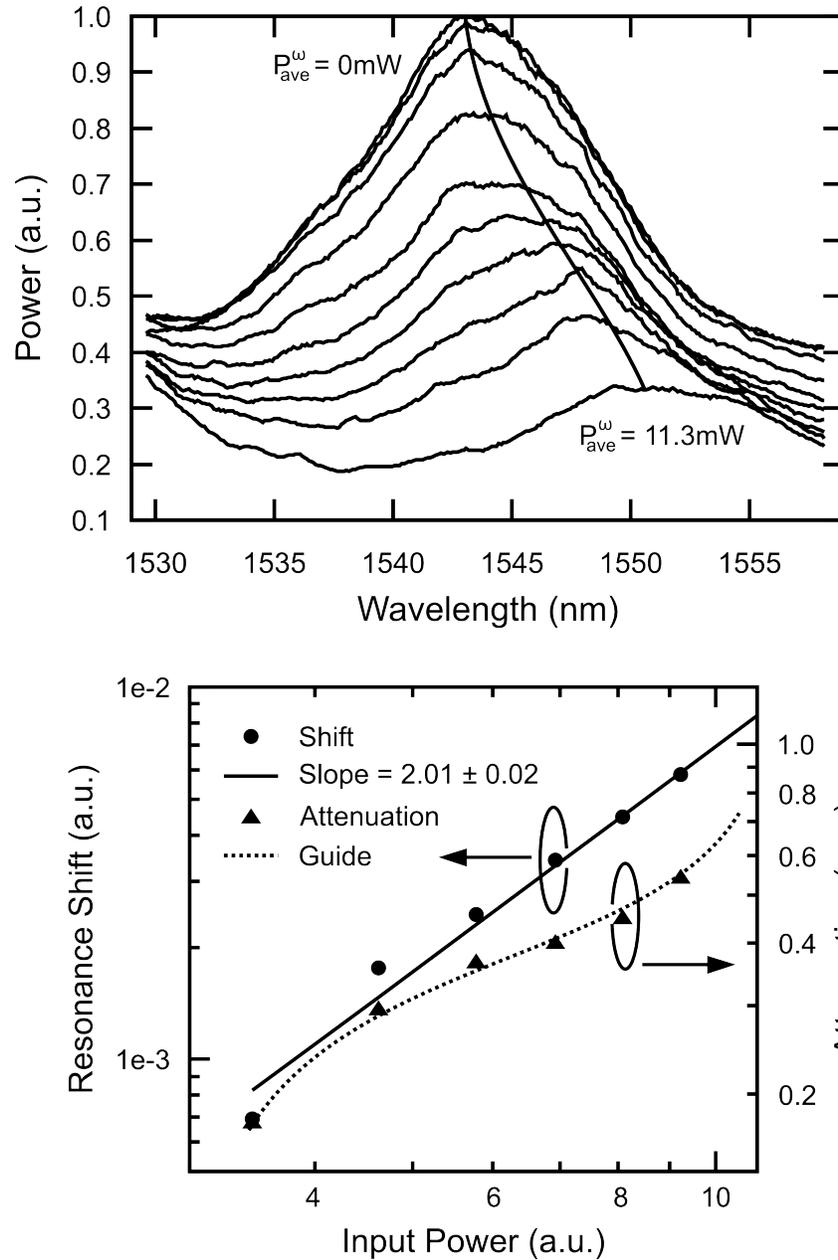


Figure 4.21: (a) Broadband transmission spectrum of the TFPR as the femtosecond laser power coupled to the device is increased from $P_{ave}^{\omega} = 0\text{mW}$ to $P_{ave}^{\omega} = 11.3\text{mW}$. (b) Logarithmic x - and y -axis plots relating the input power to the wavelength shift and resonance amplitude attenuation. The wavelength shift has a slope of 2.01 ± 0.02 , while the attenuation does not show a linear trend.

of the TFPR by $\Delta\lambda$ was determined, and was found to follow a linear relation of the form, $\Delta\lambda_{max} = 7.26 \times 10^{-2}\Delta T$. Through these results, the steady-state temperature of the TFPR was estimated to increase by $\Delta T = 104\text{K}$ when the input power was $P_{ave}^\omega = 11.3\text{mW}$. To gain further insight, a thermal stress analysis was performed using COMSOL Multiphysics to calculate the temperature distribution and the thermal expansion of the device resulting from uniform optical power dissipation². An exemplary temperature distribution for a thermal power of $P_{thermal} = 3.0\text{mW}$ is shown in Fig. 4.22(b), which indicates that the maximum temperature increase in the device is $\Delta T = 91.5\text{K}$. By varying the thermal power, the linear influence of the thermal power source on the device operating temperature was found to be, $\Delta T = 30.5P_{thermal}[\text{K}]$. Based on this trend, it is estimated that for an input power of $P_{ave}^\omega = 11.3\text{mW}$, approximately $P_{thermal} = 3.50\text{mW}$ remains as a steady-state heat source in the TFPR.

4.6 Summary

This chapter focused on obtaining a thorough understanding of the nonlinear optical properties of SOI waveguides. By using different experimental techniques and detection equipment, it was possible to observe and study several nonlinear optical effects in SOI waveguides. Nonlinear loss mechanisms, including TPA and FCA, were investigated and their implications on both linear routing and all-optical modulation were determined. The free-carrier recovery time in the SOI waveguides was measured to be $\tau_r = 265\text{ps}$, which would limit the modulation bandwidth to approximately 3.8GHz. Strong XAM was observed in a $L = 65\mu\text{m}$ waveguide at $P_{peak}^\omega \leq 33\text{W}$, but as the waveguide length was decreased to $L = 25\mu\text{m}$ and $L = 6.5\mu\text{m}$, peak powers up to $P_{peak}^\omega \sim 40\text{W}$ were insufficient for exciting a strong XAM signature, demonstrating limitations in the minimum footprint area of SOI waveguide modulators.

Subsequent measurements were performed to investigate the nonlinear refractive index of SOI waveguides, which manifested itself as SPM. Strong spectral broadening due to SPM was generated in a $L = 610\mu\text{m}$ SOI waveguide, but the SPM signature became considerably weaker for $L = 200\mu\text{m}$ waveguides. The strongest pulse broadening was observed in waveguides with widths in the range $340\text{nm} \leq w \leq 760\text{nm}$. For a smaller waveguide with a width, $w = 260\text{nm}$, strong TPA gave rise to high FCA and both of these effects saturated the SPM. For waveguides with $w > 760\text{nm}$, a combination of lower mode confinement and multimode operation give rise to less-pronounced pulse broadening. Several interesting power-scaling trends were observed in certain spectral bands, which showed potential for ultrafast wavelength-dependent all-optical switching and other functionalities. Notably, this would allow for entirely unique wavelength-dependent nonlinear functionalities to be performed in a single structure, drastically increasing the potential

²These simulations were performed with assistance from Michael Nielsen.

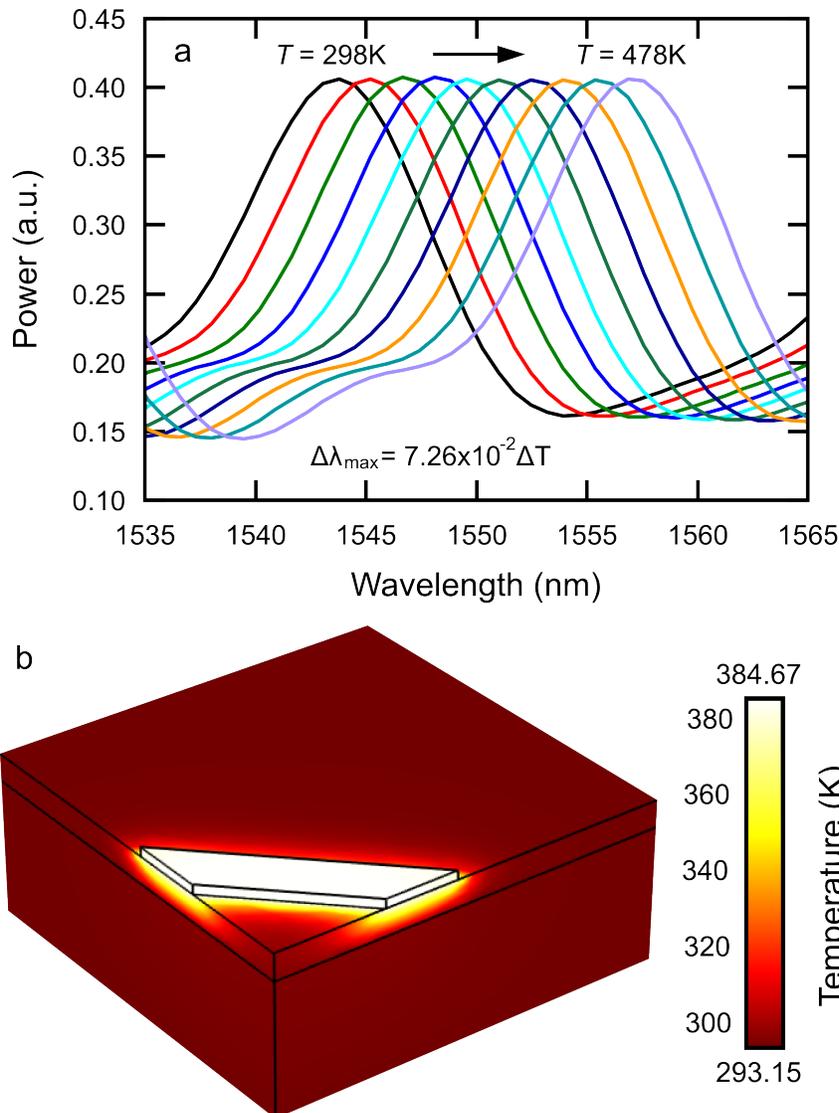


Figure 4.22: (a) Broadband transmission spectrum for operating temperatures ranging from $298\text{K} \leq T \leq 478\text{K}$ in steps of $\Delta T = 20\text{K}$, obtained via finite-difference time-domain simulations, in the absence of nonlinear loss. The peak wavelength of the resonance redshifts linearly with increasing temperature according to the relation $\Delta\lambda_{\max} = 7.26 \times 10^{-2} \Delta T$ [nm]. (b) Temperature distribution for a thermal power of 3mW distributed uniformly in the TFPR.

functionality of nonlinear SOI waveguide devices for integrated optical circuitry.

A final investigation was performed to study the passive resonant characteristics, ultrafast nonlinear loss mechanisms, and steady-state temperature change arising from non-radiative recombination of free-carriers in an ultracompact TFPR. By coupling up to $P_{ave}^\omega = 11.3\text{mW}$ from a femtosecond laser into the resonator, the amplitude of the TFPR resonance was attenuated by 65.9% and its spectral position was red-shifted by $\Delta\lambda = 7.57\text{nm}$.

These investigations demonstrate some of the capabilities and limitations of silicon photonics circuitry operating at telecommunications band wavelengths. Evidently, SOI waveguides enable low-loss signal transmission across chip-scale length scales, which is highly desirable for passive circuitry. However, the rather large footprint area and slow recovery time of SOI photonic waveguide modulators limit the utility of these structures for high-density, all-optical, active circuitry, where a compact footprint area and ultrafast recovery time are important. Therefore, it is necessary to investigate alternative optical circuits that are still compatible with electronics and Si photonics technology, but offer enhanced performance in these areas.

References

- [1] A. R. Motamedi, A. H. Nejadmalayeri, A. Khilo, F. X. Kärtner, and E. P. Ippen. Ultrafast nonlinear optical studies of silicon nanowaveguides. *Opt. Express*, 20:4085–4101, 2012.
- [2] K. S. Kim, R. H. Stolen, W. A. Reed, and K. W. Quoi. Measurement of the nonlinear index of silica-core and dispersion-shifted fibers. *Opt. Lett.*, 19:257–259, 1994.
- [3] S. Palomba, H. Harutyunyan, J. Renger, R. Quidant, N. F. van Hulst, and L. Novotny. Nonlinear plasmonics at planar metal surfaces. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369:3497–3509, 2011.
- [4] H. K. Tsang, C. S. Wong, T. K. Liang, I. E. Day, S. W. Roberts, A. Harpin, J. Drake, and M. Asghari. Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength. *Appl. Phys. Lett.*, 80:416–418, 2002.
- [5] M. Dinu, F. Quochi, and H. Garcia. Third-order nonlinearities in silicon at telecom wavelengths. *App. Phys. Lett.*, 82:2954, 2003.
- [6] D. J. Moss, L. Fu, I. Littler, and B. J. Eggleton. Ultrafast all-optical modulation via two-photon absorption in silicon-on-insulator waveguides. *Electron. Lett.*, 41:320–321, 2005.
- [7] T. Liang, L. Nunes, T. Sakamoto, K. Sasagawa, T. Kawanishi, M. Tsuchiya, G. Priem, D. van Thourhout, P. Dumon, R. Baets, and H. Tsang. Ultrafast all-optical switching by cross-absorption modulation in silicon wire waveguides. *Opt. Express*, 13:7298–7303, 2005.
- [8] T. K. Liang, L. R. Nunes, M. Tsuchiya, K. S. Abedin, T. Miyazaki, D. van Thourhout, W. Bogaerts, P. Dumon, R. Baets, and H. K. Tsang. High speed logic gate using two-photon absorption in silicon waveguides. *Opt. Commun.*, 265:171–174, 2006.
- [9] R. Espinola, J. Dadap, Jr. R. Osgood, S. McNab, and Y. Vlasov. C-band wavelength conversion in silicon photonic wire waveguides. *Opt. Express*, 13:4341–4349, 2005.

- [10] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, T. Tsuchizawa, T. Watanabe, J. I. Takahashi, and S. I. Itabashi. Four-wave mixing in silicon wire waveguides. *Opt. Express*, 13:4629–4637, 2005.
- [11] M. Foster, A. Turner, J. Sharping, B. Schmidt, M. Lipson, and A. Gaeta. Broad-band optical parametric gain on a silicon photonic chip. *Nature*, 441:960–963, 2006.
- [12] Q. Lin, J. Zhang, P. M. Fauchet, and G. P. Agrawal. Ultrabroadband parametric generation and wavelength conversion in silicon waveguides. *Opt. Express*, 14:4786–4799, 2006.
- [13] M. A. Foster, A. C. Turner, R. Salem, M. Lipson, and A. L. Gaeta. Broad-band continuous-wave parametric wavelength conversion in silicon nanowaveguides. *Opt. Express*, 15:12949–12958, 2007.
- [14] R. Salem, M. Foster, A. Turner, D. Geraghty, M. Lipson, and A. L. Gaeta. Signal regeneration using low-power four-wave mixing on silicon chip. *Nat. Photon.*, 2:35–38, 2008.
- [15] R. Dekker, A. Driessen, T. Wahlbrink, C. Moormann, J. Niehusmann, and M. Först. Ultrafast kerr-induced all-optical wavelength conversion in silicon waveguides using 1.55 μm femtosecond pulses. *Opt. Express*, 14:8336–8346, 2006.
- [16] O. Boyraz, P. Koonath, V. Raghunathan, and B. Jalali. All optical switching and continuum generation in silicon waveguides. *Opt. Express*, 12:4094–4102, 2004.
- [17] I. W. Hsieh, X. Chen, J. I. Dadap, N. C. Panoiu, Jr. R. M. Osgood, S. J. McNab, and Y. A. Vlasov. Cross-phase modulation-induced spectral and temporal effects on co-propagating femtosecond pulses in silicon photonic wires. *Opt. Express*, 15:1135–1146, 2007.
- [18] O. Boyraz, T. Indukuri, and B. Jalali. Self-phase-modulation induced spectral broadening in silicon waveguides. *Opt. Express*, 12:829–834, 2004.
- [19] G. W. Rieger, K. S. Virk, and J. F. Young. Nonlinear propagation of ultrafast 1.5 μm pulses in high-index-contrast silicon-on-insulator waveguides. *Applied Physics Letters*, 84:900–902, 2004.
- [20] A. Cowan, G. Rieger, and J. Young. Nonlinear transmission of 1.5 μm pulses through single-mode silicon-on-insulator waveguide structures. *Opt. Express*, 12:1611–1621, 2004.
- [21] E. Dulkeith, Y. A. Vlasov, X. Chen, N. C. Panoiu, and Jr. R. M. Osgood. Self-phase-modulation in submicron silicon-on-insulator photonic wires. *Opt. Express*, 14:5524–5534, 2006.

- [22] I. W. Hsieh, X. Chen, J. I. Dadap, N. C. Panoiu, R. M. Osgood, S. J. McNab, and Y. A. Vlasov. Ultrafast-pulse self-phase modulation and third-order dispersion in silicon photonic wire-waveguides. *Opt. Express*, 14:12380–12387, 2006.
- [23] L. Yin and G. P. Agrawal. Impact of two-photon absorption on self-phase modulation in silicon waveguides. *Opt. Lett.*, 32:2031–2033, 2007.
- [24] R. Claps, D. Dimitropoulos, Y. Han, and B. Jalali. Observation of raman emission in silicon waveguides at 1.54 μm . *Opt. Express*, 10:1305–1313, 2002.
- [25] R. Claps, D. Dimitropoulos, and B. Jalali. Stimulated raman scattering in silicon waveguides. *Electron. Lett.*, 38:1352–354, 2002.
- [26] R. Claps, D. Dimitropoulos, V. Raghunathan, Y. Han, and B. Jalali. Observation of stimulated raman amplification in silicon waveguides. *Opt. Express*, 11:1731–1739, 2003.
- [27] H. Rong, A. Liu, R. Jones, O. Cohen, D. Hak, R. Nicolaescu, A. Fang, and M. Paniccia. An all-silicon raman laser. *Nature*, 433:292–294, 2005.
- [28] H. Rong, R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang, and M. Paniccia. A continuous-wave raman silicon laser. *Nature*, 433:725–728, 2005.
- [29] R. L. Espinola, M. C. Tsai, James T. Yardley, and Jr. Osgood, R.M. Fast and low-power thermo-optic switch on thin silicon-on-insulator. *IEEE Photon. Tech. Lett.*, 15:1366–1368, 2003.
- [30] T. Baehr-Jones, M. Hochberg, C. Walker, E. Chan, D. Koshinz, W. Krug, and A. Scherer. Analysis of the tuning sensitivity of silicon-on-insulator optical ring resonators. *Lightwave Technology, Journal of*, 23:4215–4221, 2005.
- [31] T. Chu, H. Yamada, S. Ishida, and Y. Arakawa. Compact 1xn thermo-optic switches based on silicon photonic wire waveguides. *Opt. Express*, 13:10109–10114, 2005.
- [32] M. W. Pruessner, T. H. Stievater, M. S. Ferraro, and W. S. Rabinovich. Thermo-optic tuning and switching in soi waveguide fabry-perot microcavities. *Opt. Express*, 15:7557–7563, 2007.
- [33] Z. Jiatae, Z. Yong, W. Wanjun, H. Yinlei, Z. Qiang, Y. Jianyi, W. Minghua, and J. Xiaoqing. Analysis of the thermo-optic effect in lateral-carrier-injection soi ridge waveguide devices. *Journal of Semiconductors*, 31:064009, 2010.
- [34] A. Densmore, S. Janz, R. Ma, J. H. Schmid, D.-X. Xu, A. Delâge, J. Lapointe, M. Vachon, and P. Cheben. Compact and low power thermo-optic switch using folded silicon waveguides. *Opt. Express*, 17:10457–10465, 2009.

- [35] Y. Shoji, K. Kintaka, S. Suda, H. Kawashima, T. Hasama, and H. Ishikawa. Low-crosstalk 2x2 thermo-optic switch with silicon wire waveguides. *Opt. Express*, 18:9071–9075, 2010.
- [36] H. M. H. Chong and R. M. de la Rue. Tuning of photonic crystal waveguide microcavity by thermo-optic effect. *Photonics Technology Letters, IEEE*, 16:1528–1530, 2004.
- [37] V. Yurii and S. McNab. Losses in single-mode silicon-on-insulator strip waveguides and bends. *Opt. Express*, 12:1622–1631, 2004.
- [38] M. Harb, R. Ernstorfer, T. Dartigalongue, C. T. Hebeisen, R. E. Jordan, and R. J. D. Miller. Carrier relaxation and lattice heating dynamics in silicon revealed by femtosecond electron diffraction. *The Journal of Physical Chemistry B*, 110:25308–25313, 2006.
- [39] J. R. Goldman and J. A. Prybyla. Ultrafast dynamics of laser-excited electron distributions in silicon. *Phys. Rev. Lett.*, 72:1364–1367, 1994.
- [40] H. M. van Driel. Influence of hot phonons on energy relaxation of high-density carriers in germanium. *Phys. Rev. B*, 19:5928–5932, 1979.
- [41] G. K. Celler and S. Cristoloveanu. Frontiers in silicon-on-insulator. *J. Appl. Phys.*, 93:4955, 2002.
- [42] F. Magno, F. Dell’Olio, and V. M. N. Passaro. Multiphysics investigation of thermo-optic effect in silicon-on-insulator waveguide arrays. *Proceedings of the COMSOL Users Conference*, pages 1–6, Milano, Italy, 2006.

Chapter 5

Third-Harmonic Generation in Silicon-on-Insulator Nanophotonic Structures

5.1 Introduction

The previous chapter focused on nonlinear optical effects that occur at the frequency of the input signal (i.e. $\chi^{(3)}(-\omega; \omega, -\omega, \omega)$). It was shown that femtosecond pulses with a broad spectrum could be used to generate new frequency components in silicon-on-insulator (SOI) waveguides via self-phase modulation (SPM). This interaction is described by the nonlinear susceptibility coefficient, $\chi^{(3)}(-\omega; \omega_1, -\omega_2, \omega_3)$, where the interaction of frequencies at ω_1, ω_2 , and ω_3 give rise to a new frequency at $\omega = \omega_1 - \omega_2 + \omega_3$ [1]. Using either SPM, cross-phase modulation (XPM), or four-wave mixing, it becomes possible to convert telecommunication wavelengths to a wide band of near-infrared frequencies [2–15]. Conversion of telecommunication wavelengths to visible light in silicon would require third-harmonic generation (THG) via the $\chi^{(3)}(-3\omega; \omega, \omega, \omega)$ nonlinear susceptibility.

Harmonic generation marked the beginning of nonlinear optics and it remains a powerful mechanism for converting coherent optical radiation to higher frequencies [16]. Near-infrared (NIR) wavelengths form the backbone of present day fibre optical networks and integrated photonic systems. As a critical material in the electronics industry, silicon has also become the foundation for integrated optical systems as it merges photonics with electronics onto a single platform. While a wide range of nonlinear optical processes have been demonstrated in silicon for various applications, direct light emission remains challenging due to the indirect bandgap of silicon.

Efficient light emission by THG in silicon would facilitate on-chip conversion of NIR radiation to visible wavelengths and would enable the realization of nanoscale coherent light sources for visible wavelength all-optical signal processing, optoelectronic clock synchronization, and quantum optics on a chip. For a typical bulk semiconductor material,

the relatively small magnitude of the third-order optical susceptibility, $\chi^{(3)}$, in comparison to the second-order optical susceptibility, $\chi^{(2)}$, of exotic nonlinear materials (e.g. LiNbO₃ and LiTaO₃) has made a cascaded two-step process consisting of sum frequency generation between the second-harmonic (2ω) and the fundamental (ω) more efficient than direct THG [17]. However, inversion symmetry in Si makes cascaded THG impossible in SOI waveguides and single-step THG is necessary.

While bulk silicon shows relatively weak $\chi^{(3)}$ nonlinearities compared to exotic $\chi^{(2)}$ nonlinear materials, the extreme mode confinement provided by silicon dispersion engineered waveguides increases the optical power density and enhances nonlinear interactions in the waveguide core, allowing for a single-step THG process to take place with a moderate efficiency [18]. Despite such promise, THG has proven challenging to observe in SOI waveguides and only recently was there a report of THG in a SOI dispersion optimized photonic crystal (PhC) waveguide [18]. A combination of tight modal confinement in the PhC waveguide and spatial pulse compression, due to slow light propagation, allowed for narrowband green ($517 \leq \lambda \leq 520\text{nm}$) light emission with a conversion efficiency of $\eta_{THG} \approx 10^{-7}$ (or 5×10^{-10} for 1W of peak pump power). In this case, the overall PhC waveguide dimensions were large and required intricate designs for dispersion compensation, slow light mode energy density enhancement, and visible light extraction through the Si PhC lattice. Moreover, the green light was strongly diffracted by the PhC and emitted into free-space, limiting its functionality for on-chip light guiding applications.

In the interest of developing visible light sources that are suitable for monolithic integration with complementary metal-oxide-semiconductor (CMOS) electronics and silicon photonics, it would be desirable to develop devices that have an ultracompact footprint, operate at a high third-harmonic (TH) conversion efficiency, convert a broad bandwidth of telecommunication wavelengths to their third-harmonics, and confine the visible light to the integrated optical circuit, so that it can be coupled to waveguides that are transparent at visible wavelengths (e.g. SiO₂ or SiN).

5.2 The Influence of Waveguide Mode-Confinement on Third-Harmonic Generation

5.2.1 Phase-matching Considerations

The prospect of achieving phase-matched THG in SOI waveguides was investigated by performing mode-solver simulations. Quasi-transverse-magnetic (quasi-TM) mode profile intensity distributions for a SOI waveguide with $w = h = 340\text{nm}$ are shown at $\lambda = 1550\text{nm}$ in Fig. 5.1(a) and at $\lambda = 517\text{nm}$ in Fig. 5.1(b). Approximately 56.2% of the mode energy is confined to the silicon waveguide core at $\lambda = 1550\text{nm}$, whereas virtually all of the mode energy is confined to the silicon core at $\lambda = 517\text{nm}$. The overlap between the $\lambda = 1550\text{nm}$

and $\lambda = 517\text{nm}$ modes is calculated via:

$$f(A_\omega, A_{3\omega}) = \frac{|\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_\omega E_{3\omega} dx dy|^2}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |E_\omega|^2 dx dy \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |E_{3\omega}|^2 dx dy} \quad (5.1)$$

and is found to be 0.54. The effective refractive index of the $\lambda = 1550\text{nm}$ mode is found to be $n_{eff}^\omega = 2.24$, and therefore the effective wavelength is $\lambda_{eff}^\omega = 692\text{nm}$. Similarly, at $\lambda = 517\text{nm}$ the effective refractive index is found to be $n_{eff}^{3\omega} = 4.09$, corresponding to $\lambda_{eff}^{3\omega} = 126\text{nm}$.

Phase-matched THG is achieved when the wave vector mismatch, i.e. $\Delta k = |k_{3\omega} - 3k_\omega|$, is minimized. For these modes, the wave vector mismatch is calculated to be $\Delta k = 2.25 \times 10^7 \text{rad/s}$. Therefore, phase-matched THG is not possible using the lowest-order quasi-TM modes. Curves depicting the dispersion of the quasi-transverse-electric (quasi-TE) and quasi-TM modes for the fundamental frequency excitation in the range of $1500\text{nm} \leq \lambda \leq 1600\text{nm}$ and the 3ω quasi-TE and quasi-TM modes in the range of $500\text{nm} \leq \lambda \leq 533\text{nm}$ are calculated using a mode solver and are shown in Fig. 5.2. It should be noted that due to the strong confinement of the 3ω modes to the silicon core, the effective refractive indices of the quasi-TE and quasi-TM 3ω modes were approximately equal and the curves completely overlapped. Clearly, there is no spectral region where the lowest-order modes overlap.

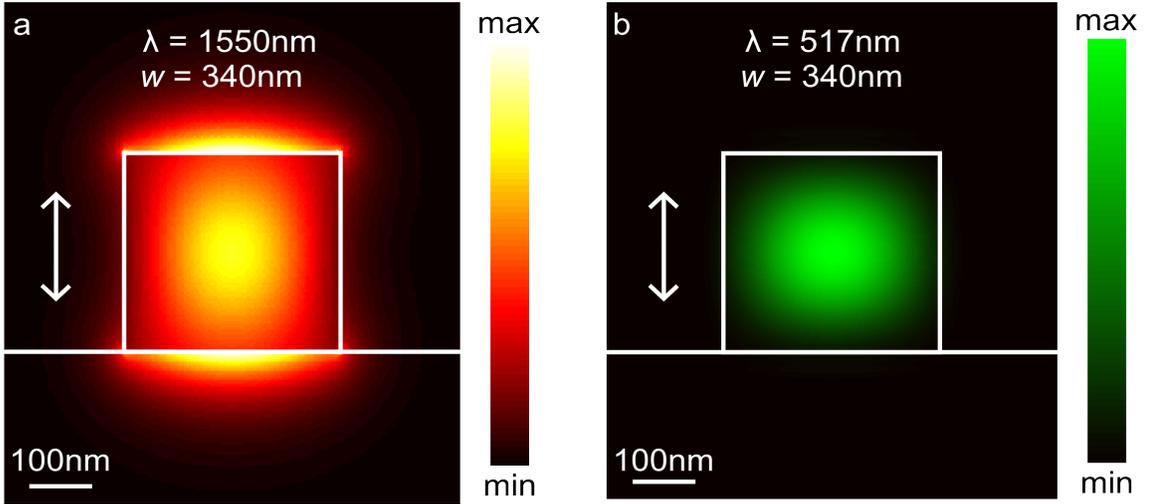


Figure 5.1: Transverse-magnetic mode intensity (I) distributions for a SOI waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$ at (a) $\lambda = 1550\text{nm}$; and (b) $\lambda = 517\text{nm}$.

Previous demonstrations of second-harmonic generation in non-SOI waveguides have achieved phase-matching with higher-order modes of the second-harmonic [19]. In general, as the mode order is increased, the effective refractive index decreases, since higher-order modes cannot be confined as well to the silicon waveguide core. Subsequent analysis of the waveguide modes reveals that a higher-order quasi-TM mode has an effective refractive index of approximately $n_{eff}^{3\omega} = 2.27$ at $\lambda^{3\omega} = 511.7\text{nm}$, which matches the effective

refractive index at $\lambda^\omega = 3\lambda^{3\omega} = 1535\text{nm}$. The mode dispersion characteristics of this higher-order mode are also shown in Fig. 5.2, and the phase-matching condition is satisfied where the lowest-order fundamental quasi-TM and higher-order 3ω curves intersect at $\lambda^\omega = 1535\text{nm}$ or $\lambda^{3\omega} = 511.7\text{nm}$.

Another important consideration is the propagation loss of these two modes. In the absence of sidewall roughness, the $\lambda = 1535\text{nm}$ mode propagates with virtually no loss. However, the $\lambda = 511.7\text{nm}$ mode undergoes extremely high propagation losses due to the inherent linear absorption of silicon at this wavelength. Through the mode-solver simulations, the propagation losses experienced by $\lambda = 511.3\text{nm}$ radiation are determined to be, $Loss \approx 6.4\text{dB}/\mu\text{m}$, corresponding to a propagation length of $L_{prop}^{3\omega} = 679\text{nm}$. Due to this high loss, phase matching over the entire length of an SOI waveguide is not attainable. Since the 3ω decays over essentially one wavelength of the fundamental, i.e. $L_{prop}^{3\omega} \approx \lambda_{eff}^\omega$, phase-matching is insignificant over a length scale of hundreds of microns, and any observed THG must arise from an essentially localized interaction. As such, any 3ω signal present at a position (x, y, z) within the silicon waveguide core can only be generated within its propagation length ($\delta = 679\text{nm}$) along the waveguide, (i.e. at $(x \pm \delta, y \pm \delta, z \pm \delta)$). Based on the propagation length of the 3ω , $L_{prop}^{3\omega} = 679\text{nm}$, it is determined that approximately 77% of the 3ω power measured from output facet of an SOI waveguide would be generated in the last $1\mu\text{m}$ portion waveguide. For essentially the same reasons, THG in bulk silicon samples is typically observed from surface reflection, where the 3ω signal is generated within the absorption depth of both ω and 3ω .

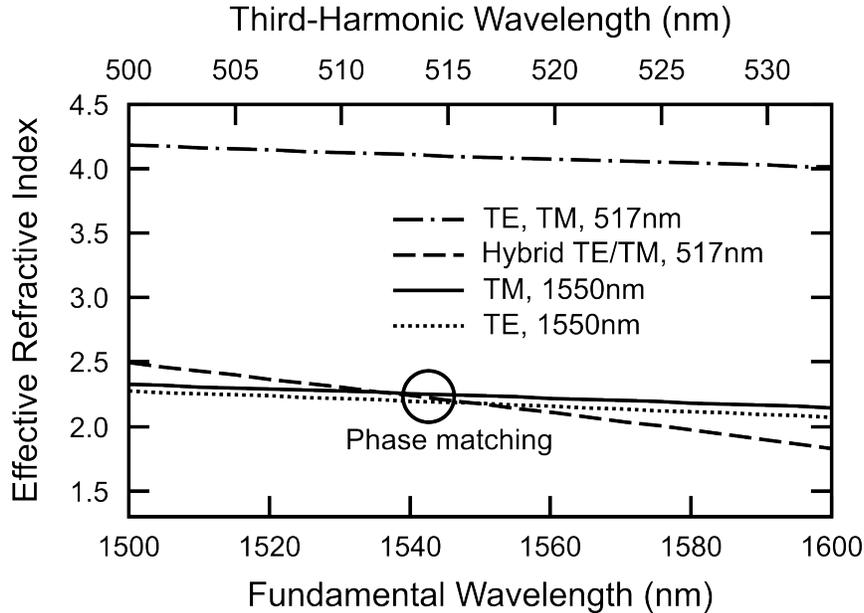


Figure 5.2: Dispersion of modes in a $w = 260\text{nm}$, $h = 340\text{nm}$ SOI waveguide. Near-infrared TE (dotted line) and TM (solid line) waveguide modes have significantly different effective refractive indices, whereas the TE and TM modes (dashed line) at green wavelengths ($500\text{nm} \leq \lambda \leq 535\text{nm}$) have virtually identical refractive indices. Effective refractive indices were calculated with a mode solver.

Finite-difference time-domain (FDTD) simulations are performed on a SOI waveguide with $w = h = 340\text{nm}$ to illustrate the propagation characteristics of the fundamental and 3ω frequencies. A time snapshot of the y-component of the electric field distribution of a quasi-TM mode at $\lambda = 1550\text{nm}$, taken in the xz -plane (as depicted in Fig. 2.15) is shown in Fig. 5.3(a). The same simulations are performed at $\lambda = 517\text{nm}$, producing the snapshot shown in Fig. 5.3(b). Clearly, the 3ω signal dies off within several oscillations of the fundamental mode.

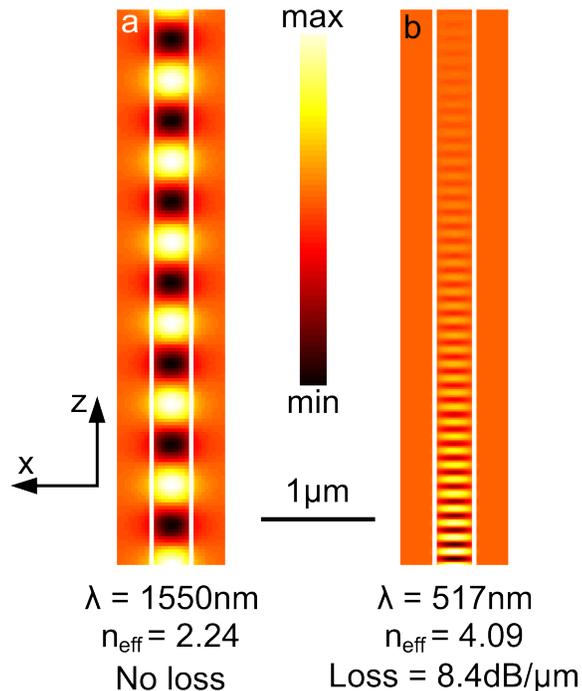


Figure 5.3: Electric field (E_y) distribution of quasi-TM modes in a $w = 340\text{nm}$, $h = 340\text{nm}$ SOI waveguide at (a) $\lambda = 1550\text{nm}$ and (b) $\lambda = 517\text{nm}$, recorded in the xz -plane, as depicted in Fig. 2.15.

5.2.2 Observation of Third-Harmonic Generation

The optimum mode confinement conditions for high-efficiency THG are determined by investigating SOI waveguides with a fixed length, $L = 200\mu\text{m}$, a fixed height, $h = 340\text{nm}$, and different widths, $w = \{260, 340, 470, 540, 760, 1000, 1510, 2050, 4150\}$ nm. A schematic depiction of the SOI waveguide with its dimensions and axes is shown in Fig. 2.15. For waveguides with $w = \{260, 340, 470, 540\}$ nm, only a single quasi-TM and a single quasi-TE mode are supported by the waveguide at $\lambda = 1550\text{nm}$. Therefore, by controlling the polarization of excitation radiation, a single mode can be excited. However, the larger waveguides with $w = \{760, 1000, 1510, 2050, 4150\}$ nm, support multiple modes that can be excited simultaneously, dividing the energy of the pulse between various modes. The effective refractive index of the different modes can vary substantially, and different modes excited by a femtosecond pulse can temporally walk-off from one

another, resulting in a weaker nonlinear interaction.

Third-harmonic generation was excited in the SOI waveguides using a $\lambda = 1550\text{nm}$ laser emitting $\tau_p = 84\text{fs}$ pulses at a repetition rate of 90MHz . Radiation was coupled to the waveguides via end-fire excitation with a $NA = 0.65$ microscope objective, where the polarization of the laser was oriented to excite quasi-TM waveguide modes. Output radiation was coupled into a lensed single-mode optical fibre with a focal spot diameter, $d = 2.5\mu\text{m}$. The optical fibre was used to deliver the output fundamental and 3ω radiation to power meters and spectrometers. Since 77% of the 3ω power is generated in the last $1\mu\text{m}$ portion of the waveguide, and because the absorption (both linear and nonlinear) of the $\lambda = 1550\text{nm}$ pulses is negligible over this length scale, this enables an accurate measure of both the 3ω power and the $\lambda = 1550\text{nm}$ power near the output facet of the waveguide, which is approximately the power that was used to excite the measured 3ω .

For peak coupled powers exceeding $P_{peak}^\omega = 37\text{W}$, a strong 3ω signal is observed. A top-view microscope image of green light emission from a $w = 260\text{nm}$, $L = 200\mu\text{m}$ waveguide is shown in Fig. 5.6(a). Spatial interference patterns with a period, $\Lambda = 3.05 \pm 0.85\mu\text{m}$, are a prominent feature of the emission and are only visible for the $w = 260\text{nm}$ waveguides. However, they are visible and have the same period for $w = 260\text{nm}$ waveguides with lengths ranging between $200\mu\text{m}$ and $1120\mu\text{m}$. Due to the small cross-sectional area of the silicon core, this waveguide supports modes with a weak light confinement within the silicon core. As such, a large fraction of the nonlinear interaction takes place near the waveguide sidewalls and near the top interface. Therefore, some of the 3ω emission is scattered due to fabrication-induced roughness and subsequently observed with the microscope camera. The emitted 3ω light intensity decays with a characteristic length, $L_{decay} = 29.5\mu\text{m}$, and after $\approx 60\mu\text{m}$, the signal is much weaker and cannot be detected by the microscope camera. The distinctive periodicity of the interference pattern suggests the presence of a second, leaky mode propagating within the waveguide. Further analysis of the waveguide modes reveals that a poorly confined mode with $n_{eff} = 0.967$ (at $\lambda = 1550\text{nm}$) is supported by this waveguide, which would beat with the primary TM mode with a period, $\Lambda = \lambda/\Delta n_{eff} = 2.78\mu\text{m}$ (i.e. not accounting for TPA plasma-induced refractive index change). The mode profile for the quasi-TM mode is shown in Fig. 5.4(a) and the mode profile for the poorly-confined leaky mode is shown in Fig. 5.4(b).

First, the propagation of each of these two modes is simulated using FDTD simulations. In the first simulation, only the quasi-TM mode is excited in the waveguide by directly injecting the mode into the cross-section of the waveguide, and the time-averaged intensity distribution is recorded in the $z = 170\text{nm}$ plane. Then the intensity is cubed to obtain the expected THG distribution, which is shown in Fig. 5.5(a). Subsequently, this same procedure is performed for the poorly-confined, leaky mode, which is depicted in Fig. 5.5(b). The electromagnetic energy near the mode input decays sharply, but subsequent analysis reveals that finite mode energy remains throughout the remainder of the waveguide. Finally, both modes are injected into the waveguide simultaneously, pro-

ducing the third-harmonic distribution in Fig. 5.5(c). Notably, the periodic interference pattern in the third-harmonic emission is only visible when both modes are present.

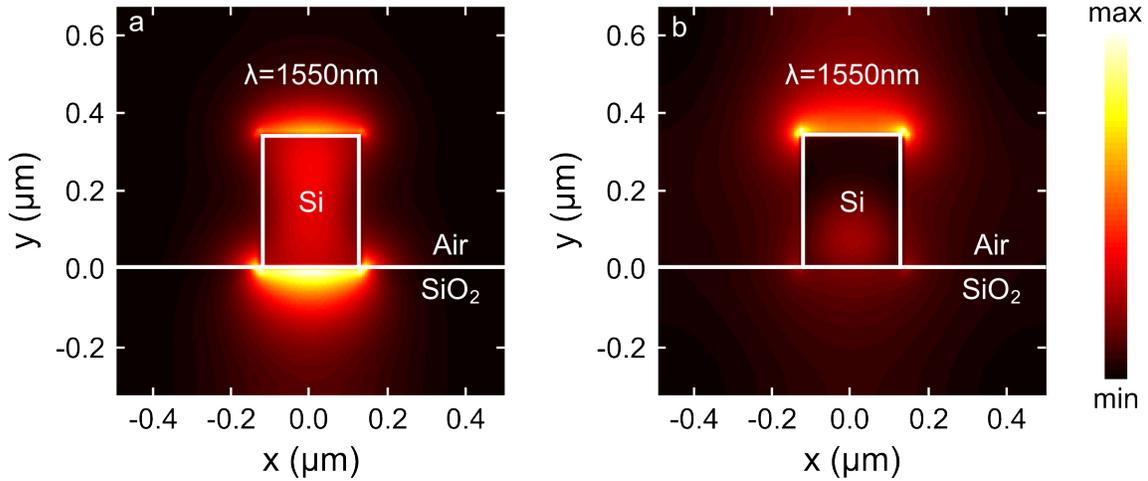


Figure 5.4: Mode profile intensity (I) distributions at $\lambda = 1550\text{nm}$ for: (a) the quasi-TM mode; and (b) the poorly-confined leaky mode.

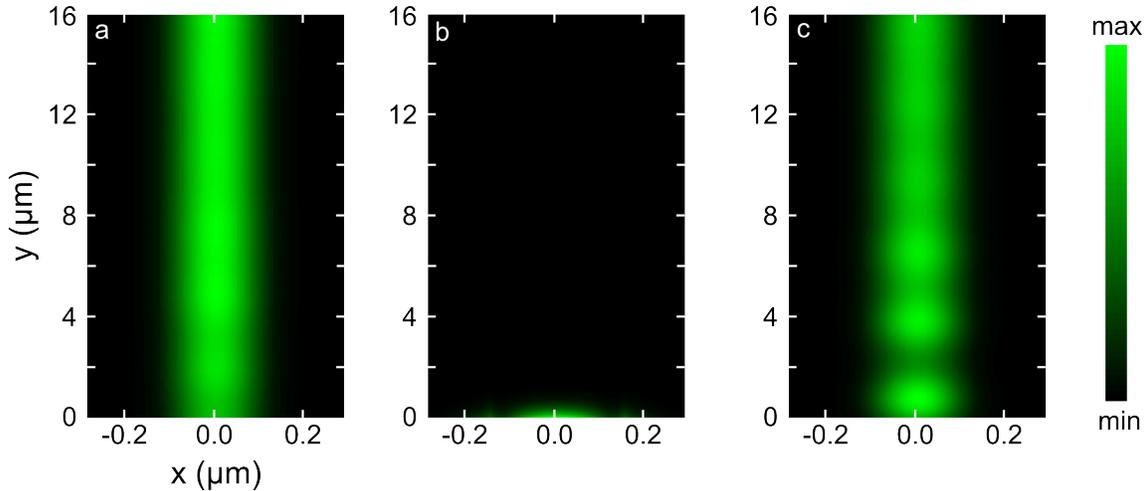


Figure 5.5: Calculated third-harmonic intensity (I) distribution for: (a) the quasi-TM mode; (b) the poorly-confined, leaky mode; and (c) both modes excited simultaneously.

Next, the excitation of the waveguide with a $NA = 0.65$ microscope objective is investigated by decomposing the refracted Gaussian laser beam into a sum of plane waves all propagating at different angles and converging towards the beam centre at the focal plane [20]. Each plane wave is proportional to $\exp(i\mathbf{k} \cdot \mathbf{r})$, where the \mathbf{k} -vector is determined by the angle of propagation. The focused beam leaving the lens is constructed by integrating all of these plane waves in the plane of the lens. The calculated fields are propagated towards the waveguide input facet using the FDTD method, and the fields propagating through the waveguide are recorded. The 3ω excitation is proportional to I^3 , which is shown in Fig. 5.6(b). A similar periodic pattern with $\Lambda = 2.62 \pm 0.26\mu\text{m}$ is visible, confirming that both the quasi-TM and the poorly-confined leaky mode are excited due

to the high-NA focusing and confirming the origin of the observed spatially-modulated green emission. A rendering of the SOI waveguide along with a mesh constructed from the experimentally measured spatial distribution of THG is shown in Fig. 5.6(c). The green light emission is collected by aligning a lensed optical fibre to the output facet of the SOI waveguide. The green light spectrum is then delivered to a spectrometer, and an exemplary trace measured at $P_{peak}^\omega = 100\text{W}$ is shown in Fig. 5.6(d).

5.2.3 Mode Confinement Dependence of Third-Harmonic Generation

Intensity maps of the 3ω power, $P_{out}^{3\omega}$, as a function of the 3ω wavelength and P_{peak}^ω are shown for each of the nine SOI waveguides, in Fig. 5.7. A superimposed spectral slice of the intensity map taken at a peak power, $P_{peak}^\omega = 100\text{W}$ provides further details of the 3ω signal spectral components. Figures 5.7(a)-(e) depict results from the multimode SOI waveguides with $w = \{4150, 2050, 1510, 1000, 760\}$ nm, respectively. The peak wavelength of the THG spectrum at the threshold input power ($P_{peak}^\omega = 37\text{W}$) is found to be at $\lambda = \{516.1, 516, 516.7, 516.9, 516.9\}$ nm for $w = \{4150, 2050, 1510, 1000, 760\}$ nm, respectively, which is virtually identical to the expected 3ω peak at $\lambda = 1550\text{nm}/3 = 516.7\text{nm}$. These multi-mode waveguides generate a narrow 3ω spectrum, with a full-width at half-maximum (FWHM) of $5.5\text{nm} \leq FWHM \leq 9.6\text{nm}$ and a spectral peak 3ω power, $P_{max}^{3\omega}$, in the range of $2.53\text{pW} \leq P_{max}^{3\omega} \leq 5.00\text{pW}$, when excited with $P_{peak}^\omega = 100\text{W}$.

The $w = \{260, 340, 470, 540\}$ nm waveguides generate a spectral peak 3ω power in the range of $0.84\text{pW} \leq P_{max}^{3\omega} \leq 2.98\text{pW}$, when excited with $P_{peak}^\omega = 100\text{W}$. Interestingly, their 3ω spectrum is different and spectral reshaping with increasing excitation power is evident, as shown in Figs. 5.7(f)-(i). The single-mode waveguides exhibit a much broader 3ω spectrum that is blue-shifted from the expected 3ω wavelength. The FWHM increases for smaller multi-mode waveguides up to $FWHM = 19\text{nm}$ for the smallest $w = 260\text{nm}$ waveguide. The peak wavelength and FWHM are calculated for each of the 3ω spectra measured at a representative $P_{peak}^\omega = 100\text{W}$, and are plotted as a function of w in Fig. 5.8.

5.2.4 Third-Harmonic Power Scaling

The $\chi^{(3)}(3\omega; \omega, \omega, \omega)$ origin of the green light emission was verified by integrating the total counts measured by the spectrometer in the wavelength range $490\text{nm} \leq \lambda \leq 535\text{nm}$. The measured green light counts were plotted as a function of the excitation power at $\lambda = 1550\text{nm}$ on logarithmic x - and y - axes. The most accurate measure of the $\lambda = 1550\text{nm}$ power used to excite the measured 3ω power is determined from the measured $\lambda = 1550\text{nm}$ signal from the output facet of the waveguide. The logarithmic plot of green light power versus output $\lambda = 1550\text{nm}$ power, shown as triangles in Fig. 5.9, follows a

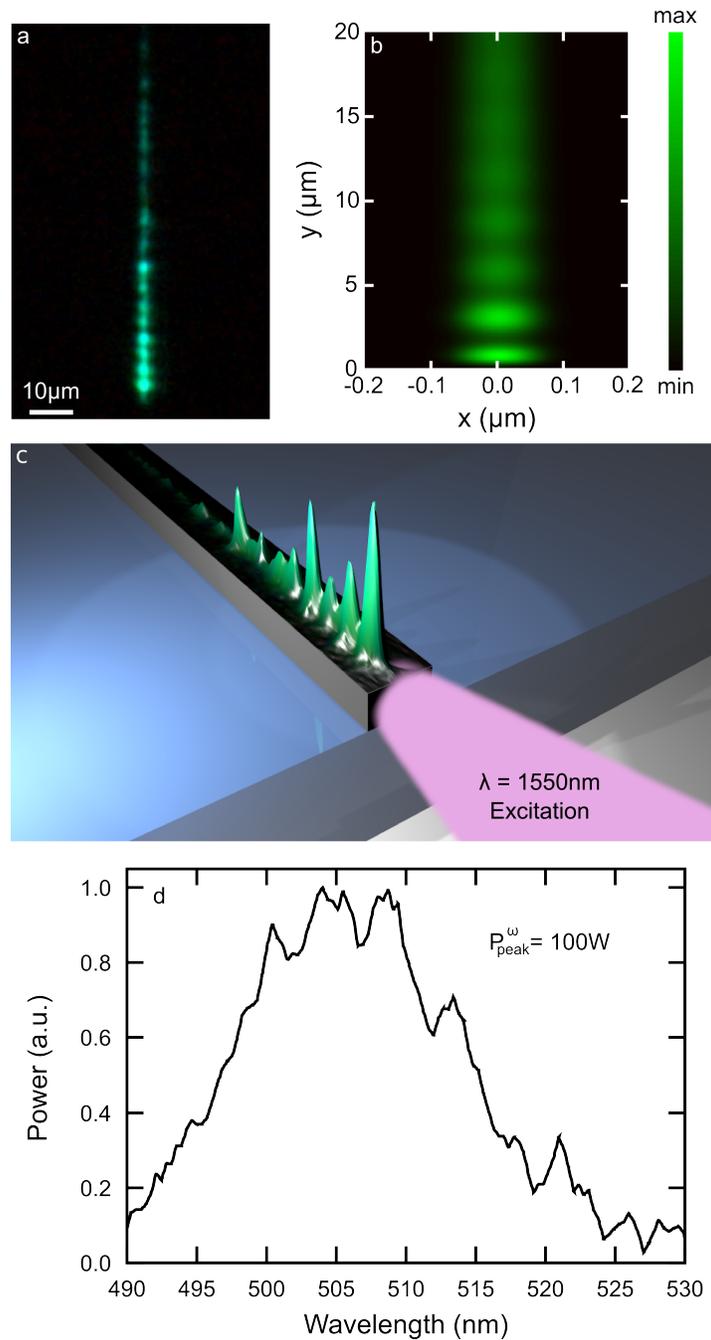


Figure 5.6: (a) Camera image of the interference pattern present in the waveguide taken at $20\times$ with a colour charge-coupled device camera. Mode beating with a period of $\Lambda = 3.05 \pm 0.85 \mu\text{m}$ is observed. (b) Third-harmonic intensity (I) interference pattern with $\Lambda = 2.62 \pm 0.26 \mu\text{m}$, calculated via FDTD simulations. (c) Artistic rendering of the sample under test along with the scattered green light that is captured with a microscope camera. The mesh plot contains the experimental data collected from the microscope camera. (d) Measured spectrum from the output facet of the SOI waveguide at $P_{peak}^{\omega} = 100W$.

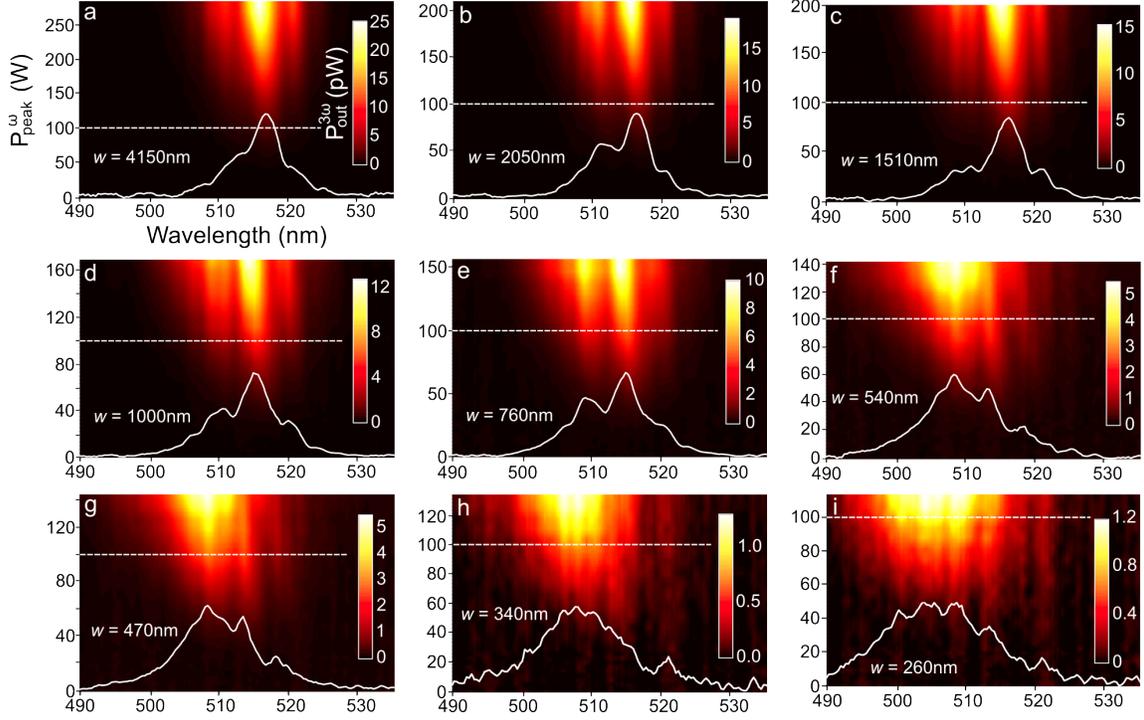


Figure 5.7: Intensity surfaces depicting the THG spectrum growth with increasing peak input power, P_{peak}^ω , for waveguides with $w =$ (a) 4150nm, (b) 2050nm, (c) 1510nm, (d) 1000nm, (e) 760nm, (f) 540nm, (g) 470nm, (h) 340nm, and (i) 260nm. The white line overlay portrays the third-harmonic spectrum measured at the dashed section, corresponding to $P_{peak}^\omega = 100\text{W}$.

linear trend, and a linear fit to this data (shown by the solid line) produces a slope of 3.02 ± 0.10 , confirming that the measured green light emission arises from THG. The output 3ω power is also plotted as a function of the excitation power (i.e. the power coupled to the input facet of the waveguide), which is shown by the circles in Fig. 5.9. A linear fit to this data returns a slope of 2.20 ± 0.05 , which is less than the ideal value of 3, revealing that there is loss in the system. In general, this loss includes both linear and nonlinear components. Linear losses due to sidewall roughness, two-photon absorption (TPA) and free-carrier absorption (FCA) attenuate the $\lambda = 1550\text{nm}$ along the length of the waveguide and effectively decrease the $\lambda = 1550\text{nm}$ power at the output facet that is available for the THG process, reducing the overall efficiency. Pump attenuation due to TPA accumulates along the entire propagation distance, z , of the waveguide according to [1]:

$$\frac{dI}{dz} = -\beta_{TPA}I^2, \quad (5.2)$$

where $\beta_{TPA} = 0.68\text{cm/GW}$ is the TPA coefficient and I is the local $\lambda = 1550\text{nm}$ intensity. Additionally, the resulting free-carrier population can further attenuate the $\lambda = 1550\text{nm}$ pulses via FCA:

$$\frac{dI}{dz} = -\sigma_{FCA}NI, \quad (5.3)$$

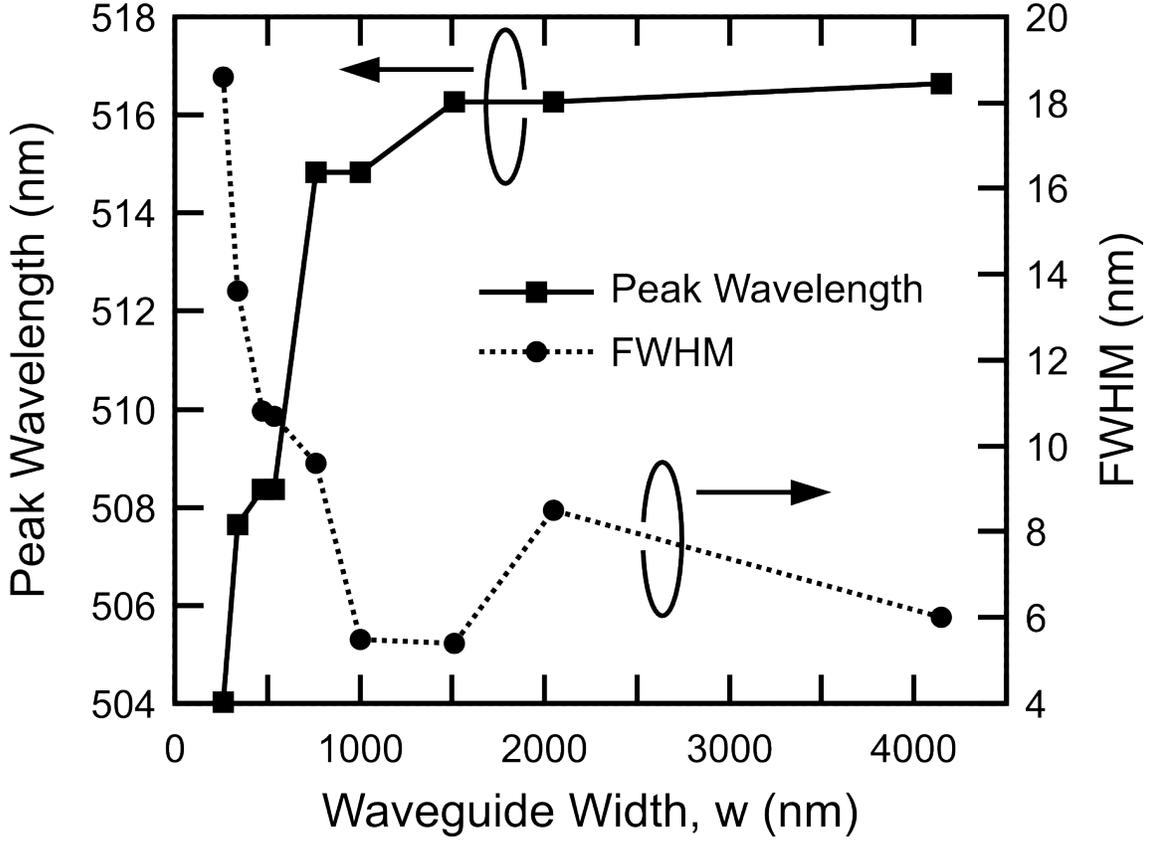


Figure 5.8: Plots of third-harmonic spectrum peak wavelength (squares and solid line) and FWHM (circles and dotted line) as a function of waveguide width, w , for a fixed $P_{peak}^\omega = 100W$.

where $\sigma_{FCA} = 1.45 \times 10^{-17} \text{cm}^2$ is the FCA cross-section and N is the free-carrier density, which is calculated to be $N \approx 5 \times 10^{19} \text{cm}^{-3}$ for $P_{peak}^\omega = 55W$ in the $w = 260 \text{nm}$ waveguide. While FCA will absorb some of the 3ω radiation, it is expected that the intrinsic losses of silicon are much higher than the FCA-induced losses to the 3ω signal. Reducing the length of the waveguide to the order of several microns would circumvent these losses.

5.2.5 Discussion of Third-Harmonic Spectral Reshaping

Further physical insight into the spectral reshaping of the THG for increasing P_{peak}^ω can be gained by normalizing each spectral slice of the intensity maps from Fig. 5.7. In this way, each spectral slice has a peak amplitude of unity, and it becomes easier to observe the evolution of the 3ω spectral intensity distribution as the input peak power is increased, as shown in Fig. 5.10. For the large multi-mode waveguides with $w = \{4150, 2050, 1510\} \text{nm}$, the peak of the 3ω spectrum exhibits a discernible blue-shift from $\lambda_1 = \{516.9, 516.9, 516.7\} \text{nm}$ to $\lambda_2 = \{516.5, 515.8, 515.5\} \text{nm}$ as the peak input power is increased from $P_{peak}^\omega = 37W$ to $P_{peak}^\omega = 150W$, while the FWHM remains relatively constant as the peak power is increased. For the smaller two multi-mode waveguides with $w = \{1000, 760\} \text{nm}$, a similar blue-shift in the peak wavelength of the 3ω spectrum is

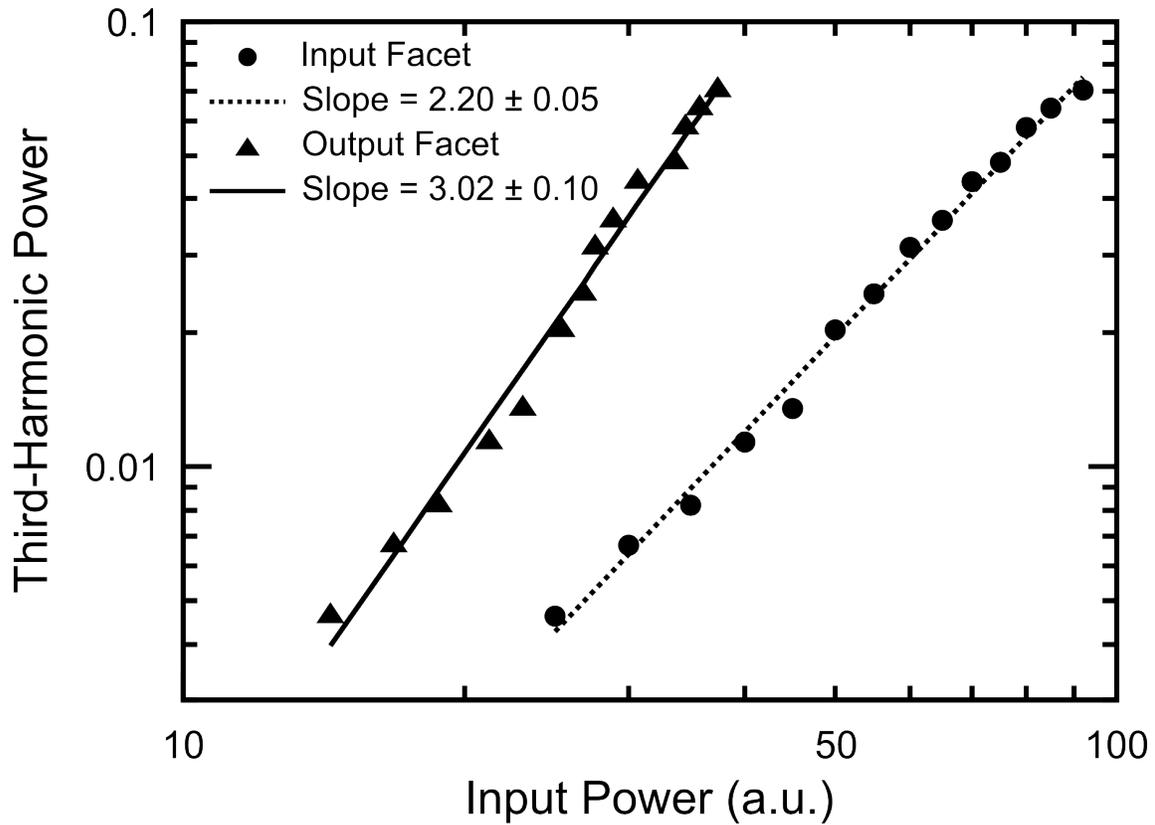


Figure 5.9: Plots depicting the measured third-harmonic power as a function of the excitation $\lambda = 1550\text{nm}$ power, plotted on logarithmic x - and y - axes. Two data sets are plotted. The first represents the $\lambda = 1550\text{nm}$ power coupled to the input facet of the waveguide (circles), which scales with a slope of 2.20 ± 0.05 . The second represents the $\lambda = 1550\text{nm}$ power at the output facet of the waveguide, which is the most accurate measure of the power used to excite the THG, and scales with a slope of 3.02 ± 0.10 .

observed. Meanwhile, as P_{peak}^ω is increased, a secondary peak appears at $\lambda = 510\text{nm}$ which grows in amplitude with increasing P_{peak}^ω . For the $w = 760\text{nm}$ waveguide, the power of the peak at $\lambda = 510\text{nm}$ increases to 86.3% of the original peak power at $\lambda = 515.6\text{nm}$ for an input peak power, $P_{peak}^\omega = 150\text{W}$.

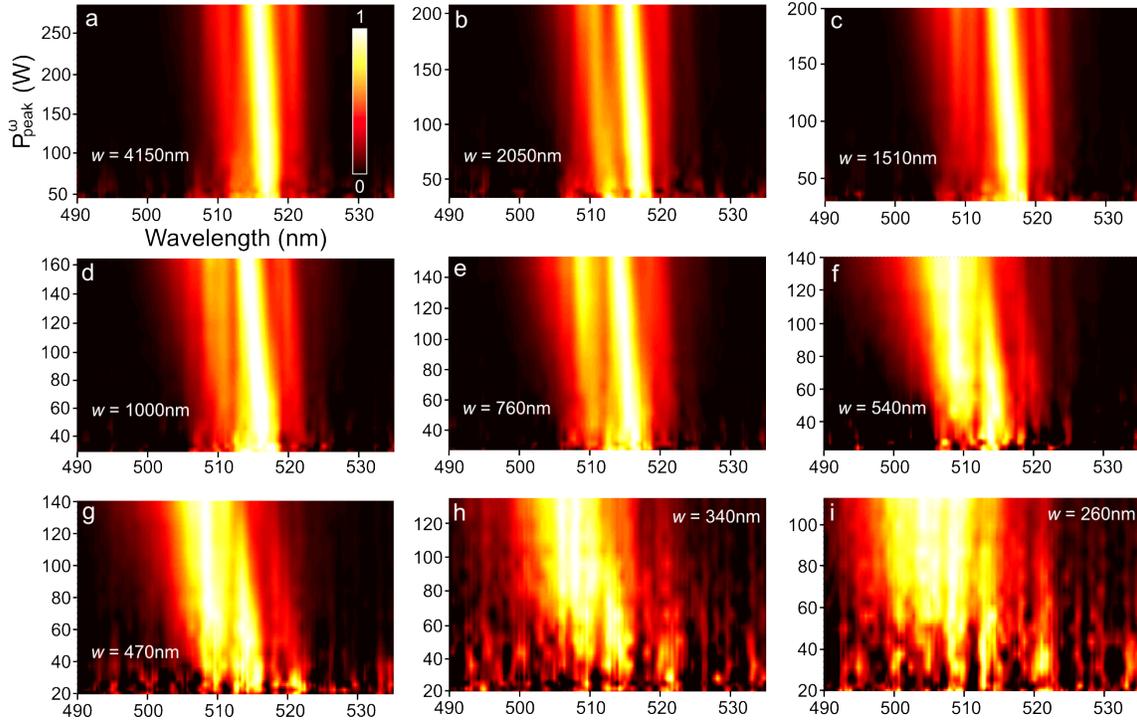


Figure 5.10: Intensity surfaces depicting the normalized third-harmonic spectrum with increasing peak input power for waveguides with $w =$ (a) 4150nm, (b) 2050nm, (c) 1510nm, (d) 1000nm, (e) 760nm, (f) 540nm, (g) 470nm, (h) 340nm, and (i) 260nm.

This trend continues for single-mode SOI waveguides with smaller $w = \{540, 470\}$ nm, as shown in Fig. 5.104(f)-(g), respectively. In these SOI waveguides, at the threshold power, $P_{peak}^\omega = 37\text{W}$, two 3ω peaks are clearly visible at wavelengths of $\lambda = 509\text{nm}$ and $\lambda = 514\text{nm}$. As P_{peak}^ω is increased, a blue-shift is observed in the peak at $\lambda = 514\text{nm}$ and its amplitude relative to the peak at $\lambda = 509\text{nm}$ decreases. Nonetheless, the wavelength of the peak at $\lambda = 509\text{nm}$ remains relatively unchanged as P_{peak}^ω is increased. At $P_{peak}^\omega = 140\text{W}$, virtually all of the 3ω energy is observed from a broad peak at $\lambda = 509\text{nm}$. In the case of the single-mode waveguide with a width, $w = 340\text{nm}$, shown in Fig. 5.10(h), the peak at $\lambda = 514\text{nm}$ is barely discernible, even at low input peak powers near the threshold for observing THG. However, as the peak power is increased, only a single, broad peak at $\lambda = 507.6\text{nm}$ with $FWHM = 16.6\text{nm}$ is observed. Interestingly, as shown in Fig. 5.10(i), THG from the smallest waveguide with $w = 260\text{nm}$ exhibits a different behaviour. Regardless of the P_{peak}^ω , a broad spectrum centred at $\lambda = 505.5\text{nm}$ with $FWHM \approx 18.7\text{nm}$ is observed with no indication of any blue-shift as P_{peak}^ω is increased.

With the exception of the $w = 260\text{nm}$ SOI waveguide, there is a distinct blue-shift in the 3ω spectrum as the peak input power is increased, up to $\Delta\lambda \approx 13\text{nm}$. The strong

blue-shift and spectral broadening are evidence of a rapidly decreasing effective refractive index of the waveguide mode, and demonstrate that multiple strong nonlinear interactions (i.e. THG, TPA, FCA, SPM) develop in the waveguides. A majority of the 3ω radiation is generated near the peak of the NIR pulses, and promptly experiences two conditions that contribute to this large blue-shift. Firstly, the 3ω radiation is simultaneously generated with a large TPA-excited free-carrier density, $N = 10^{18} - 10^{20}\text{cm}^{-3}$, near the peak of the $\lambda = 1550\text{nm}$ pulses. The presence of these free-carriers gives rise to a change in the complex refractive index, $\tilde{n} = n_r + in_i$, according to:

$$\Delta n_r = \frac{e^2 \lambda^2 N}{8\pi^2 c^2 \varepsilon_0 n_r} \left[\frac{1}{m_e^*} + \frac{1}{m_h^*} \right] \quad (5.4)$$

$$\Delta n_i = \frac{e^3 \lambda^3 N}{16\pi^3 c^3 \varepsilon_0 n_r} \left[\frac{1}{(m_e^*)^2 \mu_e} + \frac{1}{(m_h^*)^2 \mu_h} \right], \quad (5.5)$$

where e is the elementary charge, c is the speed of light, ε_0 is the permittivity of free space, n_r and n_i are the real and imaginary components of the refractive index of silicon in the absence of free-carriers, respectively, $m_e^* = 2.37 \times 10^{-31}\text{kg}$ and $m_h^* = 3.55 \times 10^{-31}\text{kg}$, are the effective masses of electrons and holes, respectively, and $\mu_e = 1400\text{cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$ and $\mu_h = 450\text{cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$ are the mobilities of electrons and holes, respectively, for silicon. For $N_e = N_h = 5 \times 10^{19}\text{cm}^{-3}$ (corresponding to $P_{peak}^\omega = 55\text{W}$ in the $w = 260\text{nm}$ waveguide), a value of $\Delta\tilde{n} = -9.10 \times 10^{-3} + i1.73 \times 10^{-5}$ is calculated at $\lambda = 517\text{nm}$. Such a large change in Δn_r over a sub-100fs timescale gives rise to significant blue-shifting and reshaping of the 3ω spectrum. More extreme blue-shifting is observed as the waveguide core dimensions are decreased, which have a greater average intensity in the silicon waveguide core and support only a single mode. As the mode confinement is increased, there is both a higher free-carrier density generated and a stronger interaction between the 3ω signal and the free-carrier plasma, both broadening the 3ω spectrum and shifting the spectral energy closer to $\lambda = 509\text{nm}$.

Secondly, XPM between the intense NIR pulses and the 3ω radiation will further blue-shift the 3ω spectrum. Since a large fraction of the 3ω radiation is generated near the peak of the $\lambda = 1550\text{nm}$ pulses and propagates at a slower group velocity than the $\lambda = 1550\text{nm}$ pulse, it will immediately interact with the temporal back of the $\lambda = 1550\text{nm}$ pulses. As the $\lambda = 1550\text{nm}$ pulse propagates faster than the 3ω radiation, the local intensity of the $\lambda = 1550\text{nm}$ radiation (as experienced by the 3ω radiation) decreases rapidly, and due to the optical Kerr effect, so does the local refractive index. Since the 3ω generated radiation experiences this rapid decrease in refractive index, its spectrum will blue-shift.

Further insight into the overall nonlinear interactions taking place can be gained by re-examining the $\lambda = 1550\text{nm}$ pulse spectra transmitted through the waveguides, as described in Section 4.4. Figure 5.11(a) depicts the NIR light transmitted through the $w = 260\text{nm}$, $L = 200\mu\text{m}$ waveguide at input powers of $P_{peak}^\omega = 10\text{W}$ (where no 3ω was

detected) and $P_{peak}^\omega = 100\text{W}$, on a log-log scale, along with a spectrum taken without the SOI waveguide. As P_{peak}^ω is increased from 10W to 100W, several new side lobes appear on the blue side of the spectrum, while the remainder of the $P_{peak}^\omega = 100\text{W}$ spectrum is essentially unchanged from both the $P_{peak}^\omega = 10\text{W}$ spectrum and the spectrum taken without the waveguide. The signatures of SPM are much more pronounced in the $w = 340\text{nm}$ waveguide, as shown in Fig. 5.11. At a low input power of $P_{peak}^\omega = 10\text{W}$, the NIR spectrum resembles the ones obtained from the $w = 260\text{nm}$ waveguide. When the input power is increased to $P_{peak}^\omega = 100\text{W}$, the NIR spectrum is broadened and strong sidebands are observed on both the red and blue side of the spectrum, clear evidence of SPM. Notably, the $w = 340\text{nm}$ waveguide exhibited a much stronger signature of SPM than the $w = 260\text{nm}$ waveguide. Yet, the deviation of the 3ω spectrum from the calculated spectrum was more pronounced for the $w = 260\text{nm}$ waveguide than for the $w = 340\text{nm}$ waveguide. This is direct evidence that the observed blue-shift does not arise from SPM of the NIR pulses, but from a combination of free-carrier plasma dispersion and XPM between the NIR and 3ω pulses.

5.2.6 Calculation of Conversion Efficiency

A key parameter to consider is the conversion efficiency of NIR photons to 3ω photons, η_{THG} , for each SOI waveguide. The output 3ω power is determined by calibrating the spectrometer with a coherent laser source at $\lambda = 532\text{nm}$. This radiation was generated from a commercially available solid state green laser ($\lambda = 532\text{ nm}$) that used a second-harmonic generation crystal to double the frequency of a $\lambda = 1064\text{nm}$ laser diode. A system of neutral density filters is used to vary the $\lambda = 532\text{nm}$ power coupled from free-space to the lensed optical fibre. The $\lambda = 532\text{nm}$ power coupled to the fibre is measured using a precision power meter and the corresponding counts detected by the spectrometer are recorded as a function of input power, yielding a linear trend with a calibration factor of $59\text{fW} \cdot \text{count}^{-1} \cdot \text{s}^{-1}$. The coupling efficiency of 3ω radiation from the end-facet of the waveguides to the lensed optical fibre is calculated via FDTD simulations and is found to vary between 11.5% for the $w = 260\text{nm}$ waveguide and 29.7% for the $w = 4150\text{nm}$ waveguide.

The conversion efficiency of each of the waveguides was calculated for a fixed peak input power, $P_{peak}^\omega = 100\text{W}$, and plotted in Fig. 5.12. Interestingly, the maximum η_{THG} was observed for the $w = 470\text{nm}$ waveguide, and was found to be $\eta_{THG} = 2.2 \times 10^{-6}$ at $P_{peak}^\omega = 100\text{W}$. For $w < 470\text{nm}$, the maximum intensity of the waveguide mode, I_{max}^ω , is higher (between $7.38 \times 10^{15}\text{W}/\text{m}^2 \leq I_{max}^\omega \leq 7.68 \times 10^{15}\text{W}/\text{m}^2$ at $P_{peak}^\omega = 100\text{W}$) and gives rise to stronger nonlinear loss (TPA and FCA), which attenuates the fundamental power before it reaches the output facet of the waveguide, resulting in a lower conversion efficiency. For $w > 470\text{nm}$, the lower intensity (between $1.15 \times 10^{15}\text{W}/\text{m}^2 \leq I_{max}^\omega \leq 6.59 \times 10^{15}\text{W}/\text{m}^2$ at $P_{peak}^\omega = 100\text{W}$) of the waveguide mode limits the conversion efficiency. Therefore, the

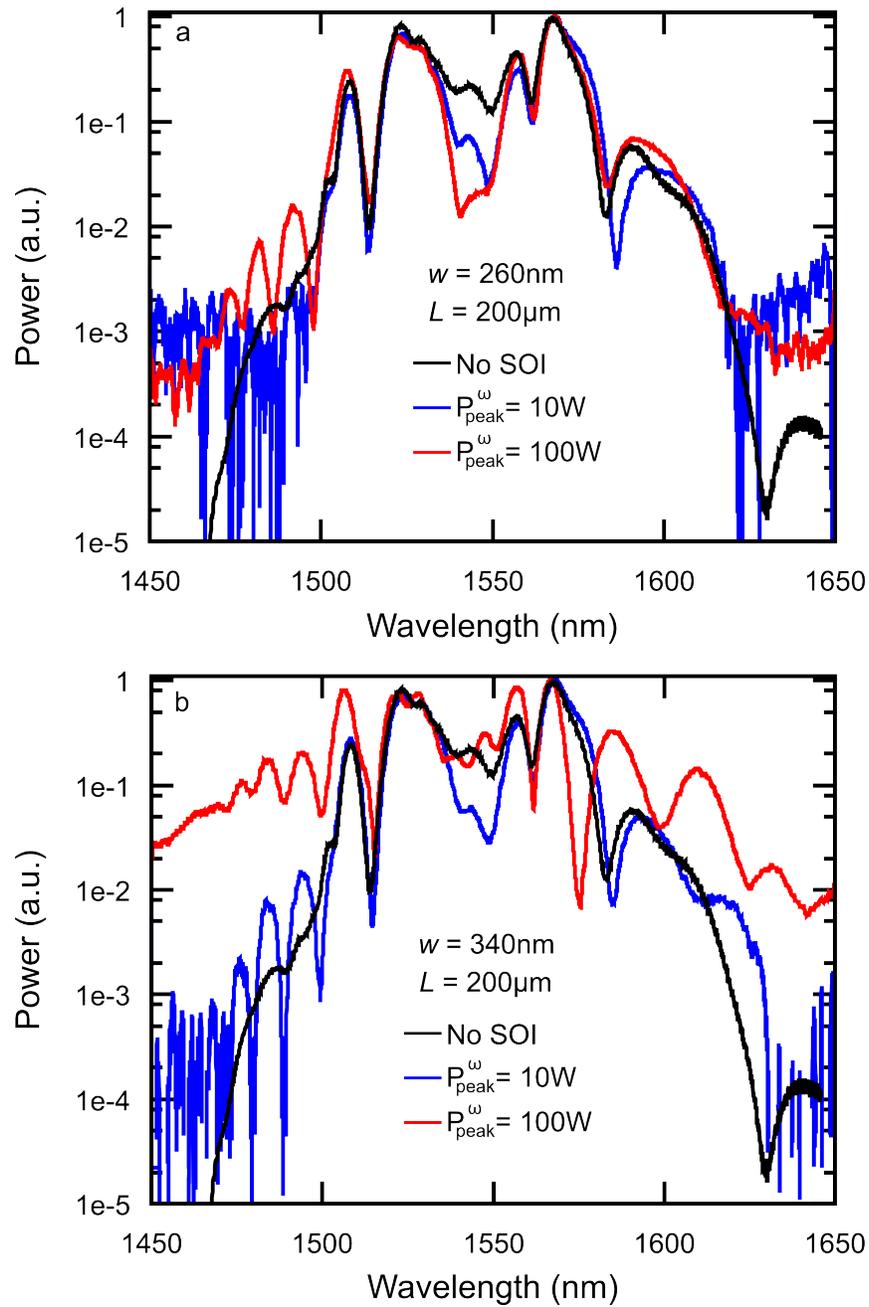


Figure 5.11: Transmitted fundamental spectra through a $L = 200\mu\text{m}$, $w = 260\text{nm}$, $h = 340\text{nm}$ waveguide for peak powers, $P_{peak}^{\omega} = 10\text{W}$ (blue) and $P_{peak}^{\omega} = 100\text{W}$ (red), along with a spectrum taken with no SOI waveguide. (b) Transmitted fundamental spectra through a $L = 200\mu\text{m}$, $w = 340\text{nm}$, $h = 340\text{nm}$ waveguide for peak powers, $P_{peak}^{\omega} = 10\text{W}$ (blue) and $P_{peak}^{\omega} = 100\text{W}$ (red), along with a spectrum taken with no SOI waveguide.

$w = 470\text{nm}$ waveguide provides an optimum balance between minimizing nonlinear loss and providing a high fundamental intensity. Increasing the peak power in the $w = 470\text{nm}$ SOI waveguide to $P_{peak}^\omega = 147\text{W}$ yielded a conversion efficiency of $\eta_{THG} = 3.9 \times 10^{-6}$, corresponding to a normalized conversion efficiency of $\eta_{norm} = \eta_{THG}/(P_{peak}^\omega)^2 = 1.8 \times 10^{-10}\text{W}^{-2}$. This represents an increase in η_{THG} by more than an order of magnitude and approximately the same η_{norm} that was demonstrated previously in a silicon PhC waveguide, in a structure with a footprint area that is reduced by a factor of 6.5. In addition, this structure is non-resonant and is suitable for converting a wide range of excitation frequencies to their third-harmonics, provided that the excitation photon energy is less than the bandgap of silicon, $E_g = 1.1\text{eV}$, and the wavelength is short enough that a significant fraction of the pulse energy is confined to the silicon core.

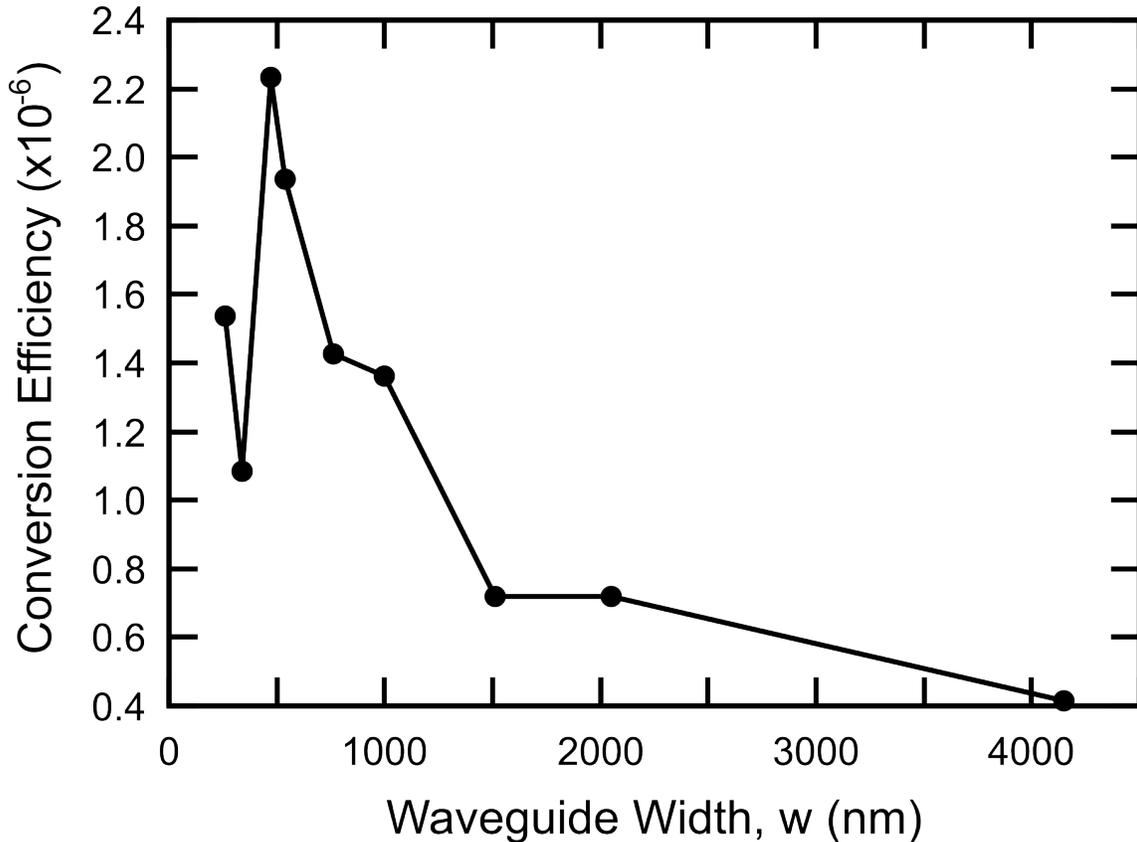


Figure 5.12: Third-harmonic generation conversion efficiency as a function of waveguide width, w , where each waveguide is excited with the same peak power, $P_{peak}^\omega = 100\text{W}$.

5.3 Third-Harmonic Generation in Ultrashort, High-Confinement SOI Waveguides

Based on the previous section, it is clear that high efficiency THG requires a combination of high waveguide mode confinement and a very short waveguide length. Under these

conditions, high mode confinement will give rise to a high intensity for exciting THG and a short length minimizes nonlinear losses. Furthermore, such a device would serve as a very compact source for visible light that could be integrated onto a CMOS electronic or silicon photonic chip with a very small footprint area. The potential for high-efficiency THG under these conditions is assessed by investigating very short SOI waveguides with a width, $w = 340\text{nm}$, a height, $h = 340\text{nm}$, and lengths, $L = \{1.8, 3.7, 5.2, 6.5, 7.6\}\mu\text{m}$, fabricated on a characterization beam as discussed in Section 3.3. This sample enabled the $\lambda = 1550\text{nm}$ femtosecond pulses to be coupled to the very short SOI waveguides, a strong third-harmonic signal to be generated in the silicon core of the waveguide, and the third-harmonic signal to be out-coupled into a lensed optical fibre, as shown in Fig. 5.13. At a peak excitation power, $P_{peak}^\omega = 37\text{W}$, bright green light emission was clearly visible from the end facets of the waveguides under normal room lighting.

An exemplary 3ω spectrum measured from the $L = 7.6\mu\text{m}$ waveguide is depicted in Fig. 5.13, along with the theoretical 3ω spectrum calculated from the free-space $\lambda = 1550\text{nm}$ femtosecond laser pulse spectrum (discussed in Appendix C). It should be noted that the calculated spectrum does not incorporate waveguide dispersion or non-uniform coupling of different spectral components of the pulse to the waveguide. However, it still reproduces most of the spectral features with relatively good agreement. Top-view microscope images of guided green light emission captured from representative $w = 340\text{nm}$ waveguides with lengths, $L = \{1.8, 3.7, 5.2, 6.5, 7.6\}\mu\text{m}$, are displayed in Figs. 2(c)-(g), respectively.

The coupled peak power of the $\lambda = 1550\text{nm}$ pulses was increased up to P_{peak}^ω , and surface plots were generated from the measured third-harmonic spectra taken at each input peak power. Surface plots for $L = \{3.7, 5.2, 6.5, 7.6\}\mu\text{m}$ are shown in Figs. 5.14(a)-(d), respectively, and a spectral slice for $P_{peak}^\omega = 100\text{W}$ is shown in a white curve overlay. In the same manner as the previous section, each spectral slice was normalized to visualize the spectral evolution with increasing power, and the results are shown in Fig. 5.14. Near the threshold for observing THG, the waveguides with lengths $L = \{3.7, 5.2, 6.5, 7.6\}\mu\text{m}$ have peak wavelengths of $\lambda = \{517.4, 517.7, 517.0, 517.0\}\text{nm}$. As the peak power is increased to $P_{peak}^\omega = 145\text{W}$, a slight blue-shift in the peak wavelength to $\lambda = 516.3\text{nm}$ is observed for the $L = 3.7\mu\text{m}$ waveguide. None of the other three waveguides produce a measurable change in the peak wavelength.

The third-harmonic conversion efficiency is calculated for each of the waveguides, and for a fixed peak power, $P_{peak}^\omega = 100\text{W}$, conversion efficiencies of $\eta_{THG} = \{7.4 \times 10^{-6}, 1.2 \times 10^{-5}, 8.7 \times 10^{-6}, 4.4 \times 10^{-6}\}$ are measured for waveguides with $L = \{3.7, 5.2, 6.5, 7.6\}\mu\text{m}$, respectively. The greatest conversion efficiency is measured in the $L = 5.2\mu\text{m}$ waveguide, and longer waveguides produce a slightly lower conversion efficiency, signifying that nonlinear loss can attenuate the pump pulses even over a very short length scale of several microns. Strangely, the shortest $L = 3.7\mu\text{m}$ waveguide produces a lower conversion efficiency than both the $L = 5.2\mu\text{m}$ and the $L = 6.5\mu\text{m}$ waveguides. This, in combination

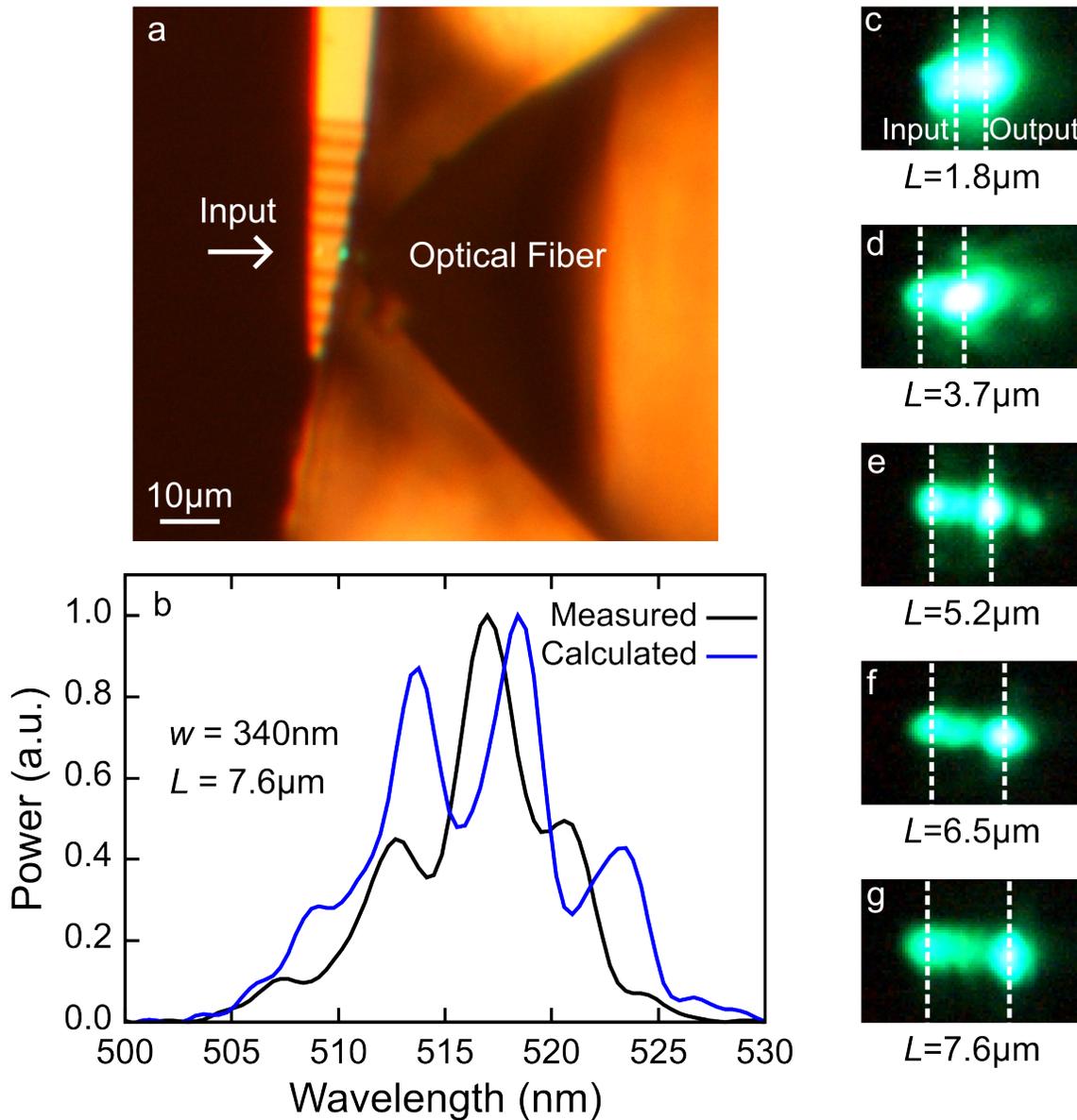


Figure 5.13: (a) Short SOI waveguide with a width, $w = 340\text{nm}$, a height, $h = 340\text{nm}$, and lengths in the range, $1.8\mu\text{m} \leq L \leq 7.6\mu\text{m}$. Femtosecond $\lambda = 1550\text{nm}$ pulses are coupled to the $L = 5.2\mu\text{m}$ waveguide facet on the left-hand side. Green light emission is clearly visible (under normal room light illumination) at the waveguide end facets and is collected with an optical fibre on the right-hand side. (b) Measured and calculated third-harmonic spectra for a $L = 7.6\mu\text{m}$ waveguide. Green light emission is observed with a $20\times$ microscope camera from SOI waveguides with lengths, $L =$ (c) $1.8\mu\text{m}$, (d) $5.2\mu\text{m}$, (e) $6.5\mu\text{m}$, and (f) $7.6\mu\text{m}$.

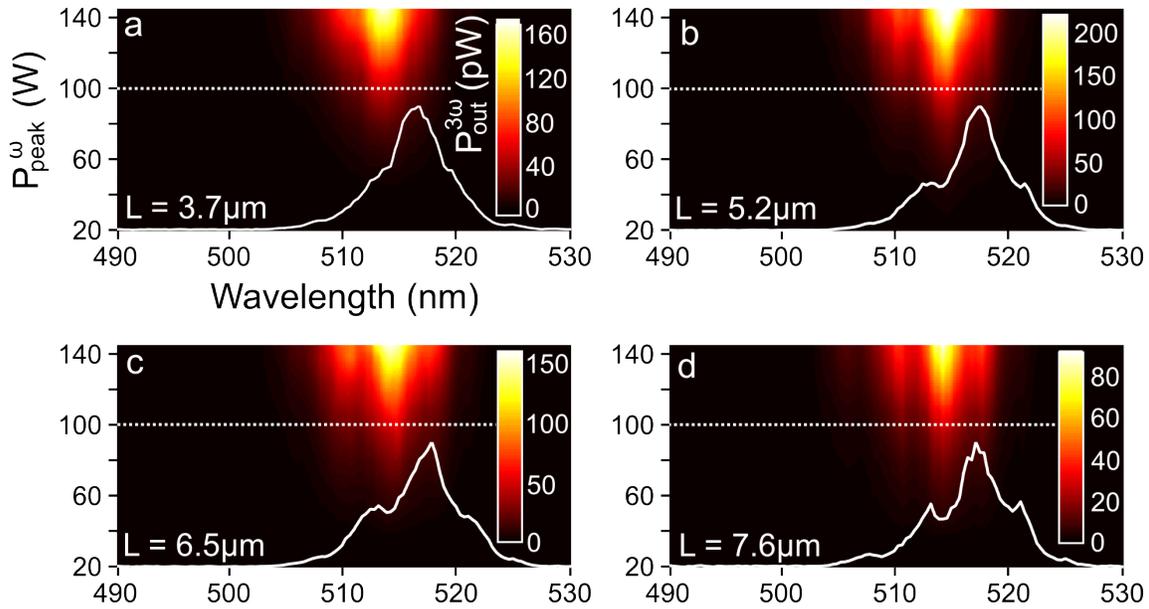


Figure 5.14: Intensity surfaces depicting the THG spectrum growth with increasing peak input power, P_{peak}^{ω} , for SOI waveguides with $L =$ (a) $3.7 \mu\text{m}$, (b) $5.2 \mu\text{m}$, (c) $6.5 \mu\text{m}$, and (d) $7.6 \mu\text{m}$. The white line overlay portrays the third-harmonic spectrum measured at the dashed section, corresponding to $P_{peak}^{\omega} = 100 \text{W}$.

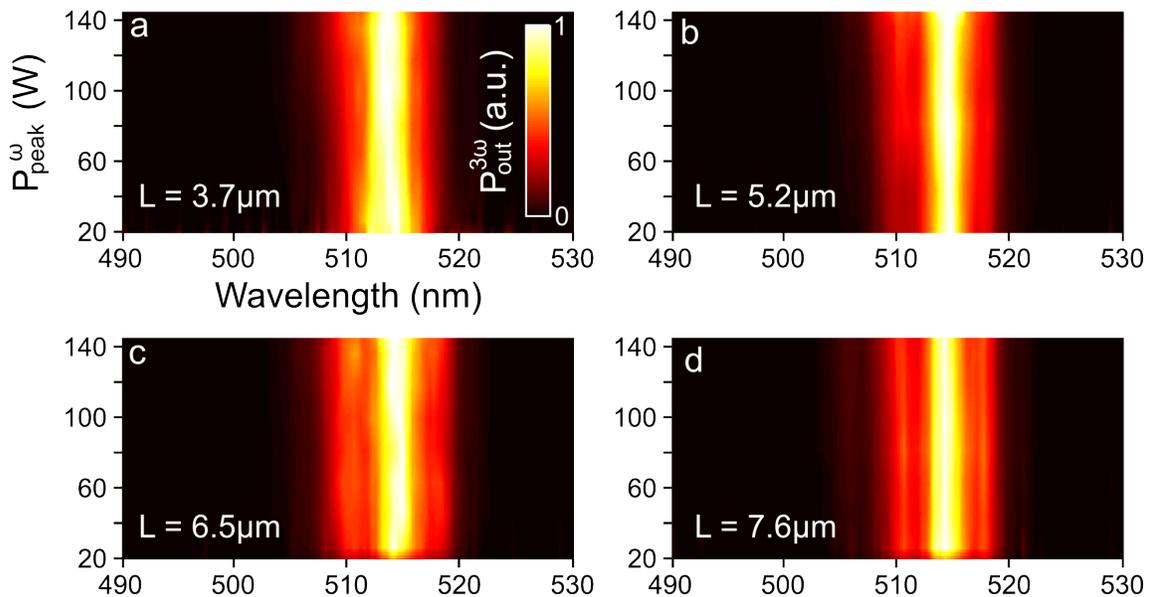


Figure 5.15: Intensity surfaces depicting the normalized third-harmonic spectrum with increasing peak input power for SOI waveguides with $L =$ (a) $3.7 \mu\text{m}$, (b) $5.2 \mu\text{m}$, (c) $6.5 \mu\text{m}$, and (d) $7.6 \mu\text{m}$.

with the small, yet measurable, blue-shift in the peak wavelength of its spectrum, suggest that the intensity present near the output facet of this waveguide is significantly higher than what is present in the other waveguides, and that nonlinear loss can both decrease the conversion efficiency and blue-shift the 3ω spectrum wavelength.

The highest absolute conversion efficiency was measured in the $L = 5.2\mu\text{m}$ waveguide at a peak input power, $P_{peak}^\omega = 145\text{W}$, and was calculated to be, $\eta_{THG} = 2.8 \times 10^{-5}$, corresponding to a normalized conversion efficiency of $\eta_{norm} = 1.3 \times 10^{-9}\text{W}^{-2}$. This is both the highest absolute and relative conversion efficiency that have been demonstrated in a silicon structure of any kind, to date. These measurements were demonstrated in a waveguide with an ultracompact footprint area of $1.77\mu\text{m}^2$, which is approximately 344 smaller than the Si PhC that was discussed earlier [18].

5.4 Summary

In summary, this chapter has presented comprehensive investigations into THG in SOI waveguides in the green region of the visible spectrum. Interference patterns in the scattered green light from a $w = 260\text{nm}$ waveguide are observed and are attributed to interference between the TM mode and a leaky mode of the waveguide. Detailed studies on the influence of mode confinement on THG are performed in $L = 200\mu\text{m}$ waveguides. In large, multi-mode waveguides a narrow third-harmonic peak is observed at 3ω ($\lambda = 517\text{nm}$). Curiously, as the cross-sectional dimensions of the waveguide are decreased, the THG peak wavelength is blue-shifted from $\lambda = 517\text{nm}$ to $\lambda \approx 509\text{nm}$, signifying that the 3ω radiation experiences a rapidly decreasing refractive index. In addition, the FWHM increases from $FWHM = 5.5\text{nm}$ to $FWHM = 19\text{nm}$. For short waveguides with $L \leq 10\mu\text{m}$, the 3ω spectrum agrees well with the expected spectrum from the laser pulses.

The observed THG is extremely broadband, and can be achieved in a CMOS electronic and silicon photonic compatible structure that is simple to fabricate, occupies an ultracompact footprint area of $1.77\mu\text{m}^2$, and generates the 3ω radiation in a manner that allows it to be routed on a chip. As non-resonant structures, the SOI waveguides can be used to convert a wide range of wavelengths to their third-harmonics, permitting generation of the entire visible spectrum in a single structure. It was shown that a conversion efficiency up to $\eta_{THG} = 2.8 \times 10^{-5}$ can be achieved, which is more than two orders of magnitude greater than the highest previous demonstration of THG in a silicon structure. These results demonstrate a unique, enabling platform for integrated light sources spanning the visible to NIR wavelengths.

References

- [1] R. Dekker, N. Usechak, M. Forst, and A. Driessen. Ultrafast nonlinear all-optical processes in silicon-on-insulator waveguides. *J. Phys. D: Appl. Phys.*, 40:R249–R271, 2007.
- [2] R. Espinola, J. Dadap, Jr. R. Osgood, S. McNab, and Y. Vlasov. C-band wavelength conversion in silicon photonic wire waveguides. *Opt. Express*, 13:4341–4349, 2005.
- [3] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, T. Tsuchizawa, T. Watanabe, J. I. Takahashi, and S. I. Itabashi. Four-wave mixing in silicon wire waveguides. *Opt. Express*, 13:4629–4637, 2005.
- [4] M. Foster, A. Turner, J. Sharping, B. Schmidt, M. Lipson, and A. Gaeta. Broad-band optical parametric gain on a silicon photonic chip. *Nature*, 441:960–963, 2006.
- [5] Q. Lin, J. Zhang, P. M. Fauchet, and G. P. Agrawal. Ultrabroadband parametric generation and wavelength conversion in silicon waveguides. *Opt. Express*, 14:4786–4799, 2006.
- [6] M. A. Foster, A. C. Turner, R. Salem, M. Lipson, and A. L. Gaeta. Broad-band continuous-wave parametric wavelength conversion in silicon nanowaveguides. *Opt. Express*, 15:12949–12958, 2007.
- [7] R. Salem, M. Foster, A. Turner, D. Geraghty, M. Lipson, and A. L. Gaeta. Signal regeneration using low-power four-wave mixing on silicon chip. *Nat. Photon.*, 2:35–38, 2008.
- [8] R. Dekker, A. Driessen, T. Wahlbrink, C. Moormann, J. Niehusmann, and M. Först. Ultrafast kerr-induced all-optical wavelength conversion in silicon waveguides using 1.55 μm femtosecond pulses. *Opt. Express*, 14:8336–8346, 2006.
- [9] O. Boyraz, P. Koonath, V. Raghunathan, and B. Jalali. All optical switching and continuum generation in silicon waveguides. *Opt. Express*, 12:4094–4102, 2004.
- [10] O. Boyraz, T. Indukuri, and B. Jalali. Self-phase-modulation induced spectral broadening in silicon waveguides. *Opt. Express*, 12:829–834, 2004.

- [11] G. W. Rieger, K. S. Virk, and J. F. Young. Nonlinear propagation of ultrafast 1.5 μm pulses in high-index-contrast silicon-on-insulator waveguides. *Applied Physics Letters*, 84:900–902, 2004.
- [12] A. Cowan, G. Rieger, and J. Young. Nonlinear transmission of 1.5 μm pulses through single-mode silicon-on-insulator waveguide structures. *Opt. Express*, 12:1611–1621, 2004.
- [13] E. Dulkeith, Y. A. Vlasov, X. Chen, N. C. Panoiu, and Jr. R. M. Osgood. Self-phase-modulation in submicron silicon-on-insulator photonic wires. *Opt. Express*, 14:5524–5534, 2006.
- [14] I. W. Hsieh, X. Chen, J. I. Dadap, N. C. Panoiu, R. M. Osgood, S. J. McNab, and Y. A. Vlasov. Ultrafast-pulse self-phase modulation and third-order dispersion in silicon photonic wire-waveguides. *Opt. Express*, 14:12380–12387, 2006.
- [15] L. Yin and G. P. Agrawal. Impact of two-photon absorption on self-phase modulation in silicon waveguides. *Opt. Lett.*, 32:2031–2033, 2007.
- [16] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich. Generation of optical harmonics. *Phys. Rev. Lett.*, 7:118–119, 1961.
- [17] W. Seka, S. D. Jacobs, J. E. Rizzo, R. Boni, and R. S. Craxton. Demonstration of high efficiency third harmonic conversion of high power nd-glass laser radiation. *Opt. Commun.*, 34:469–473, 1980.
- [18] C. Corcoran, C. Monat, C. Grillet, D. J. Moss, B. J. Eggleton, T. P. White, L. O’Faolain, and T. F. Krauss. Green light emission in silicon through slow-light enhanced third-harmonic generation in photonic-crystal waveguides. *Nature Photon.*, 3:206–210, 2009.
- [19] J. S. Levy, M. A. Foster, A. L. Gaeta, and M. Lipson. Harmonic generation in silicon nitride ring resonators. *Opt. Express*, 19:11415–11421, 2011.
- [20] M. Mansuripur. Distribution of light at and near the focus of high-numerical-aperture objectives. *J. Opt. Soc. Am. A*, 3:2086–2093, 1986.

Chapter 6

Ultrafast Nonlinear Effects and Strong-field Phenomena in Silicon-loaded Nanoplasmonic Waveguides

6.1 Introduction

Perhaps the most immediate and far-reaching application of nonlinear plasmonics is active nanoplasmonic circuitry, where efficient excitation of nonlinearities is leveraged to modulate the propagating light signal [1–7]. Pioneering experimental investigations into nonlinear optical phenomena were performed in bulk materials, where transparent materials permitted a nonlinear interaction to accumulate under appropriate phase matching conditions, and lossy materials lent themselves to localized reflection interactions occurring within the skin depth of the constituent frequencies [8]. Subsequently, fibre optics and chip-scale waveguides enabled high electromagnetic intensities to be maintained over long interaction lengths, generating pronounced nonlinear signatures [9, 10]. Amassing strong nonlinear signatures over short interaction lengths and at low input power has proven challenging, yet such devices are important elements of next-generation integrated optical nanocircuitry. Subwavelength confinement, electric field enhancement, and high sensitivity provided by nanoplasmonic structures make them excellent candidates for compact, all-optical circuitry with gate speeds exceeding those of conventional electronics by orders of magnitude.

As described in the previous chapters, a number of third-order nonlinear interactions have been investigated in Si photonic waveguides, including two-photon absorption (TPA) [11, 12], free-carrier absorption (FCA), Raman amplification [13, 14], the optical Kerr effect, four-wave mixing [15, 16], cross-phase modulation [17, 18], and self-phase modulation [19, 20]. Third-harmonic generation (THG) has proven much more challenging to study,

with only a single demonstration of THG in a silicon-on-insulator (SOI) photonic crystal structure [21] prior to the publication of this thesis and extensive investigations presented in the previous chapter. Third-harmonic generation has also been observed in plasmonic nanoparticles [22–24], but integration of these particles to optical circuitry has not yet been achieved. Third-harmonic generation enhanced by strong confinement in a Si-loaded nanoplasmonic waveguide offers a novel platform for generating on-chip ultrafast green light sources.

Despite prior promising developments, the maximum integration density of Si photonic devices is limited by the diffraction limit, radiation loss that accrues around sharp bends, and cross-talk between adjacent waveguides. Moreover, the nonlinear interaction length of active devices is typically on the order of hundreds of microns, and such a footprint hinders its potential for large-scale integrated circuitry. The strong coupling of electromagnetic energy to the free electrons of noble metal features in nanoplasmonic waveguides enables highly confined waveguide modes at subwavelength dimensions and propagation around sharp bends. It is natural to consider Si-loaded plasmonic waveguides fabricated on an SOI platform, which would maintain the possibility for monolithic integration with complementary metal-oxide-semiconductor (CMOS) electronic devices and Si photonic devices [25]. Electric field enhancement at the metal-Si interfaces would bolster $\chi^{(3)}$ nonlinear interactions, making possible the development of optical devices with advanced functionality.

Furthermore, Si-loaded nanoplasmonic waveguides enable advanced functionalities in hybrid electronic-photonic circuitry, since the nanoplasmonic devices include high conductivity metallic features. In this respect, hybrid plasmonic-electronic devices can be integrated within the current electronic device architecture where the nanoscale metallic-dielectric waveguides allow for the transport of both electrical and optical signals simultaneously on the same platform. This is of particular importance because the plasmonic device's size and the nonlinear interactions can be brought down to several nanometers, making them compatible with advanced CMOS processor circuit architectures [26, 27]. The Si-based nanoplasmonic waveguides can also be efficiently interfaced to less lossy plasmonic waveguides or silicon photonic waveguides for routing the signal over longer distances [25]. Although plasmonic losses limit the nonlinear interaction length to several microns, compensation provided by the electric field enhancement enables marked nonlinear interaction in a compact device.

6.2 Passive Operation

¹ A key challenge to nanoplasmonic device implementation has been their efficient integration with existing photonic systems and the macroscopic world. Various schemes such

¹A version of this section has been published: S. Sederberg, V. Van, and A. Y. Elezzabi, *Applied Physics Letters* **96**, 121101 (2010). Copyright (2010) American Institute of Physics.

as a SOI photonic waveguide combined with a plasmonic taper, a prism wedge structure, a grating, a nanoplasmonic antenna, evanescent coupling from a dielectric waveguide, or end-fire coupling from a tapered SOI waveguide have been used previously to couple light into plasmonic waveguides. End-fire coupling from a SOI waveguide effectively interfaces silicon photonic waveguides with plasmonic waveguides on a platform that is largely compatible with CMOS processing.

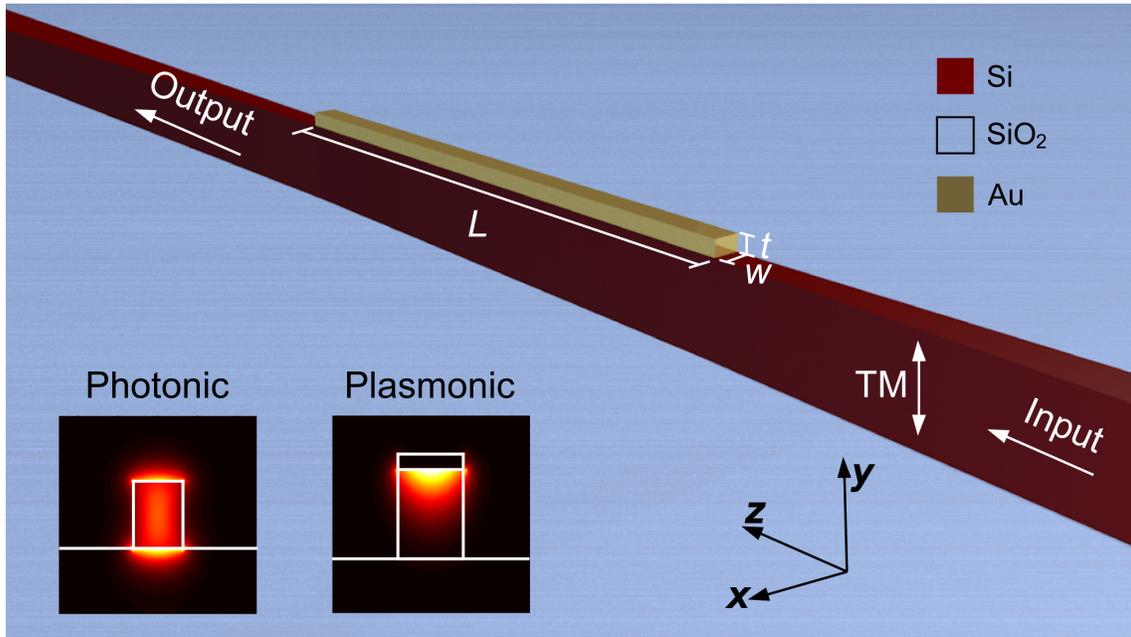


Figure 6.1: Schematic depiction of a silicon-based nanoplasmonic waveguide interfaced to silicon photonic input and output couplers. The insets depict the mode profiles intensity (I) distributions for the photonic and plasmonic waveguides.

The plasmonic waveguide under investigation consists of a $300\text{nm} \times 340\text{nm}$ SOI waveguide capped by a 50nm layer of gold on the top. In order to excite a pure plasmonic mode in this structure, y -polarized light at $\lambda = 1550\text{nm}$ is coupled into it using a $2\mu\text{m}$ wide SOI waveguide that tapers linearly to a width of $w = 300\text{nm}$ over a distance of $55\mu\text{m}$, as shown in Fig. 6.1. An identical taper collects the light that is transmitted by the nanoplasmonic waveguide and delivers it to a detector. The input and output SOI waveguides are aligned to input/output lensed single mode fibres (SMF) using piezo-controlled nanopositioning stages. The nanoplasmonic waveguide and couplers are fabricated according to the process discussed in Section 3.5. Scanning electron micrographs of fabricated straight and s-bend plasmonic waveguides are shown in Fig. 6.2.

A fundamental concern that must be addressed is whether the proposed coupling scheme will excite the desired plasmonic mode. Three-dimensional (3D) finite-difference time-domain (FDTD) simulations were performed on a Si photonic waveguide interfaced to a silicon-based plasmonic waveguide to investigate this matter. The simulation geometry is represented in Fig. 6.1. Here, a quasi-TM mode at $\lambda = 1550\text{nm}$ is excited in a $300\text{nm} \times 340\text{nm}$ SOI waveguide. After propagating $2\mu\text{m}$ through the SOI photonic waveguide,

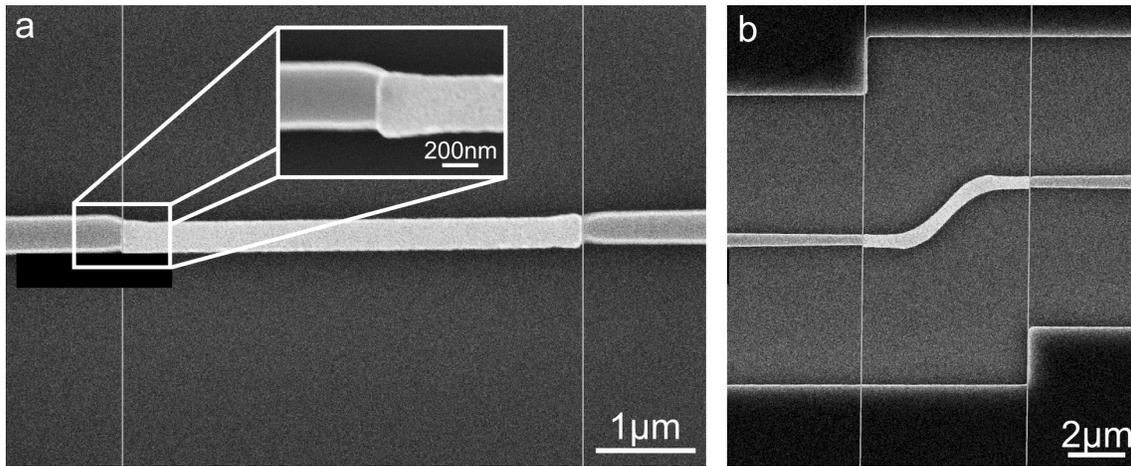


Figure 6.2: (a) Scanning electron micrograph of a silicon-based plasmonic waveguide interfaced to silicon photonic input and output tapered waveguides. The silicon-based plasmonic waveguide has a width, $w = 300\text{nm}$, a height, $h = 340\text{nm}$, and a length, $L = 4\mu\text{m}$. The inset reveals a $\sim 35\text{nm}$ misalignment between the photonic and plasmonic waveguides. (b) Scanning electron micrograph of a silicon-based plasmonic s-bend, with a transverse displacement of $\Delta x = 2\mu\text{m}$ and an overall length of $L = 6.3\mu\text{m}$.

the mode encounters the silicon-based plasmonic waveguide that consists of the same $300\text{nm} \times 340\text{nm}$ Si core that is capped by a $t = 50\text{nm}$ layer of gold and has a length of $L = 5\mu\text{m}$. Fields are recorded in various planes through the structure to observe the excitation of the plasmonic mode and to quantify the coupling efficiencies and propagation losses. Figure 6.3(a) displays a time snapshot of the electric fields propagating through the Si photonic waveguide ($-2\mu\text{m} \leq z \leq 0\mu\text{m}$), exciting the plasmonic waveguide ($z = 0\mu\text{m}$), the plasmonic mode propagation ($0\mu\text{m} \leq z \leq 5\mu\text{m}$), and coupling back into the output Si photonic waveguide ($5\mu\text{m} \leq z \leq 6.5\mu\text{m}$), taken in the $x = 0.0\mu\text{m}$ plane. Another snapshot taken in the $y = 330\text{nm}$ plane is shown in Fig. 6.3(b), depicting the tightly confined plasmonic fields close to the Au-Si interface. Figure 6.3(c) shows the time-averaged electric field distribution in the $z = -0.5\mu\text{m}$ plane, representing the quasi-TM mode profile in the Si photonic waveguide. The excitation of the plasmonic mode occurs via diffraction at the edge of the plasmonic waveguide at $z = 0\mu\text{m}$, and the time-averaged electric field distribution in the $z = 0.0\mu\text{m}$ plane is shown in Fig. 6.3(d). Close examination of this figure reveals that strong diffraction occurs directly adjacent to the edge of the gold film at $z = 0.0\mu\text{m}$ and electric fields are enhanced by a factor of ~ 2.12 in this region. This diffraction effectively excites the plasmonic mode with a coupling efficiency of 76.5%. Despite the fact that the Si core dimensions of the Si-based plasmonic waveguide are the same as those of the Si photonic waveguide, there is no excitation of a photonic mode in the Si-based plasmonic waveguide. The time-averaged electric field distribution taken at $z = 0.5\mu\text{m}$ is shown in Fig. 6.3(e). The electric fields are strongly confined to the Au-Si interface, representing a plasmonic excitation. Based on the recorded fields, the propagation distance of $\lambda = 1550\text{nm}$ radiation through the Si-based plasmonic waveguide

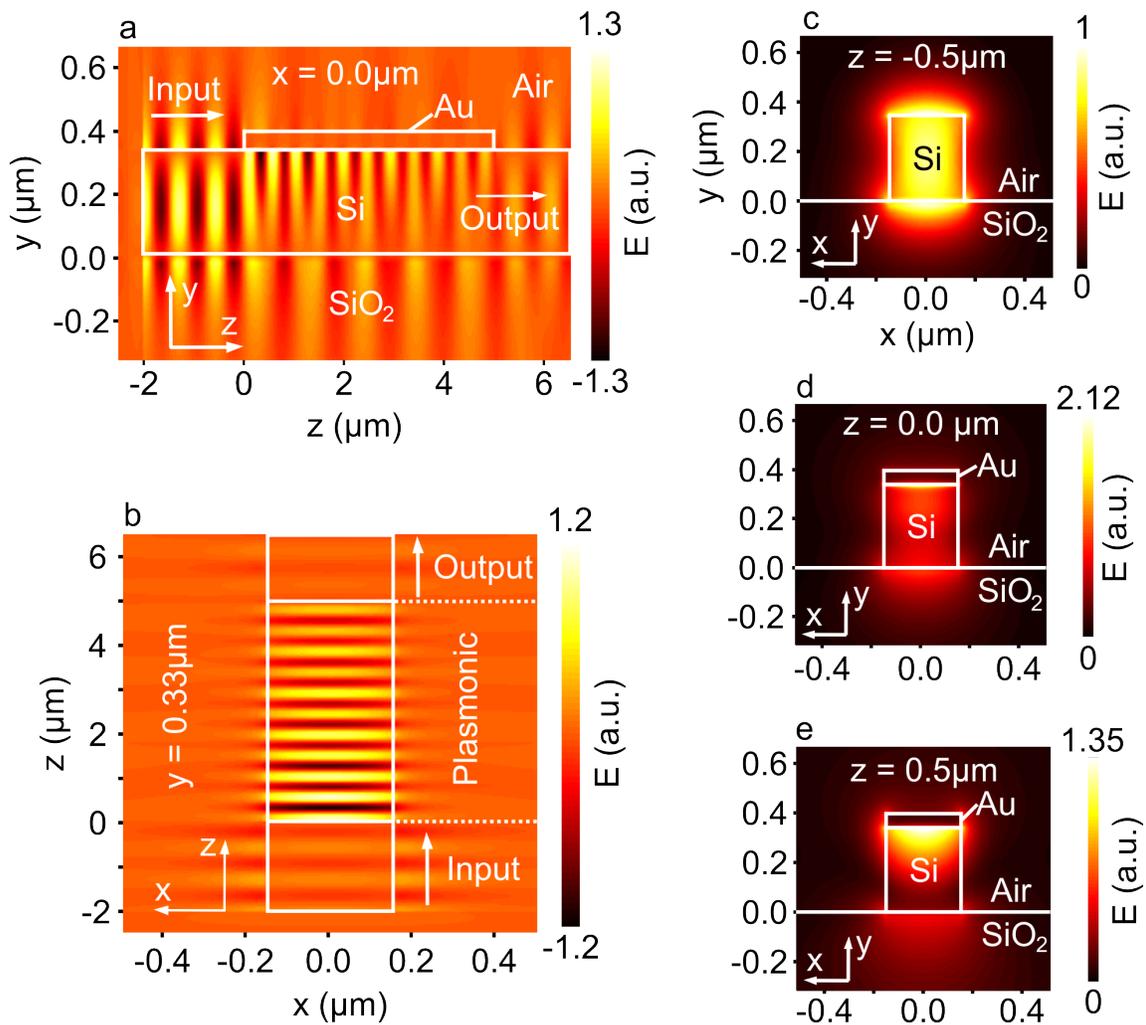


Figure 6.3: Simulation of interfacing between silicon photonic and silicon-based plasmonic waveguides. (a) Snapshot of electric fields (E_y) in the $x = 0\mu\text{m}$ plane, depicting the excitation of a plasmonic mode with a silicon photonic input coupler and the subsequent out-coupling to the output silicon photonic waveguide. (b) Snapshot of electric fields (E_y) in the $z = 330\text{nm}$ plane, highlighting the excitation of a plasmonic mode and the tight confinement near the Au-Si interface. (c) Electric field ($|E|$) distribution of the silicon photonic mode taken $z = -500\text{nm}$ before the beginning of the plasmonic waveguide. (d) Electric field ($|E|$) distribution at the interface between the silicon photonic and silicon-based plasmonic waveguides, showing strong diffraction at the Au interface. (e) Electric field ($|E|$) distribution of the plasmonic mode, taken $z = 500\text{nm}$ after the beginning of the plasmonic waveguide.

was determined to be $L_{prop} = 4.5\mu\text{m}$.

Subsequent mode solver simulations were performed to confirm that only the plasmonic mode was excited with a high efficiency. The Si-based plasmonic waveguide supports a total of 4 modes (including those with mixed TE and TM polarizations), with intensity distributions shown in Fig. 6.4. The lowest-order plasmonic mode is shown in Fig. 6.4(a), and a second-order plasmonic mode is shown in Fig. 6.4(b). Figure 6.4(c) depicts a quasi-TE photonic-like mode that suffers relatively high losses due to the close proximity of the metallic features. Finally, Fig. 6.4(d) shows a mode that is confined mainly to the Si-SiO₂ interface. A summary of the effective refractive index, propagation losses and TE polarization fraction of each of the modes is summarized in Table 6.1. Clearly, the plasmonic mode profile obtained from the mode solver (Fig. 6.4(a)) most closely resembles the mode excited by coupling radiation from a SOI photonic waveguide to a nanoplasmonic waveguide (Fig. 6.3(e)). Therefore, it is concluded that the other three modes calculated by the mode solver are not excited with a high efficiency.

Mode #	Effective Refractive Index	Propagation Loss (dB/ μm)	TE Polarization Percentage
1	3.185	0.881	6
2	1.974	0.788	45
3	1.580	0.518	70
4	1.442	0.158	11

Table 6.1: Summary of Mode Propagation Characteristics

Four different lengths, $L = 2, 3, 4,$ and $5\mu\text{m}$, of silicon-based plasmonic waveguides with widths of $w = 300\text{nm}$ were used to investigate this structure experimentally. For each waveguide, radiation near $\lambda = 1550\text{nm}$ was end-fire coupled into the input facet of the Si waveguide using a lensed SMF with a focal spot diameter, $d = 2.5\mu\text{m}$. The transmitted radiation was outcoupled from the output facet of the Si waveguide and delivered to a photodetector. Using a custom-designed Matlab graphical user interface, the laser wavelength was scanned in the range, $1520\text{nm} \leq \lambda \leq 1615\text{nm}$, and the transmitted power at each wavelength was measured. The code used to automate these measurements is presented in Appendix F.

Transmitted powers for this set of waveguides at $\lambda = 1550\text{nm}$ are shown in Fig. 6.5(a), from which a propagation distance of $2.00\mu\text{m}$ ($Loss = 2.17\text{dB}/\mu\text{m}$) was extracted. Notably, this value is lower than the predicted $4.5\mu\text{m}$ from the FDTD calculation. The greater experimental losses arise from the presence of a 5nm thick Cr adhesion layer between the Si and Au films. Chromium is much lossier than Au at $\lambda = 1550\text{nm}$, and due to the intense interaction at the Si-Cr-Au interfaces, this is detrimental to the propagation length. A sputtered Cr film with a 5nm thickness will be marginally continuous and will not be accurately represented as a continuous film in the simulations. Through subsequent simulations, it was found that a 2.5nm thick continuous Cr adhesion layer reproduced the propagation distance of $L_{prop} = 2.0\mu\text{m}$. As shown in Fig. 6.5(b), the propagation

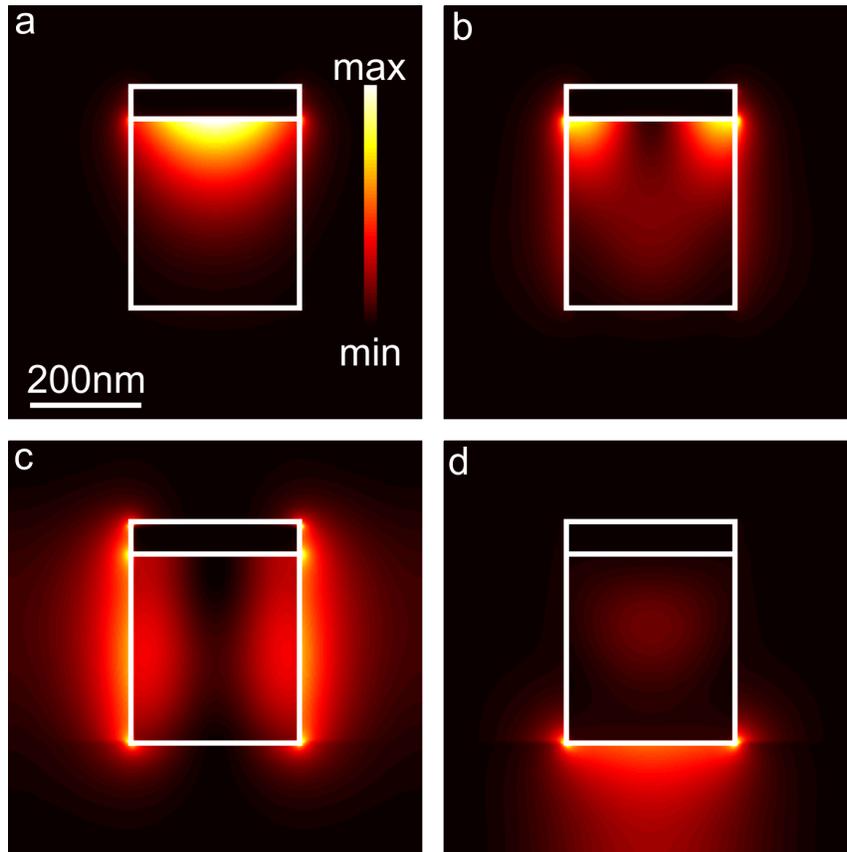


Figure 6.4: Intensity (I) distributions of the modes supported by a Si-based nanoplasmonic with $w = 300\text{nm}$ and $h = 340\text{nm}$.

losses decrease from $Loss = 3.01\text{dB}/\mu\text{m}$ at $\lambda = 1520\text{nm}$ to $1.54\text{dB}/\mu\text{m}$ at $\lambda = 1615\text{nm}$, demonstrating that this structure is not restricted to operation at a single wavelength and is suitable for wavelength multiplexing schemes.

The y -intercept ($8.3\mu\text{W}$) of the propagation curve in Fig. 6.5(a) provides the overall coupling losses resulting from the input/output facets of the SOI waveguide, taper, and impedance and mode mismatches between the SOI waveguides and silicon-based plasmonic waveguides. To deduce the coupling losses due to the SOI input facet, measurements are performed on a straight $2\mu\text{m} \times 340\text{nm} \times 3\text{mm}$ SOI reference waveguide without any Au features. The power transmitted through the reference waveguide at $\lambda = 1550\text{nm}$ was measured to be $P_{out}^{\omega} = 56.5\mu\text{W}$. Based on these values, the coupling efficiency between the SOI platform and the silicon-based plasmonic platform is determined to be 38.0% per interface, which is approximately half of the simulated value. As shown in the inset of Fig. 6.2(a), a slight misalignment of approximately 35nm occurred between subsequent layers of electron beam lithography, reducing the coupling efficiency. In the ideal case of perfect alignment, it is anticipated that the coupling efficiency would approach the theoretical value.

Along with the ability to guide a desired mode, it is important that a waveguide is able to transport optical energy around sharp bends to allow for routing that resembles

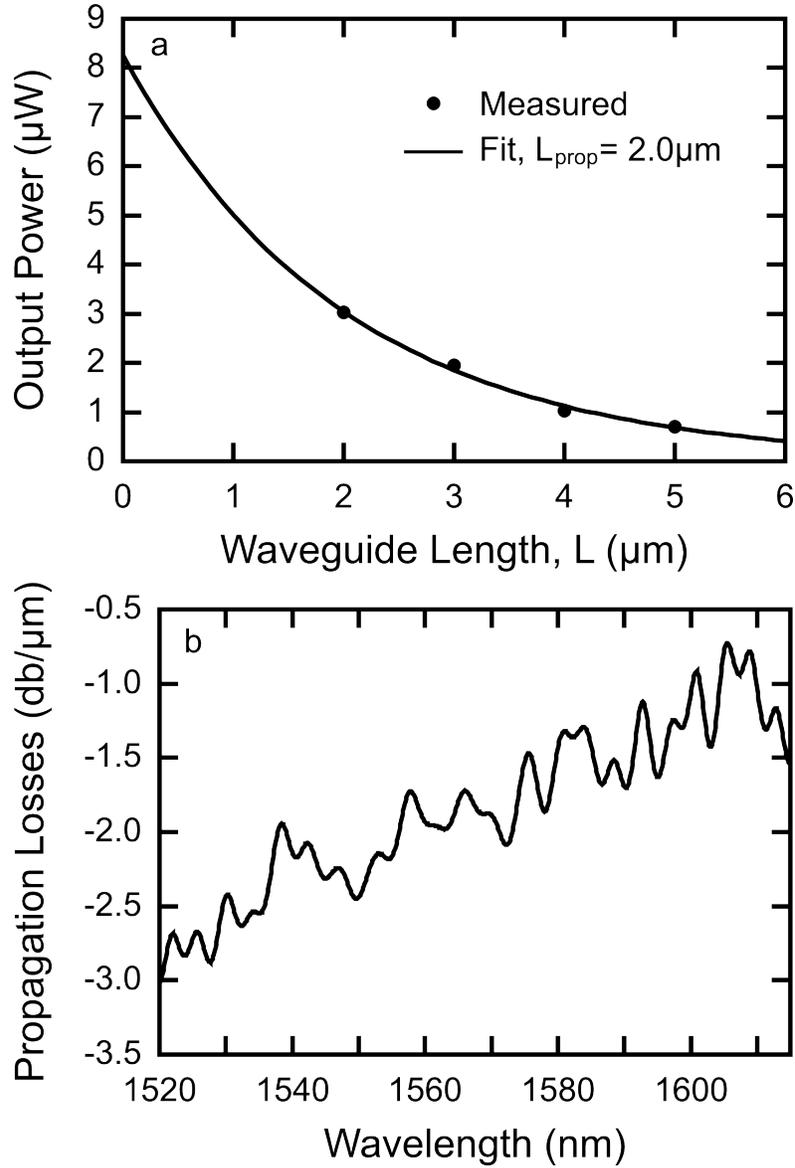


Figure 6.5: Measured power transmission through $w = 300\text{nm}$, $h = 340\text{nm}$ silicon-based plasmonic waveguides. (a) Power transmitted through waveguides with lengths of $L = \{2.0, 3.0, 4.0, 5.0\}\mu\text{m}$ at $\lambda = 1550\text{nm}$. An exponential fit to the data points reveals a propagation length of $L_{\text{prop}} = 2.0\mu\text{m}$ and a coupling efficiency of 38.0% per interface. (b) Broadband propagation losses, measured in the range of $1520\text{nm} \leq \lambda \leq 1615\text{nm}$.

modern CMOS circuits. Routing in this manner was demonstrated with an s-bend silicon plasmonic waveguide having a width of $w=400\text{nm}$, a transverse displacement of $\Delta x = 2\mu\text{m}$ and an overall length of $6.3\mu\text{m}$ is fabricated, as shown in Fig. 6.2(b). The power transmitted through this structure at $\lambda = 1550\text{nm}$ was measured to be $0.41\mu\text{W}$ and a bending loss of 14.6% is calculated. In addition to demonstrating efficient guiding around sharp bends, this result provides additional confirmation that a guided mode is being observed in this experiment.

6.3 Active Device Design

In the interest of exciting nonlinearities in a silicon-based nanoplasmonic waveguide, it is advantageous to decrease the Si core dimensions to achieve higher mode confinement and more pronounced nonlinear interactions. A perspective view of a more compact nanoplasmonic waveguide, with width, $w = 95\text{nm}$, height, $h = 340\text{nm}$, length, $L = 5\mu\text{m}$, and a $t = 60\text{nm}$ gold cap is shown in Fig. 6.6(a). Figure 6.6(b) depicts the plasmonic mode profile, which has an effective index, $n_{eff} = 2.65$, and propagates a distance, $L_{prop} = 3.1\mu\text{m}$ ($Loss = 1.4\text{dB}/\mu\text{m}$), before attenuating to e^{-1} of its initial amplitude. The intensity distributions along the two longitudinal cross-sections in Figs. 6.6(c),(d) portray the exponential decay of electromagnetic energy as the plasmonic mode propagates through the structure.

To measure ultrafast nonlinear dynamics, it is crucial to realize a means to excite a nonlinear interaction that is isolated to the structure of interest and not intertwined with a nonlinear excitation in an on-chip coupler. In the case of the previous section, the femtosecond pulses were propagated through a $\sim 1.5\text{mm}$ long Si photonic waveguide before being coupled into the Si-based nanoplasmonic waveguide. Significant nonlinear signatures will accumulate in such a long structure, making it impossible to accurately separate the nonlinear interactions taking place in the Si-based nanoplasmonic waveguide from those taking place in the Si photonic coupler. By eliminating the on-chip coupler altogether, free-space radiation can be coupled directly into the waveguide using a microscope objective and out-coupled using a lensed SMF. This scheme is realized by fabricating nanoplasmonic waveguides on a narrow beam that is etched to a depth of $d_{etch} = 75\mu\text{m}$, allowing for in-plane access with lensed SMFs. The processes used to fabricate these structures are discussed in Section 3.3. An artistic rendering of the sample along with the excitation and detection scheme is shown in Fig. 6.7(a). A scanning electron micrograph (SEM) of a nanoplasmonic waveguide before integration onto the beam is shown in Fig. 6.7(b). A beam including several nanoplasmonic waveguides embedded in silica cladding is shown in Fig. 6.7(c), and the inset shows a cross-sectional SEM of the nanoplasmonic waveguide end facet. Cross-sectional dimensions of the Si core are measured to be $w \times h = 95\text{nm} \times 340\text{nm}$ and the Au cap has a thickness, $t = 60\text{nm}$.

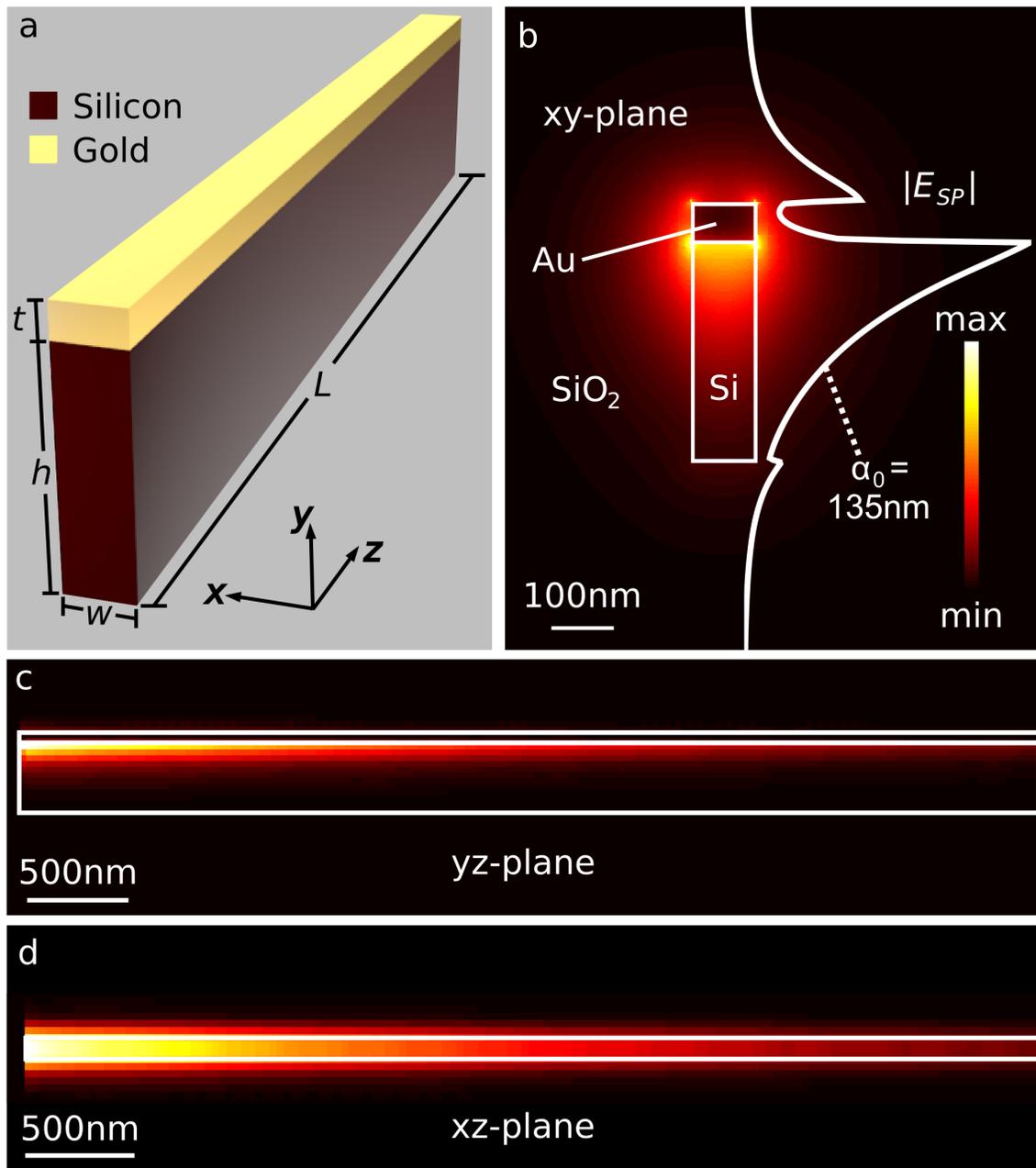


Figure 6.6: (a) Schematic depiction of the Si-based nanoplasmonic waveguide with $w = 95\text{nm}$, $h = 340\text{nm}$, and $L = 5\mu\text{m}$. (b) Time-averaged electric field amplitude ($|E|$) distribution in the xy -plane of the structure. (c) Time-averaged intensity (I) distribution in the yz -plane of the structure. (d) Time-averaged intensity (I) distribution in the xz -plane of the waveguide at $y = h/2 = 170\text{nm}$.

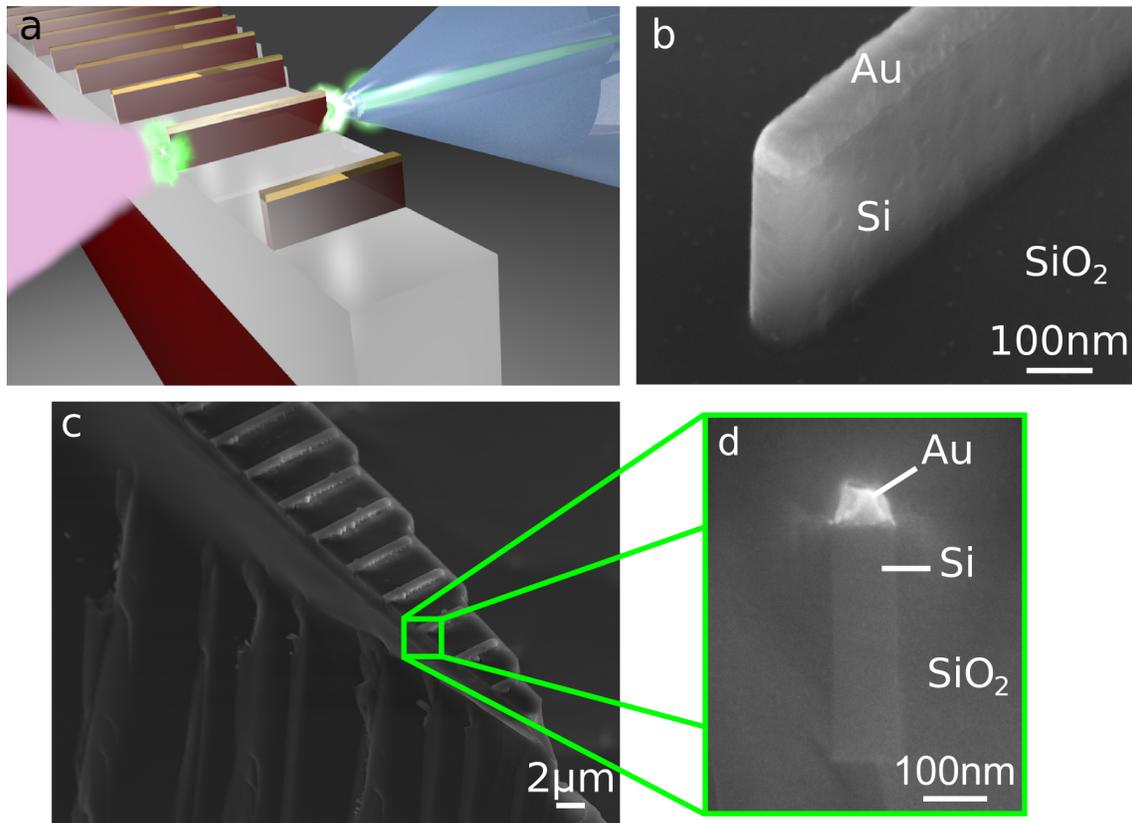


Figure 6.7: (a) Schematic depiction of the conceived device, excitation scheme, and observed visible emission. (b) Scanning electron micrograph of a single nanoplasmonic waveguide before deposition of SiO₂ cladding. (c) Scanning electron micrograph of several nanoplasmonic waveguides fabricated on a narrow beam. A close-up of the cross-sectional geometry of the waveguide end-facet is shown in the inset. (d) Optical microscope image of visible emission from above the nanoplasmonic waveguides.

6.3.1 Free-space Coupling to the Plasmonic Mode

First, it is essential to investigate the prospect of exciting plasmons in the Si-based nanoplasmonic waveguide using tightly focused free-space radiation at $\lambda = 1550\text{nm}$. Assuming validity of the paraxial approximation, the diffraction-limited spot size achievable with a microscope objective with $NA = 0.85$, is determined by the following formula:

$$2w_0 = \frac{2M^2\lambda}{\pi NA} = 1.16\mu\text{m}, \quad (6.1)$$

where NA is the numerical aperture of the microscope objective, w_0 is the spot radius, and M^2 is the beam quality factor (a measure of how closely the transverse beam intensity distribution matches the TEM_{00} mode), which is assumed to be $M^2 = 1$. However, due to the tight focusing of the $NA = 0.85$ microscope objective, inaccuracies can be introduced through the paraxial approximation, warranting a more detailed investigation.

High- NA Focusing

Finite-difference time-domain simulations are performed to ensure the most accurate representation of the tightly-focused fields incident on the input facet of the nanoplasmonic waveguides. The intensity distribution is obtained by considering a laser beam with a Gaussian transverse intensity distribution that is incident on a thin lens with a diameter, $d = 100\mu\text{m}$, which is much greater than the laser radiation wavelength. The effect of the lens is to cause refraction of the incident beam. Once refracted, the beam can be represented as a sum of plane waves all propagating at different angles and converging towards the beam centre at the focal plane. Each plane wave is proportional to $\exp(i\mathbf{k} \cdot \mathbf{r})$ and the \mathbf{k} -vector can easily be related to the angle of propagation of each plane wave. Therefore, the influence of a high- NA lens on an incident Gaussian beam can be determined by integrating all of these plane waves in the plane of the thin lens. The resulting fields are then propagated through the simulation space using the FDTD method and the full-vectorial distribution of electromagnetic fields is recorded along the optical axis and in the transverse plane at the beam focus. These simulations are performed using a commercially-available software package (Lumerical FDTD Solutions).

The intensity distribution in a plane along the optical axis is shown on a logarithmic scale in Fig. 6.8, confirming that the beam is focused to a compact spot. The E_x , E_y , and E_z field components at the focus of the lens are shown in Figs. 6.9(a)-(c), respectively, and the intensity distribution is shown in Fig. 6.9(d). A transverse slice of the intensity distribution at the focus is shown on a linear scale in Fig. 6.10 along with the Gaussian function with a radius, $w_0 = 580\text{nm}$. Based these plots, some minor deviations from the paraxial approximation are evident.

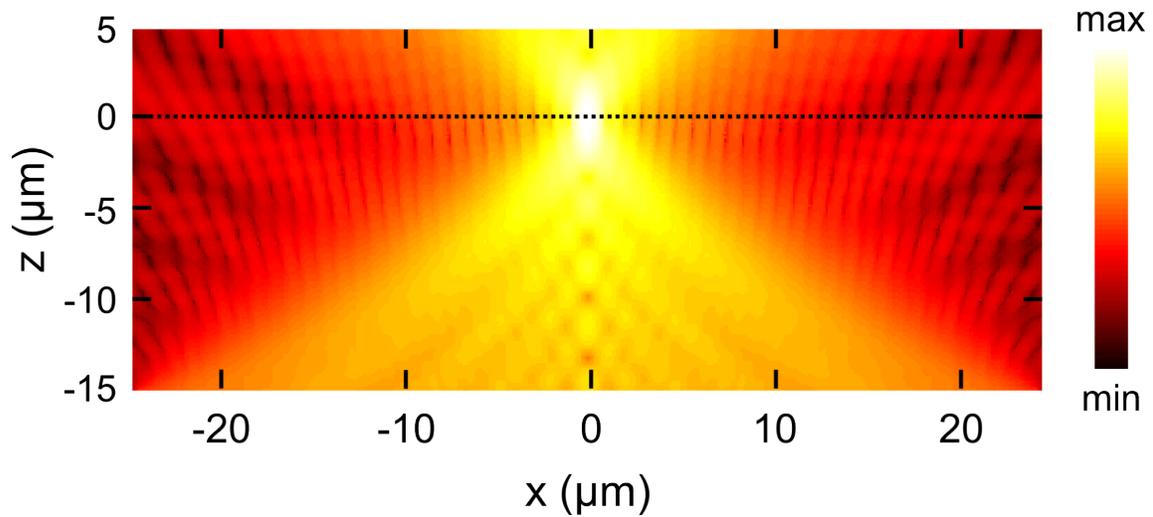


Figure 6.8: Intensity (I) distribution of focusing by a $NA=0.85$ microscope objective, shown on a logarithmic scale.

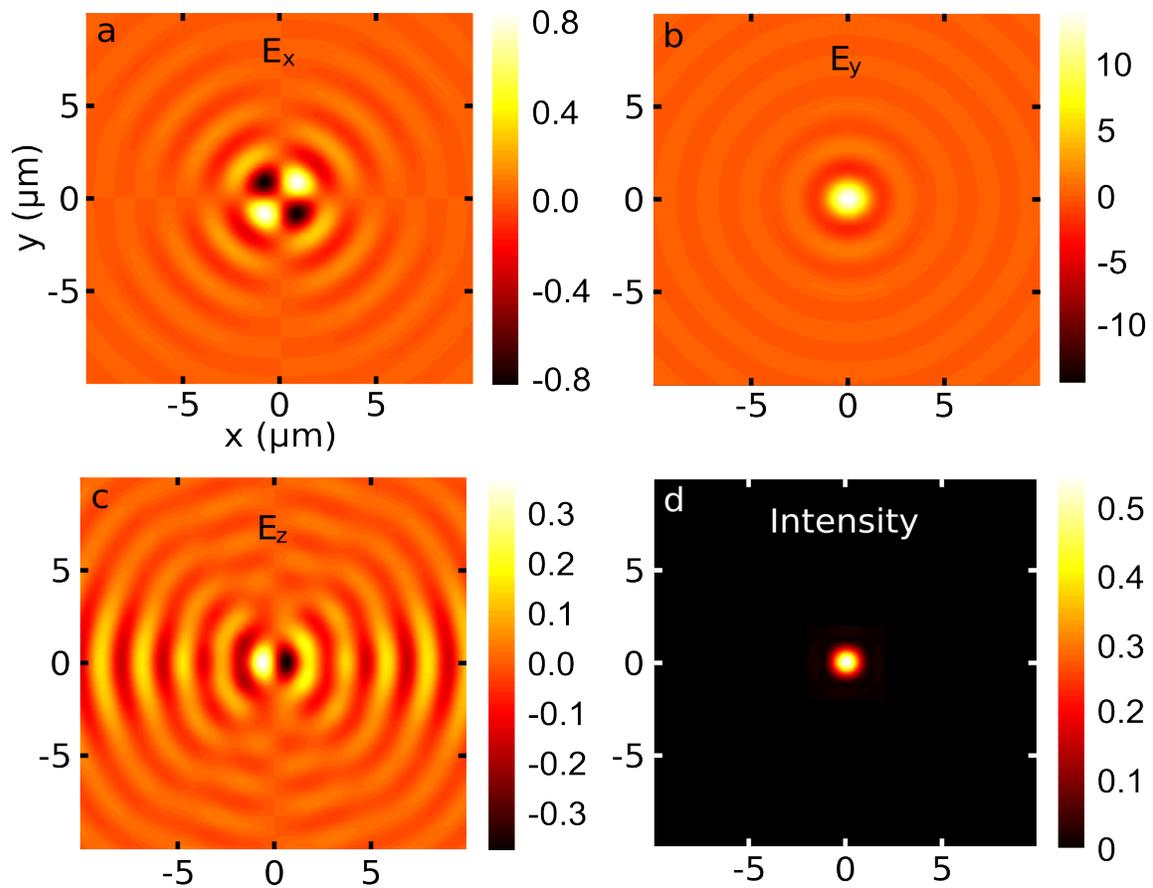


Figure 6.9: Electric field distributions at the focal plane of a $NA = 0.85$ microscope objective for: (a) E_x , (b) E_y , and (c) E_z . The intensity (I) distribution at the focal plane is shown in (d).

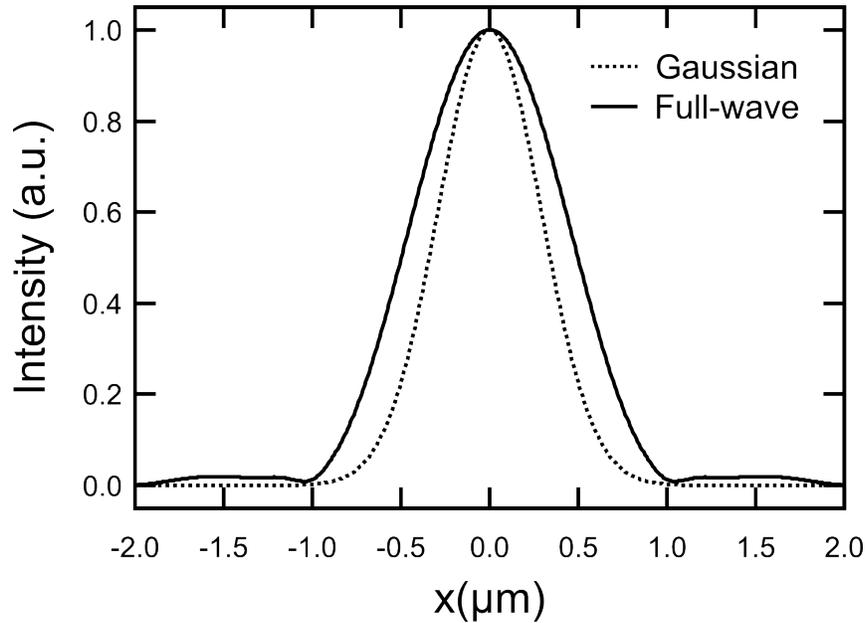


Figure 6.10: Deviation of the focal spot from the ideal Gaussian distribution with $w_0 = 580\text{nm}$ at the focus of a $NA = 0.85$ microscope objective.

Nanoplasmonic Waveguide Excitation

The same high- NA modelling is then used to excite the input facet of a nanoplasmonic waveguide with $w = 100\text{nm}$, $h = 340\text{nm}$, and $L = 9.0\mu\text{m}$. A y -polarized Gaussian beam is focused on the input facet of the nanoplasmonic waveguide using fields obtained from the $NA = 0.85$ lens. A time snapshot of the E_y component of the electric field incident on the structure and coupling to the nanoplasmonic waveguide is shown in Fig. 6.11(a). The excitation of the nanoplasmonic mode occurs via diffraction at the edge of the Au cap of the nanoplasmonic waveguide, and the fields at the input facet of the waveguide are enhanced by a factor of 2.4 relative to the free-space excitation. The mode propagates through the Si core and decays with a length scale of $L_{prop} = 3.1\mu\text{m}$, matching the value obtained from the mode solver. Due to the sub-diffraction dimensions of the input facet of the nanoplasmonic waveguide, not all of the radiation can be focused onto the facet and some scatters around the waveguide, coupling onto the Au-air interface and Si-SiO₂ interface. It should be stressed that the nanoplasmonic mode excited in the Si core is the only true waveguide mode excited, and the other radiation decays much more rapidly than the nanoplasmonic mode. Figure 6.11(b) shows the E_y electric field component in the $y = 0.33\mu\text{m}$ plane, directly below the Au-Si interface. This demonstrates the tight confinement of fields to the Si core, as well as the field enhancement relative to the excitation field and the scattered radiation surrounding the waveguide. As shown in Fig. 6.11(c), after propagating $2\mu\text{m}$ through the nanoplasmonic waveguide, the power flow is tightly confined to the Si-Au interface, and this is where the dominant nonlinear interactions will occur.

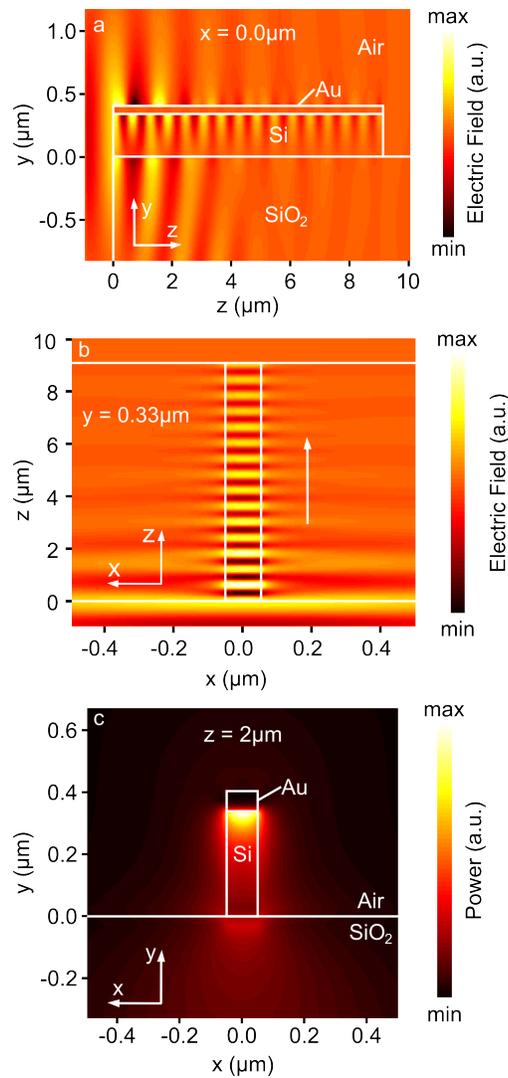


Figure 6.11: Simulation of the excitation of the silicon-based nanoplasmonic waveguide with a $NA = 0.85$ microscope objective. (a) Time snapshot of focused free-space electric fields (E_y) incident on the input facet of the nanoplasmonic waveguide and coupling to a tightly confined mode at the Si-Au interface. Fields also couple to poorly-defined surface modes on the air-Au and Si- SiO_2 interfaces, but become negligible within $\approx 2 \mu\text{m}$. (b) Time snapshot of E_y taken in the $y = 330 \text{nm}$ plane, demonstrating the coupling and enhancement of focused incident radiation to a plasmonic mode at the Au-Si interface. (c) Intensity (I) distribution of the excited plasmonic mode profile in the $z = 2 \mu\text{m}$ plane.

6.3.2 Estimation of Electric Field Strength

In these experiments, the free-space laser power incident on the microscope objective is varied with a continuously variable neutral density filter wheel. The time-averaged incident power is varied in the range $2.0 \leq P_{fs}^\omega \leq 55.0\text{mW}$. Assuming a sech^2 temporal profile of the laser pulses, the peak power is given by:

$$P_{peak,fs}^\omega \approx 0.88 \frac{P_{fs}^\omega}{f_{rep}\tau_p}, \quad (6.2)$$

where $f_{rep} = 90\text{MHz}$ is the repetition rate of the laser and $\tau_p = 84\text{fs}$ is the pulse duration. Using these values, the peak power is varied in the range, $0.23 \leq P_{peak,fs}^\omega \leq 6.40\text{kW}$.

Due to the deviation from a perfect Gaussian at the focus of the $NA = 0.85$ microscope objective, the intensity at the focus must be calculated from the full-field simulations. The total power, P_0 , associated with the intensity distribution, $I(x, y)$, shown in Fig. 6.9(d), is calculated via the following equation:

$$P_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) dx dy. \quad (6.3)$$

By multiplying the intensity distribution, $I(x, y)$, of Fig. 6.9(d) by a factor of P_{peak}^ω/P_0 , the intensity distribution at the focus of the lens for a pulse with P_{peak}^ω is obtained. Using this expression, the spatial intensity distribution at the focus of the lens is calculated for the range of powers used in the experiments, $0.23 \leq P_{peak,fs}^\omega \leq 6.40\text{kW}$. Doing so reveals that the peak excitation intensity incident on the nanoplasmonic waveguides is varied in the range, $1.91 \times 10^{14} \leq I_{peak,fs}^\omega \leq 5.24 \times 10^{15}\text{W/m}^2$. The local intensity can then be used to determine the electric field amplitude by the following equation:

$$I_{peak,fs}^\omega = \frac{c\epsilon_0}{2} |E_{peak,fs}^\omega|^2 \quad (6.4)$$

From this, the peak electric field amplitude incident on the input facet of the nanoplasmonic waveguides is estimated to be varied in the range, $3.79 \times 10^8 \leq |E_{peak,fs}^\omega| \leq 1.99 \times 10^9\text{V/m}$.

A final important factor to consider is the electric field enhancement provided by the highly confined nanoplasmonic mode. From the FDTD simulations performed in the previous section, it is estimated that the electric fields are enhanced by a factor of 2.4 relative to the fields incident on the input facet of the nanoplasmonic waveguide. Based on this, the enhanced electric field strength propagating through the silicon core is estimated to be varied in the range, $9.10 \times 10^8 \leq |E_{SP}| \leq 4.77 \times 10^9 (\pm 2.7 \times 10^8)\text{V/m}$.

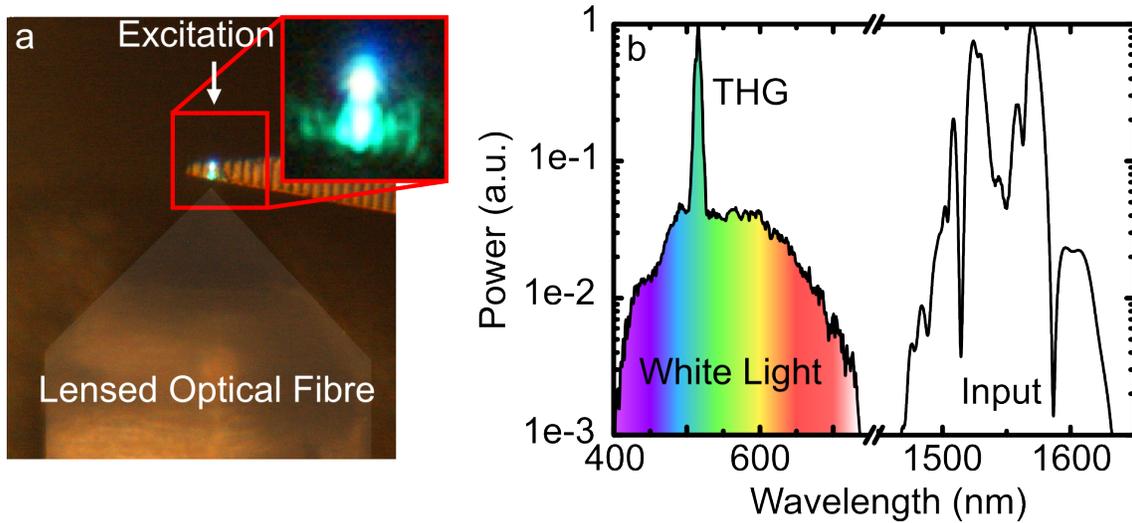


Figure 6.12: Spectra of the ultrafast laser (Input), along with measured spectra of visible light collected from the end-facet of the silicon-based nanoplasmonic waveguide. The visible light consists of a strong THG ($500\text{nm} \leq \lambda \leq 530\text{nm}$) peak and a broad white light spectrum ($375\text{nm} \leq \lambda \leq 725\text{nm}$).

6.4 Third-Harmonic Generation

Free-space radiation from a laser emitting $\tau_p \sim 84\text{fs}$ pulses centred at $\lambda = 1550\text{nm}$ is end-fire coupled into the waveguides using a $NA = 0.85$ microscope objective. For peak input powers above $P_{peak}^\omega = 9.9\text{W}$, strong visible light emission is observed from the waveguide end facets under normal room lighting. Scattered visible emission is collected from above the sample by a $20\times$ microscope objective and delivered to a charge-coupled device camera, as shown in Fig. 6.12(a). Bright white light emission with a strong blue component is visible at the input facet of the nanoplasmonic waveguide and distinct THG (green light) is observed only at the output facet. The light emission was present in every nanoplasmonic waveguide that was tested regardless of the waveguide length and was distinctly different than the visible light emission that was observed in the Si photonic waveguides (i.e. only THG, Chapter 5). Notably, some of these waveguides were longer than the propagation length of both the fundamental and the third-harmonic (3ω) radiation. The $\lambda = 1550\text{nm}$ laser pulses propagate through the nanoplasmonic waveguide with an effective wavelength, $\lambda_{eff}^\omega = 586\text{nm}$ and decay with a characteristic length, $L_{prop} = 3.1\mu\text{m}$, while the 3ω radiation ($500\text{-}530\text{nm}$) propagates with $\lambda_{eff}^{3\omega} = 170\text{nm}$ and decays over a length of 640nm ($1.1\lambda_{eff}^\omega$). The nonlinear interaction is essentially localized, as the 3ω radiation is absorbed as it interacts with the fundamental radiation on a scale of approximately one wavelength and the typical conditions of phase matching were not required to observe THG.

The distinct difference in the spectral components of emitted light from the input and output facets of the waveguide (which are $4.5\mu\text{m}$ apart) provides clear evidence that the observed emission arises from the Si core of the nanoplasmonic waveguide and is

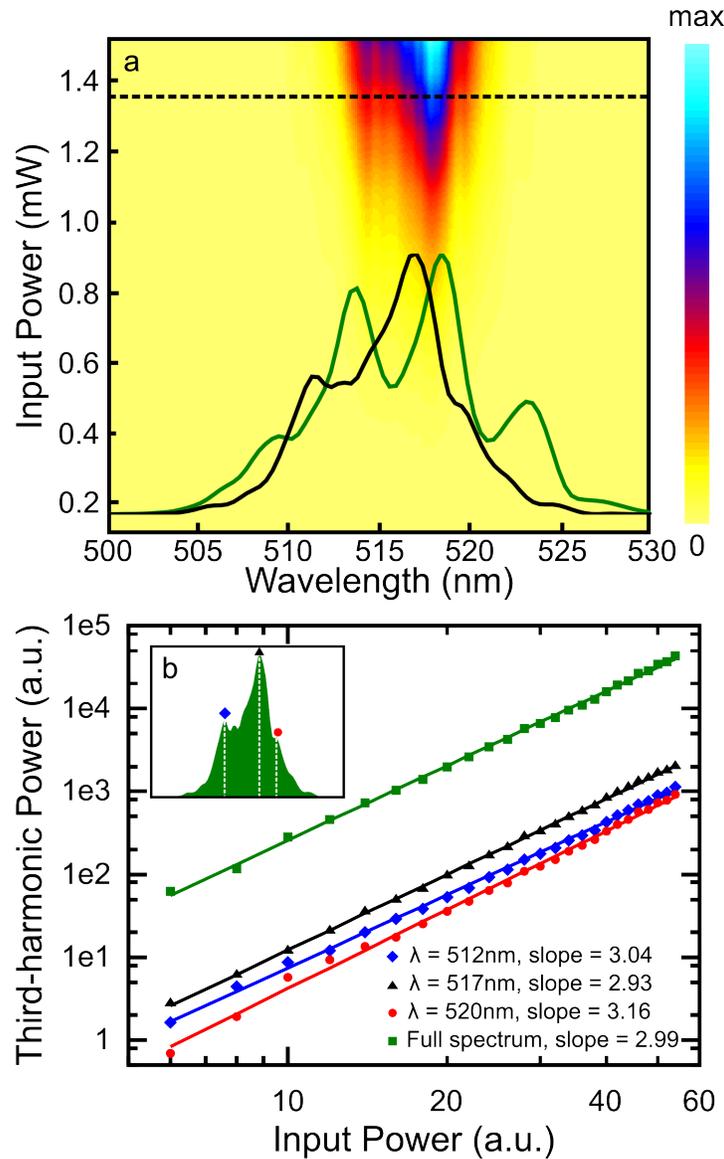


Figure 6.13: (a) Surface plot of third-harmonic power as a function of wavelength and input power measured at the output facet of the nanoplasmonic waveguide. The black overlaid spectrum is a 3ω spectrum measured at an average input power, $P_{in}^\omega = 1.36$ mW, while the green spectrum is the 3ω spectrum calculated from the laser pulse spectrum. (b) Plots on logarithmic x - and y -axes of detected 3ω power versus input power for the entire 3ω spectrum (green) and individual wavelength, $\lambda = \{512, 517, 520\}$ nm.

not simply excited at the input facet, where the intensity is highest, and scattered into the output fibre. Instead, the white light emission is stronger than the 3ω at the input facet. As the $\lambda = 1550\text{nm}$ pulses propagate through the device, the nonlinear interaction evolves and as they reach the output facet of the waveguide, the 3ω emission is stronger than the white light emission, as indicated by the images and out-coupled spectra. This distinct evolution in the nonlinear interaction can only occur through a guided mode in the nonlinear Si core of the waveguide and would not be observed from scattered light from the input facet of the waveguide.

The nature of the visible light emission is investigated by coupling the radiation emitted from the output facet of the nanoplasmonic waveguide to a lensed SMF, and delivering it to a spectrometer. Finite-difference time-domain simulations determine that $\sim 4.1\%$ of the visible light generated in the nanoplasmonic waveguide is coupled into the lensed optical fibre. The laser spectrum is shown together with the visible emission in Fig. 6.12. The visible emission consists of a strong 3ω signal that spans $500\text{nm} \leq \lambda \leq 530\text{nm}$ and a broad white light background spanning $375\text{nm} \leq \lambda \leq 725\text{nm}$. The broad 3ω spectrum is further evidence that phase matching does not play a role in the THG.

A surface depicting the spectral power of the TH, $P_{out}^{3\omega}$, as a function of the average input laser power, P_{in}^{ω} , is shown in Fig. 6.13(a). The overlaid spectra depict a measured 3ω spectrum for $P_{in}^{\omega} = 1.36\text{mW}$, (dashed line in the figure), and a calculated 3ω spectrum using the measured fundamental pulse spectrum (discussed in Appendix C). Regardless of P_{in}^{ω} , the 3ω spectrum spans $500\text{nm} \leq \lambda \leq 530\text{nm}$, and the peak wavelength is located at $\lambda = 517\text{nm}$. The total 3ω power is obtained by integrating the 3ω spectrum and a log-log plot of P_{in}^{ω} versus $P_{out}^{3\omega}$ is found to scale with a slope of 2.99 ± 0.02 , as shown in Fig. 6.13(b). Similarly, the slopes for specific wavelengths in the 3ω spectrum, $\lambda = \{512, 517, 520\}\text{nm}$, are found to be 3.04 ± 0.01 , 2.93 ± 0.01 , and 3.16 ± 0.04 , respectively. The $P_{out}^{3\omega} \propto (P_{in}^{\omega})^3$ dependence confirms that the green light originates from a $\chi^{(3)}(-3\omega; \omega, \omega, \omega)$ nonlinear optical interaction and demonstrates that the entire 3ω spectrum grows uniformly with increasing input power [28]. To date, there has been only a single observation of THG in a SOI device from a large ($5\mu\text{m}$ -width \times $80\mu\text{m}$ -length) Si photonic crystal waveguide, where slow light propagation enhancement and tight modal confinement allowed for a THG conversion efficiency, $\eta_{THG} \approx 1 \times 10^{-7}$ [21]. A calibration is performed on the spectrometer (discussed in Appendix E), allowing the absolute 3ω output power and conversion efficiency to be calculated. Remarkably, the THG conversion efficiency from the nanoplasmonic structure is $\eta_{THG} = 2.29 \times 10^{-5}$, but in an ultracompact footprint ($95\text{nm} \times 4.5\mu\text{m}$) that is reduced by a factor of ~ 1400 with respect to the Si PhC [21]. For this nanoplasmonic structure, large electric field enhancement and tight spatial confinement are achieved at the Si-Au interface, which provide the necessary conditions for observing high efficiency THG.

Figure 6.14 depicts the measured conversion efficiency as a function of P_{peak}^{ω} on logarithmic x - and y -axes. A linear fit to this data reveals a slope of 1.92 ± 0.02 , which is

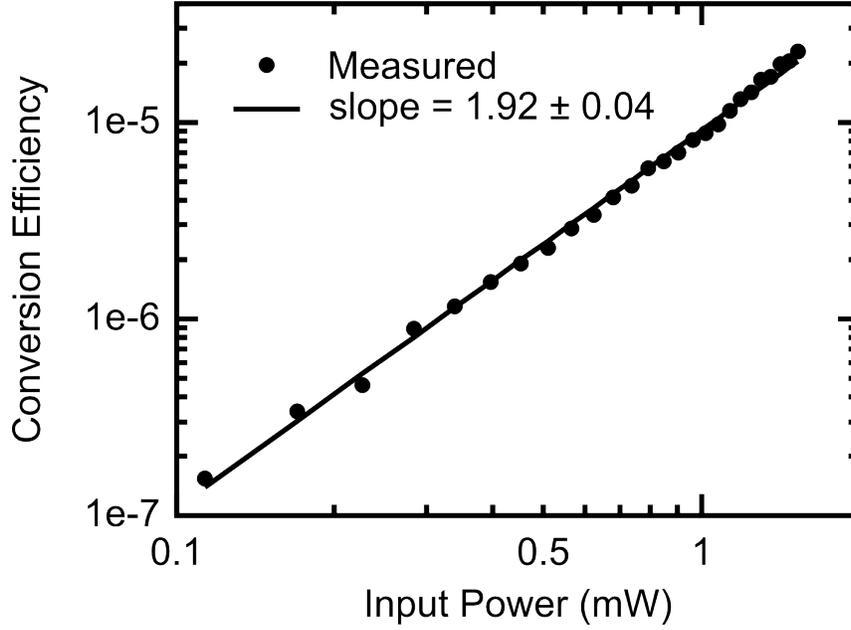


Figure 6.14: Conversion efficiency as a function input power, plotted on logarithmic x - and y -axes. A linear fit to the data points reveals a slope of 1.92, which is very close to the expected value of 2.

close to the expected value of 2. It is also important to note the normalized conversion efficiency, i.e. the conversion efficiency at a peak input power of $P_{peak}^\omega = 1\text{W}$, which is calculated via:

$$\eta_{norm} = \frac{\eta}{(P_{peak}^\omega)^2} \quad (6.5)$$

Based on a statistical analysis of the conversion efficiencies measured at each P_{peak}^ω , a normalized conversion efficiency of $\eta_{norm} = (6.96 \pm 0.69) \times 10^{-10}\text{W}^{-2}$ is determined.

6.5 White Light Emission

It is important to elucidate the nature of the bright white light emission since its generation process is key in revealing the ultrafast carrier dynamics and their interaction with the highly-confined plasmonic field, E_{SP} . A surface plot showing the white light emission spectrum for varying P_{in}^ω is displayed in 6.15(a). The overlay depicts the white light spectrum obtained for $P_{peak}^\omega = 165 \pm 3\text{W}$ measured at the output facet of the nanoplasmonic waveguide, where the spectral contributions due to THG have been removed for clarity. The white light emission spans the range $375\text{nm} \leq \lambda \leq 650\text{nm}$ ($1.9\text{eV} \leq h\nu \leq 3.3\text{eV}$) and has a peak that lies within the THG spectrum ($500\text{nm} \leq \lambda \leq 530\text{nm}$). Depending on the excitation alignment, the spectrum was observed to extend up to $\lambda = 725\text{nm}$. By rotating the incident polarization, it was observed that the white light emission arises from coupling the TM- E_{SP} mode into the nanoplasmonic waveguide. This and the absence of distinct atomic emission lines, increased scaling of the throughput light power with P_{in}^ω ,

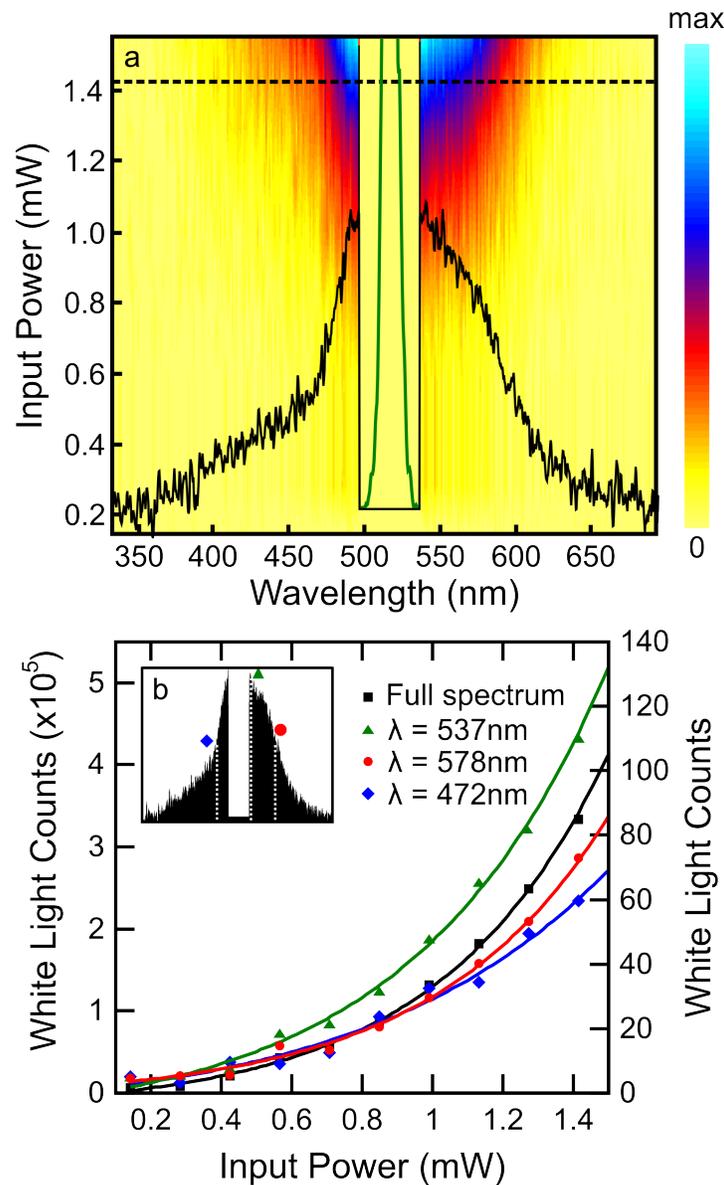


Figure 6.15: (a) Surface plot of white light photon counts measured at the output facet of the nanoplasmonic waveguide as a function of wavelength and power. The black overlaid spectrum is measured at an average input power, $P_{in}^{\omega} = 1.42$ mW, and the green curves represent the position of the 3ω components, which have been removed for clarity. (b) Plots of experimentally measured white light counts versus input power for the entire spectrum (black) and individual wavelengths, $\lambda = \{472, 537, 578\}$ nm.

and lack of damage at the input facet, preclude the presence of polarization-independent surface breakdown.

The nonlinear scaling of the white light emission is studied by out-coupling the emission from the output facet of the nanoplasmonic waveguide and delivering it to a spectrometer. By varying the input power and collecting a white light spectrum at each power, a nonlinear scaling plot of the white light power versus the excitation power is generated. Figure 6.15(b) displays the total white light power, P_{wh} , growth as a function of P_{in}^ω , along with an exponential growth fit, i.e. $P_{wh} \propto \exp(\gamma P_{in}^\omega)$, where γ is the power growth coefficient ($\sim 2.38 \pm 0.15 \text{mW}^{-1}$). The exponential power-dependence on P_{in}^ω is observed at all wavelengths within the spectrum, and the scaling trends for $\lambda = \{472, 537, 578\} \text{nm}$ are shown as well.

Notably, most light emission mechanisms follow an integer power scaling law, i.e. $P_{emis} \propto (P_{in}^\omega)^n$, where n is an integer. Further analysis confirms the exponential growth of the white light power. Figure 6.16(a) displays the P_{wh} versus P_{in}^ω on linear axes, along with an exponential growth fit. However, more insight into nonlinear interactions is always gained by plotting the data on logarithmic x - and y -axes. Therefore, the data from Fig. 6.16(a) is plotted on logarithmic x - and y -axes, as shown in Fig. 6.16(b). Careful examination reveals that the data points do not follow a straight line with a constant slope. Three linear fits are performed to different sets of data points, and the slope is observed to increase from 0.13 to 2.35. A consistently increasing slope would be expected of exponential growth. Notably, white light luminescence from the gold film would show a fourth-order dependency on P_{in}^ω [29]. Finally, the data points are plotted on a linear x -axis scale and natural logarithmic y -axis scale, which is shown in Fig. 6.16(c). In this case, the data points follow a curve that is distinctly linear. This analysis serves as conclusive evidence that the underlying physical mechanism of the white light generation gives rise to an exponential growth of the white light power for increasing excitation power.

6.6 Ponderomotive Electron Acceleration

As the femtosecond laser pulses propagate through the Si nanoplasmonic waveguides, electrons are excited from the valence band (VB) into the indirect bandgap ($E_g = 1.1 \text{eV}$) at the conduction band (CB) X-valley via TPA. The average excess energy per electron is $E_{ex} \sim 0.5 \text{eV}$ and the electron energy distribution lies near the bottom of the indirect CB edge. Since the energy of the white light emission spectrum ($1.9 \text{eV} \leq h\nu \leq 3.3 \text{eV}$) is much higher than the two-photon energy, the excited electrons must gain additional energy and climb high into the CB directly after they are excited. However, electric fields acting upon the electrons supply additional energy and allow for radiative transitions at optical frequencies. Interestingly, a strong ($> 10^9 \text{V/m}$), quasi-DC electric field can be inherently provided by the highly spatially-inhomogeneous, oscillating plasmonic field. Once electrons are placed in the CB, they will interact with the strong evanescent (non-propagating,

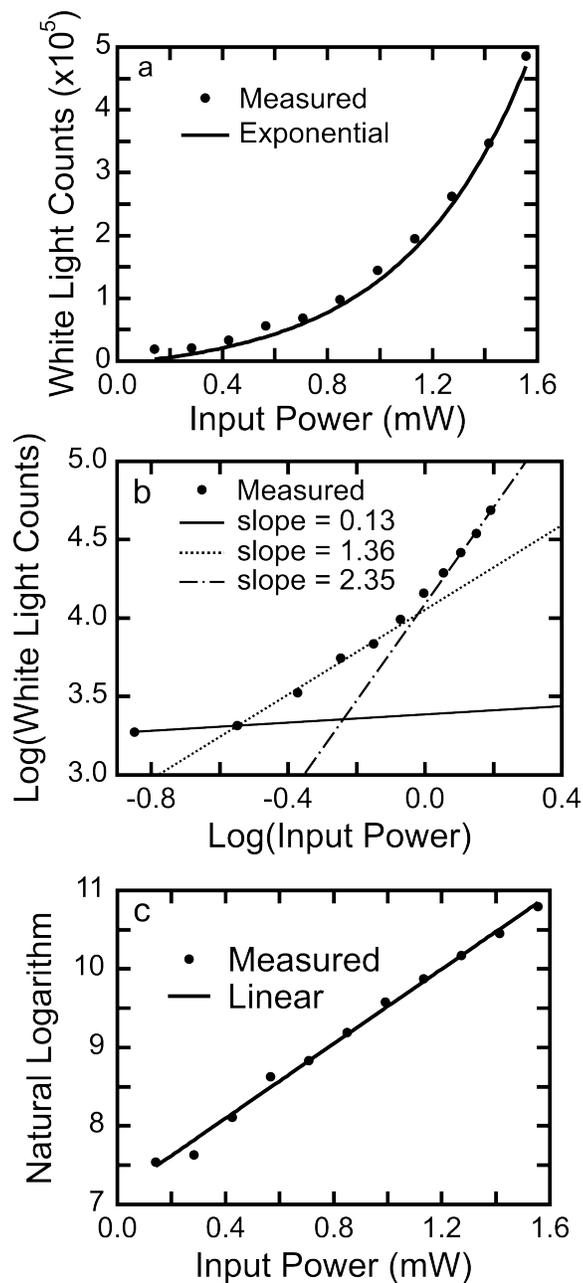


Figure 6.16: Total white light counts versus input power, shown on various axes scales: (a) Linear x - and y -axes, along with an exponential growth fit. (b) Logarithmic x - and y -axes, with various linear fits, revealing an increasing slope with increasing input power. (c) Linear x -axis and natural-logarithmic y -axis, along with a linear fit.

near-field) E_{SP} field in the Si core of the nanoplasmonic waveguide. The nonlinear E_{SP} -electron interaction in the presence of a high E_{SP} field gradient allows significant electron energy gain via ponderomotive (or cycle-by-cycle) acceleration (described in Section 2.6) [30–32]. The basic principle of ponderomotive acceleration for an electron in an E_{SP} field is illustrated in Fig. 6.17. The high E_{SP} propagating along the Au-Si interface is spatially inhomogeneous, rapidly decaying over a length scale of $\alpha_0 = 135\text{nm}$ (sub-wavelength) in the Si core. As the E_{SP} electric field oscillates over one cycle, it reverses its direction and a TPA-excited electron in the CB experiences acceleration and deceleration during the positive and negative half cycles, respectively. If the E_{SP} were spatially uniform, the electron would merely experience symmetric-but-opposing forces that would cancel over an optical cycle and no net kinetic energy would be gained. However, in a situation where the E_{SP} field decays exponentially with an extreme spatial gradient, an electron with an initial kinetic energy, E_{ex} , that starts its motion at the Au-Si interface where the E_{SP} field magnitude is highest, experiences asymmetric forces within the decay length ($\sim 135\text{nm}$) of the E_{SP} electric field. As the electron gains energy during the acceleration phase (positive cycle), it traverses into a region of a weaker E_{SP} electric field before the electric field reverses during the negative cycle. As such, the decelerating force is weaker during the deceleration phase. Effectively, the electron is “pushed” by the electric force in the direction of decreasing field amplitude by an amount that is larger than the subsequent cycle in which it is “pulled” back. During successive rapid oscillations of the E_{SP} field, the electron is pushed away from the interface in a cycle-by-cycle fashion, and the electron is imparted with a net velocity, $\mathbf{v}(t)$, along the direction of decreasing electric field amplitude as illustrated in Fig. 6.17. With this ponderomotive acceleration-deceleration being repeated over several oscillation cycles, a low-frequency (quasi-DC) ponderomotive force, F_{pond} , arises in the direction normal to the Au-Si interface and imparts a net kinetic energy gain to the electron.

6.6.1 Modelling of Ponderomotive Acceleration

The motion of electrons that are generated by two-photon absorption (TPA) can be considered in three separate stages. Immediately after the electrons are generated via TPA, they are imparted with a kinetic energy that is determined by the difference between the energy of the two photons and the indirect bandgap of silicon, $KE = 2h\nu - E_g = 0.5\text{eV}$. The electrons then interact with the intense E_{SP} field propagating through the waveguide and undergo rapid acceleration via the ponderomotive force. After interacting with the E_{SP} field, the electrons have acquired a new velocity vector, which in general, will have a different amplitude and direction than their original velocity.

It is important to note that the period of the carrier frequency of the laser pulses,

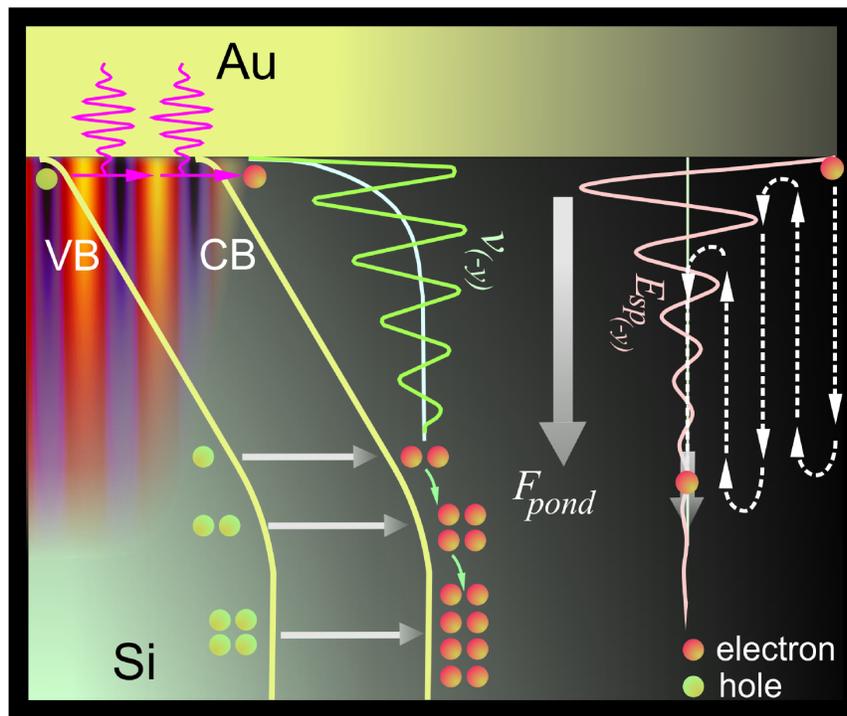


Figure 6.17: Schematic depiction of physical processes relevant to the electron acceleration, including intense, asymmetric E_{SP} field propagation, TPA, cycle-by-cycle acceleration of the electron in the E_{SP} field leading to a net ponderomotive force, and electron and hole avalanche multiplication.

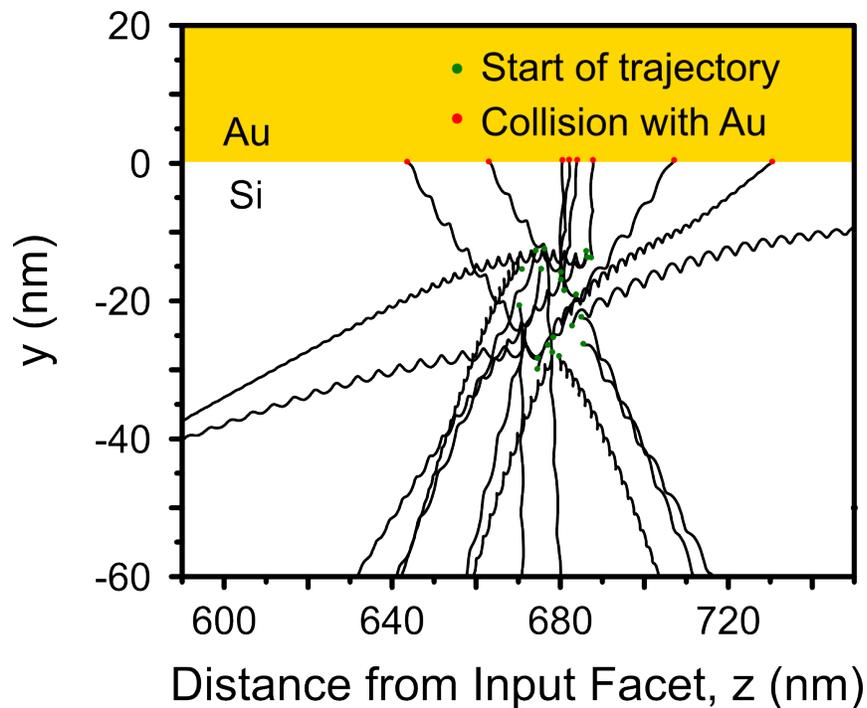


Figure 6.18: Sample calculated trajectories of 21 electrons excited via TPA at random points in time, with random positions and initial velocities. The green dots represent the starting point of the electron trajectory and the red dots denote the location of an electron colliding with the gold surface.

$T_p = 5.20\text{fs}$, is much less than the electron collision time in silicon,

$$\tau_c = \mu_e m_e^* / e = 214\text{fs}, \quad (6.6)$$

where $\mu_e = 0.145\text{m}^2/\text{V}\cdot\text{s}$, is the electron mobility and $m_e^* = 2.54 \times 10^{-31}\text{kg}$ is the electron effective mass [33]. The mean free path for electron collisions can be estimated from:

$$l_{avg} = v_s \tau_c = 28\text{nm}, \quad (6.7)$$

where $v_s = 1.3 \times 10^5\text{m/s}$ is the electron saturation velocity in silicon [34]. Thus, during the first $\sim 28\text{nm}$ of travel, electrons do not suffer any collisions and move in the Si core as ballistic electrons. That is, while interacting with the ultrafast E_{SP} field, the electrons do not follow the same behaviour as electrons drifting through Si in response to a slowly-varying electric field. In particular, they do not undergo a sufficient number of collisions for their motion to be accurately represented by a drift velocity, which can saturate for sufficiently high electric field. Instead, their motion can be considered as ballistic during the early acceleration phase and their instantaneous velocity will momentarily exceed the saturation velocity. As outlined in the previous section, the ponderomotive force acts in a quasi-DC manner on the electron. Therefore, once the electron has finished interacting with the E_{SP} field and undergoes multiple collisions, it will move with an effective drift velocity, which is limited by the saturation velocity in silicon.

A rigorous treatment of the time-dynamics of such a problem would require solution of the nonequilibrium Boltzmann transport equation (BTE), utilizing fully vectorial electromagnetic fields calculated in the nanoplasmonic waveguides and calculated silicon band structure data. Here, we are particularly interested in estimating the energies attained by electrons before they undergo their first collision, i.e. within their first $\sim 28\text{nm}$ of motion. It is during this length scale of motion that electrons acquire their initial energy, which can be enough to initiate impact ionization at their first collision, depending on the magnitude, direction, and phase of the E_{SP} field at their initial position and the direction of their initial velocity. This permits an alternative and accurate approach that does not require solution of the BTE, and allows usage of techniques similar to those used to study laser ionization and subsequent electron motion in the intense laser fields [35].

The electron acceleration is calculated in two steps:

1. First, the time-varying fields of a $\tau_p \sim 84\text{fs}$ pulse propagating through the nanoplasmonic waveguide are obtained via FDTD simulations performed using Lumerical FDTD Solutions. The cross-sectional dimensions of the waveguide are taken to be the same as the fabricated devices ($w \times h = 95 \times 340\text{nm}$) and the waveguide length is taken to be $L = 5\mu\text{m}$.
2. Next, the electron trajectories and energies are calculated from the time-varying

electromagnetic fields, according to the following equation:

$$v_f(t_i) = v(t_i) + \int_{t_1}^{\infty} \frac{eE(t)}{m_e^*} dt, \quad (6.8)$$

where $v(t_i)$ is the initial electron velocity, $v_f(t_i)$ is the final electron velocity, e is the elementary charge, m_e^* is the electron mass, and $E(t)$ is the instantaneous electric field experienced by the electron.

Approximately 1.4×10^5 electrons are released at random locations in the waveguide during the simulation. Each electron is given an initial kinetic energy, $KE = 0.50\text{eV}$, with its velocity in a random direction. The time-dependent trajectory of each electron is calculated via Eq. 6.8, and the position and velocity of the electron is recorded throughout the interaction with the pulse. In addition, the electric field experienced by the electron as it undergoes its motion is recorded. Electrons that collide with the waveguide sidewalls or Au film are no longer considered in the simulation, as they can recombine at the Au surface or be emitted into the SiO_2 (depending on their energy). Twenty-one sample electron trajectories in response to $E_{SP} = 3 \times 10^9\text{V/m}$ field are shown in Fig. 6.18. Each of these electrons is introduced to the simulation at a different point in time, and with a different starting position and velocity direction (the velocity amplitude was kept constant at 0.5eV). The trajectory of each electron is significantly different, demonstrating the sensitivity of the final energy of the electron to its initial conditions. Of the 21 sample trajectories, 8 electrons collide with the Au film before traveling a distance of $\sim 28\text{nm}$.

Figure 6.19(a) shows the electric field experienced by an electron that is produced at the temporal peak of the laser pulse as it interacts with the E_{SP} field for three different field strengths lying in the range used in the experiments, $E_{SP} = \{1 \times 10^9, 3 \times 10^9, 5 \times 10^9\}\text{V/m}$. Similarly, Fig. 6.19(b) shows the instantaneous velocity reached by the electron as it interacts with the E_{SP} field. It is important to note that these calculations are only valid for the first $\sim 28\text{nm}$ of the electron's trajectory, before it undergoes its first collision. Depending on the E_{SP} strength, this may be only a few oscillations of the E_{SP} field. Before the first collision, the average velocity of the electron can exceed v_s by over an order of magnitude.

The electron energy distribution is generated by calculating the average velocity attained by each electron during the first $\sim 28\text{nm}$ of its trajectory. Each electron is given a weight depending on the instantaneous intensity of electric fields at the location of the nanoplasmonic waveguide that it was injected at, to account for the spatially-varying number density of TPA-generated electrons in the nanoplasmonic waveguide. The number of electrons excited in the waveguide by the E_{SP} field is given by [9]:

$$\frac{dN(\mathbf{r}, t)}{dt} = \frac{\beta_{TPA}}{2h\nu} I_{SP}^2(\mathbf{r}, t) \quad (6.9)$$

where β_{TPA} is the two photon absorption coefficient in silicon, $N(\mathbf{r}, t)$ is the instantaneous,

local density of electrons, and $I_{SP}^2(\mathbf{r}, t)$ is the instantaneous, local plasmonic electric field intensity. Therefore, each electron is given a weighting factor proportional to the square of the electric field intensity at the location that it was injected at. Based on the calculated average velocities, the electrons are separated into energy bins with widths, $\Delta E = 0.01\text{eV}$, and a spectrum of the average electron energies before the first collision is obtained.

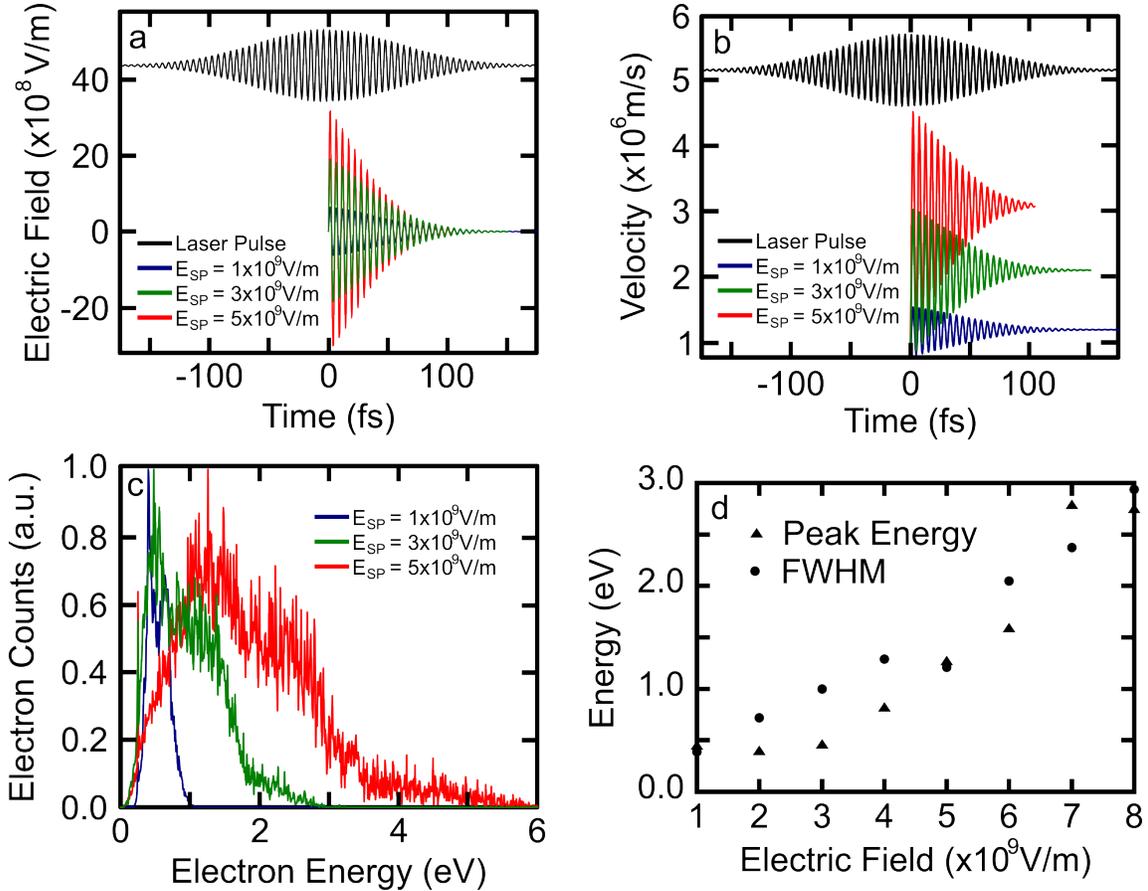


Figure 6.19: (a) Electric field experienced by an electron during the course of its trajectory in response to the E_{SP} field, shown for three different E_{SP} field strengths. (b) Instantaneous velocity reached by the electron in response to the E_{SP} field, shown for three different E_{SP} field strengths. It should be noted that these calculations neglect collisions and should only be taken as valid before the electron undergoes its first collision (the first $\sim 28\text{nm}$ of its trajectory). The red and green curves are terminated when the electron has traversed the entire Si waveguide core and collides with the SiO_2 interface. (c) Calculated electron energy spectra for three different E_{SP} values in the same range as those used in the experiments. The electron energy represents the average energy of the electrons before undergoing their first collision. (d) Plot of the electron energy spectral peak and FWHM versus E_{SP} , taken from smoothed spectra.

Electron energies are calculated from electron trajectories obtained at E_{SP} strengths in the range of $1 \times 10^9 \text{V/m} \leq E_{SP} \leq 8 \times 10^9 \text{V/m}$, and three exemplary spectra for $E_{SP} = \{1 \times 10^9, 3 \times 10^9, 5 \times 10^9\} \text{V/m}$ are shown in Fig. 6.19(c). The peak electron energy and FWHM are calculated from smoothed versions of the spectra, and are presented in Fig. 6.19(d). It is determined that a threshold field of $E_{SP} \sim 2.5 \times 10^9 \text{V/m}$ is

required before any electrons at an energy of $E_t = 2.3\text{eV}$ are present. This is in excellent agreement with the estimated experimental electric fields in the structure, which are in the range $9.10 \times 10^8 \leq |E_{SP}| \leq 4.77 \times 10^9 (\pm 2.7 \times 10^8)\text{V/m}$. For $E_{SP} = 5 \times 10^9\text{V/m}$, the electron energy distribution reaches a maximum value of $\sim 6\text{eV}$. It should be noted that these plots do not represent the thermal energy distribution (Maxwellian distribution) of the free electrons, but rather the excess kinetic energy that the electrons gain from the plasmonic field via ponderomotive acceleration. Furthermore, there is not a simple direct relationship between the calculated electron energy spectrum and the observed white light spectrum. The electron energy spectrum is similar to the ones reported previously for photoelectron emission from metal films [30, 32, 35].

Remarkably, these calculations demonstrate that TPA-excited electrons can be ponderomotively accelerated to velocities exceeding the saturation velocity in Si, and above the threshold energy for impact ionization. Free electrons can undergo collisions with valence electrons while still interacting with the E_{SP} field, initiating an electron avalanche. After the E_{SP} field has passed, approximately half of the TPA-generated electrons are imparted with a velocity in the $-y$ -direction, away from the Au-Si interface. As the electrons undergo multiple collisions, their velocity will be limited to the saturation velocity. These calculations demonstrate that the experimental conditions are suitable for an electron avalanche that is initiated by TPA-excited electrons and driven by ponderomotive electron acceleration.

6.7 Impact Ionization and Avalanche Multiplication

The transfer of kinetic energy among electrons occurs via impact ionization. This is a three-particle process that can be observed in Si when initial free-carriers (generated by TPA) are subjected to a high electric field. As the free-carriers gain energy from the ponderomotive potential, they undergo scattering interactions with bound electrons in the valence band of the semiconductor. A threshold energy of $E_t = 2.3\text{eV}$ must be imparted to the valence band electrons for impact ionization to take place [36]. It is important to note that the electrons can gain this energy before they undergo their first collision ($\sim 28\text{nm}$) and their trajectory can be considered as ballistic during this time. Sufficiently energetic interactions promote new electrons to the conduction band, resulting in a new electron-hole pair. Secondary electron-hole pairs can attain a high energy in a similar manner and go on to multiply in an “avalanche mode”, yielding exponential growth in the number of electrons. Electrons that undergo collisions before reaching E_t can emit a photon depending on the nature of the collision.

In the present case, the laser pulse excites a free-carrier population via TPA. Once excited, the free-carriers interact with the remainder of the pulse. During acceleration with the remainder of the laser pulse, the density of free-carriers multiplies in an avalanche fashion (i.e. the instantaneous rate of multiplication is proportional to the instantaneous

free-carrier density, the definition of exponential growth). Therefore, the emitted white light intensity would be expected to grow exponentially with the TPA-induced free-carrier density, or equivalently, to grow exponentially with the excitation intensity, which is what is observed in the experiments. The distinct exponential growth displayed in Figs. 6.15-6.16 signifies that the impact ionization mechanism has taken place due to ponderomotive electron acceleration.

6.8 White Light Emission Model

The emitted white light spectrum shown in Fig. 6.15(a) is featureless and resembles that observed from a Si diode, Si-junction gate field-effect transistor (JFET), and Si-bipolar junctions under high reverse bias voltage [37–57]. In these devices, the high electric field in the depletion region accelerates free-carriers to energies exceeding the threshold for impact ionization, i.e. $E_t \approx 2.3\text{eV}$ [36]. Electrons that reach this energy and collide with valence band atoms create new electron-hole pairs and return to the conduction band edge. Therefore, electrons that are accelerated to E_t do not emit radiation and the maximum observable energy in the emission spectrum is approximately $E_{max} \sim E_g + E_t = 1.1\text{eV} + 2.3\text{eV} = 3.4\text{eV}$, where $E_g = 1.1\text{eV}$ is the band gap energy of Si. This value is consistent with the cutoff of the experimentally measured spectra of 3.3eV . Electrons with energies $E < E_t$ that undergo collisions can emit a photon whose energy is dependent on the energy of the electron and the details of the collision.

A widely-accepted multimechanism model has been used to theoretically describe the white light emission that occurs during avalanche ionization in a variety of electronic devices and a variety of physical conditions [58]. Three dominant mechanisms for light emission from electron avalanches in silicon are identified in this model and are illustrated schematically in Fig. 6.20(a). The three mechanisms are as follows:

1. Direct interband (D-I) recombination involving hot electrons and holes near $k = 0$. By application of Boltzmann's equation including electron-phonon and electron-hole pair production, it has been shown that the emission intensity can be related to the photon energy by:

$$I(h\nu)_{D-I} \propto h\nu \sqrt{h\nu - E_g} \left[1 + B_2(h\nu/W) \right] e^{-h\nu/W}, \quad (6.10)$$

where B_2 is proportional to the exponential integral, \bar{E}_i , of (E/W) , and W is calculated by the following equation [50, 51, 54]:

$$W = \frac{(eE\lambda)^2}{3(1 - \cos\theta)\hbar\omega}, \quad (6.11)$$

where θ is the angle that the velocity vector of the electrons make with the electric field and \hbar is the reduced Planck constant.

2. Intraband transitions (I-T) or Bremsstrahlung radiation occurring when the Coulomb field surrounding charged centres decelerates the electron and energy is conserved by the emission of a photon. By assuming a Maxwellian carrier distribution, the relation between photon energy and emission intensity can be written as [53]:

$$I(h\nu)_{I-T} \propto e^{-h\nu/(k_B T_e)}, \quad (6.12)$$

where T_e is the electron temperature and k_B is the Boltzmann's constant.

3. Indirect interband (I-I) recombination of electrons and holes, occurring when an electron interacts simultaneously with a hole and a phonon, satisfying the energy and momentum conservation prerequisites for photon emission. In theory, the emitted spectrum can extend to energies close to the indirect bandgap ($E_g = 1.1\text{eV}$). However, in the present experiments the spectrum extends to a low-energy cut-off of approximately $E = 1.9\text{eV}$, signifying that the majority of electrons are accelerated to high energies in a ballistic manner before undergoing a collision where they can emit a photon. By considering acoustical, optical and ionizing scattering, it has been shown that the low-energy photon emission intensity can be summarized by the following set of equations [45, 46]:

$$I(h\nu)_{I-I} = I(h\nu)_{Emis} - I(h\nu)_{Abs} \quad (6.13)$$

$$I(h\nu)_{Emis} \propto n^2 \left[\nu \frac{dn}{d\nu} + n \right] \nu^2 A \exp \left[-\frac{h\nu + k_B \theta}{k_B T} \right] \exp \left[\frac{h\nu + k_B \theta}{k_B T_h} \right] \quad (6.14)$$

$$\cdot \exp \left[-\frac{aA}{2} \right] \mathbf{I}_1(aA/2) \quad (6.15)$$

$$I(h\nu)_{Abs} \propto n^2 \left[\nu \frac{dn}{d\nu} + n \right] \nu^2 B \exp \left[-\frac{h\nu - k_B \theta}{k_B T} \right] \exp \left[\frac{h\nu - k_B \theta}{k_B T_h} \right] \quad (6.16)$$

$$\cdot \exp \left[-\frac{aB}{2} \right] \mathbf{I}_1(aB/2), \quad (6.17)$$

where $A = h\nu + k_B \theta - E_g$, $B = h\nu - k_B \theta - E_g$, $a = 1/(k_B T_e) + 1/(k_B T_h)$, n is the real component of the refractive index of silicon, $\theta = 0.063\text{eV}$ is the longitudinal optical phonon energy, T is the local silicon lattice temperature where the avalanche is taking place, T_h is the hole temperature, and \mathbf{I}_1 is the modified Bessel function of order one. Notably, the three temperatures, T , T_e , and T_h depend strongly on the experimental conditions.

The similarity between the measured white light spectrum and white light spectra emitted from electron avalanches in electronic devices is studied by fitting the multimechanism model to the experimental data, where Eqs. 6.10-6.17 are simultaneously fit using a nonlinear fitting scheme. The refractive index data used for silicon is taken from experimental data [59]. It is determined that $T = 953\text{K}$, $T_e = 3096\text{K}$, and $T_h = 1021\text{K}$ provide the optimal fit to the experimental data, with the resulting spectrum shown in Fig. 6.20(b). The excellent agreement between the model and experiment confirms the

physical origin of the light emission. The emission intensity at energies below $E = 1.9\text{eV}$ is relatively weak, demonstrating that the majority of electrons are accelerated to high energies between collisions. The absence of radiation at energies greater than 3.3eV provides direct evidence of the prominent impact ionization process.

6.9 Ultrafast Nonlinear Dynamics

Additional evidence signifying new underlying physics of the phenomenon of ponderomotive plasmonic field induced electron avalanche growth can be revealed by the time dependence of the electron dynamics. In the present nanoplasmonic waveguide, electrons are generated in the Si region via TPA and the majority of these electrons exist close to the Au-Si interface where the electric field is the most intense. In the presence of a strong F_{pond} , the electrons in the Si are accelerated and swept away very quickly in the E_{SP} field ($\sim 10^9\text{V/m}$), reaching a saturation velocity, $v_s \approx 1.3 \times 10^5\text{m/s}$. Thus, within $\sim 1.0\text{ps}$, these electrons traverse a distance, $\alpha_0 = 135\text{nm}$ from the Au-Si interface, and within $\sim 2.6\text{ps}$, they would cross the entire Si core ($h = 340\text{nm}$). Therefore, it can be expected that the measured electron dynamics take place on a timescale within this range (1-2.6ps). Electron sweeping dynamics can be captured in real time via ultrafast time-resolved spectroscopy.

Ultrafast pump-probe experiments (Appendix D) are performed on the nanoplasmonic waveguides to confirm the presence of TPA and FCA, and visualize the electron sweeping dynamics. Two experimental configurations are investigated: one where the polarization of the pump is parallel to that of the probe (co-polarized) and the other where their polarization is perpendicular (cross-polarized). An artistic rendering of the cross-polarized pump-probe characterization setup is shown in Fig. 6.21. A pump-probe time domain trace obtained in the co-polarized configuration is shown in Fig. 6.22. Interference fringes appear in the traces as the pulses begin to overlap and as the pulse peaks overlap, nonlinear loss due to TPA ($Loss \propto (P_{in}^\omega)^2$) and FCA ($Loss \propto NP_{in}^\omega$), where N is the free-carrier density) produces a sharp dip in the transmission. Due to the quadratic power scaling, TPA dominates the interaction and the FWHM of the dip is dependent primarily on the pulse duration and is found to be $FWHM \approx 367\text{fs}$. Since both the pump and probe pulses are detected, this trace is essentially an autocorrelation of the laser pulse. Cross-polarized pump-probe experiments enable the pump radiation to be filtered out with a polarizer, providing more details of the nonlinear interaction and time-dynamics.

Cross-polarized pump-probe time-domain traces for pulses propagated through a nanoplasmonic waveguide for varying pump power $P_{in}^\omega = \{142, 170, 198\} \pm 28\mu\text{W}$ are shown in Fig. 6.23(a). These traces are analogous to the pump-probe traces discussed in Section 4.3.2. Interference fringes appear in the traces as the pulses begin to overlap. As the pulse peaks coincide, a combination of TPA and FCA produce a sharp drop in the transmission. As discussed in Section 6.3.1, focusing achieved by the $NA=0.85$ microscope objective

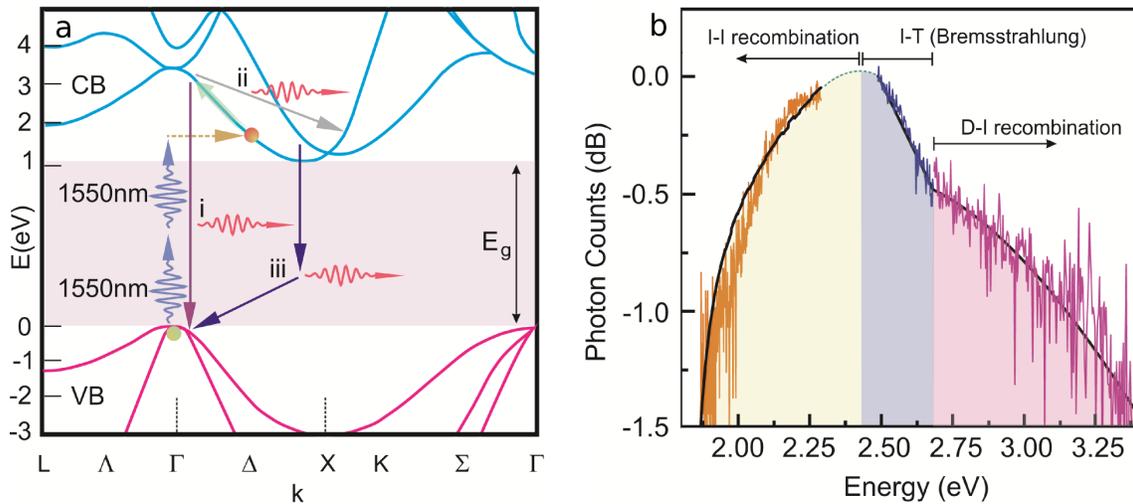


Figure 6.20: (a) Schematic depiction of photon emission processes: (i) direct interband (D-I) recombination, (ii) intraband transitions (I-T) or Bremsstrahlung radiation, and (iii) indirect interband (I-I) recombination. (b) Plot of photon counts versus photon energy, as measured in the experiments, along with the fitted curves obtained using a multimechanism model including D-I (magenta), I-T (blue), and I-I (orange).

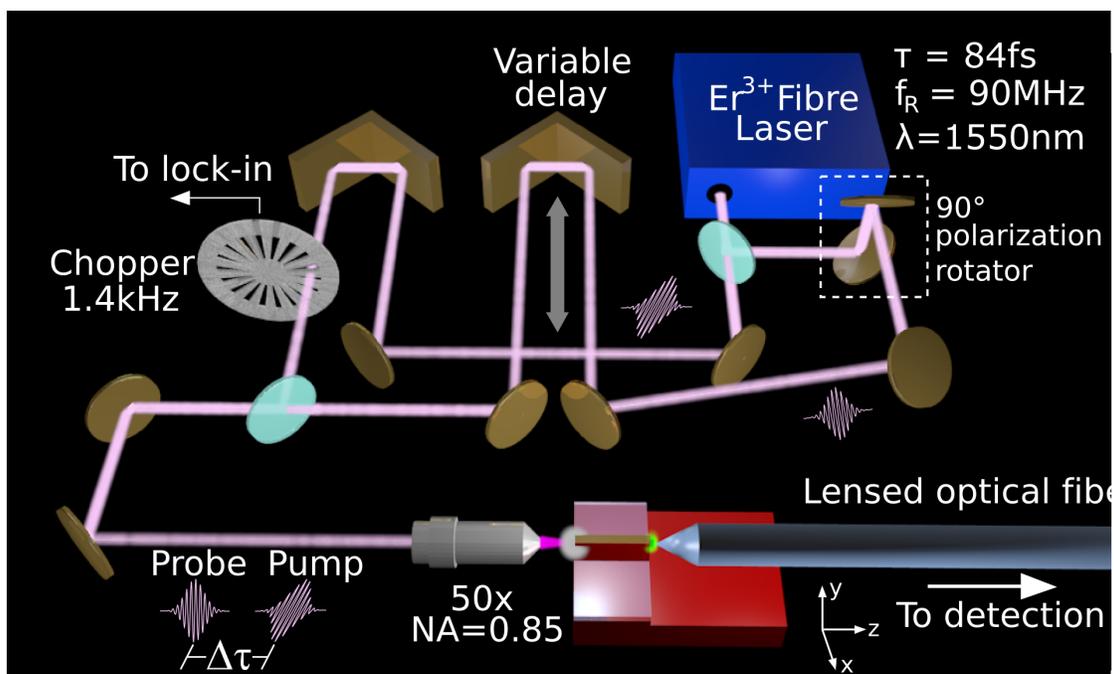


Figure 6.21: Artistic rendering of the ultrafast pump-probe characterization setup.

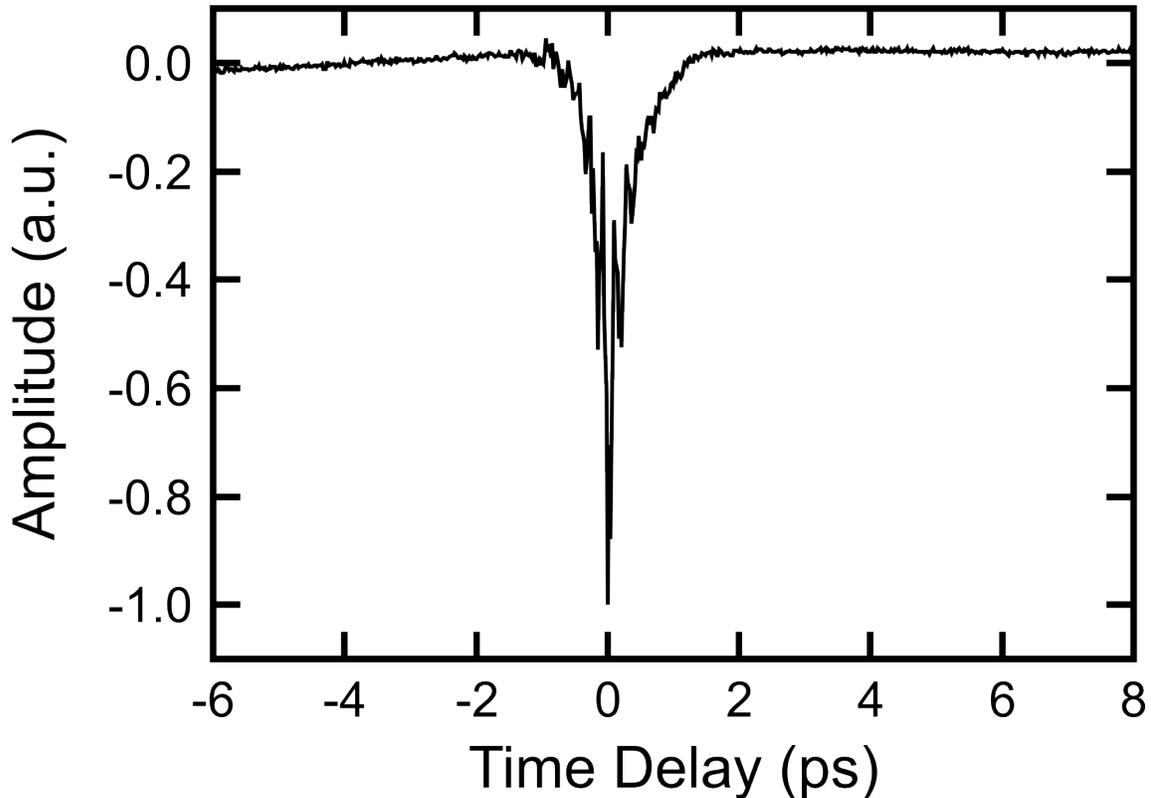


Figure 6.22: Pump-probe time-dependent power transmission for a co-polarized configuration.

produces mixed polarization components at the waveguide input facet. Therefore, some components of the pump are aligned parallel to the probe, producing the interference fringes and providing ponderomotive acceleration. As N decreases due to field sweeping and recombination (two-body, surface, Auger), the transmission signal recovers to its original amplitude. The total nonlinear loss experienced by the probe is calculated by integrating the recovery tail shown by the shaded region in Fig. 6.23(b). The total nonlinear loss is plotted against the input power on logarithmic x - and y -axes is shown in Fig. 6.23(c), and a linear fit to the data points produces a slope of 2.52 ± 0.12 , confirming the nonlinear scaling of the loss associated with TPA, FCA, and the presence of the F_{pond} driven electron avalanche. While the loss does not scale with an integer power, this demonstrates that the nonlinear interaction is not as simple as TPA and FCA alone, and demonstrates the importance of the electron avalanche multiplication. A typical pump-probe trace obtained from a bare Si photonic waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$ is shown for reference in Fig. 6.23(d).

Several distinct advantages of the plasmonic structure are apparent. The first is the free-carrier recovery time. The nanoplasmonic waveguide makes a full recovery within the scan time of $\Delta t = 22\text{ps}$, whereas the photonic waveguide recovers very little in this time window. A more detailed analysis of the recovery dynamics of the nanoplasmonic waveguide will follow shortly. Secondly, the dimensions of the nanoplasmonic waveguide

are much smaller, yet the modulation depth of the observed pump-probe time-domain trace is comparable to what was observed in the Si photonic waveguide with $L = 65\mu\text{m}$. This shows that the nanoplasmonic waveguide has a comparable modulation depth, a much faster recovery time, and occupies a footprint area that is reduced by a factor of approximately 52. This dramatically improved performance arises from a combination of plasmonic field enhancement at the Au-Si interface and the avalanche multiplication of free-carriers.

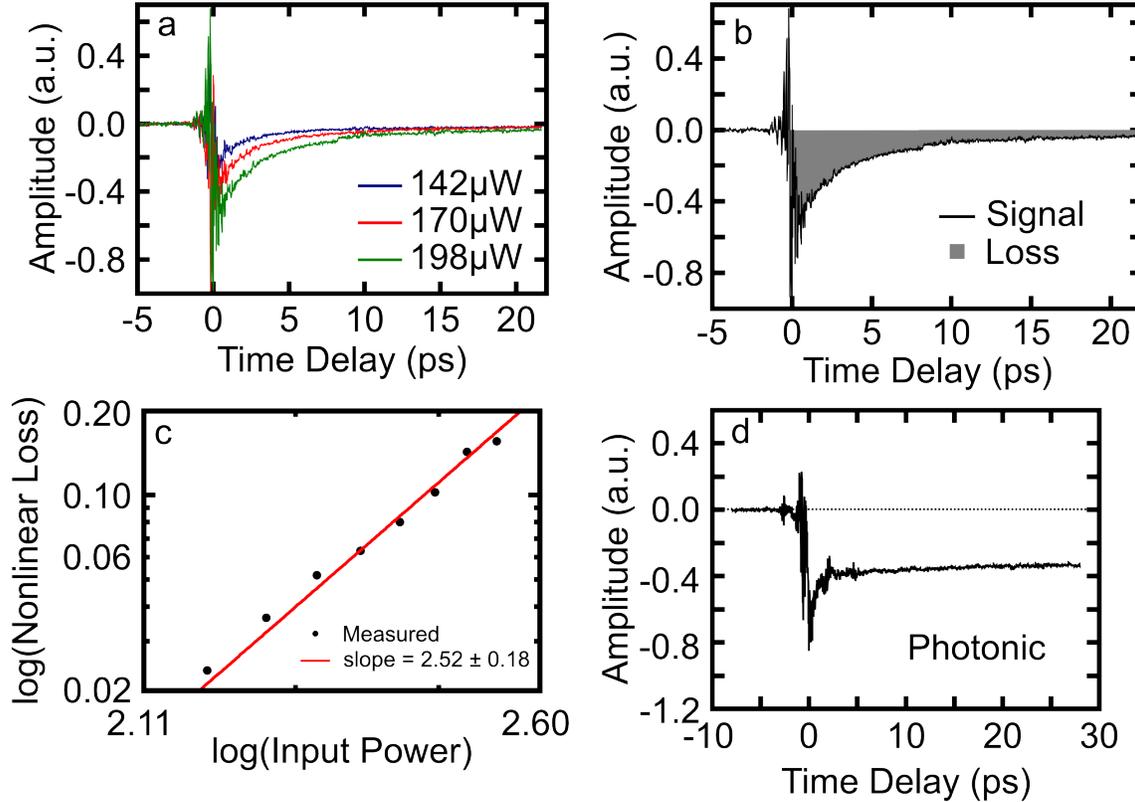


Figure 6.23: (a) Pump-probe time-dependent power transmission for the cross-polarized pump-probe configuration, for three exemplary pump powers. (b) A single pump-probe trace with a shaded area representing the total nonlinear loss experienced by the probe as the delay is varied. (c) Total nonlinear loss versus input pump power, plotted on logarithmic x - and y -axes. A linear fit to this data yields a slope of 2.52, confirming that observed excitation scales nonlinearly with the input power. (d) Pump-probe time-domain trace for a Si photonic waveguide with $w = 340\text{nm}$, $h = 340\text{nm}$, and $L = 65\mu\text{m}$, for $P_{peak}^\omega = 22.1\text{W}$.

6.9.1 Analysis of the Free-carrier Time-Dynamics

The faster recovery time of the nanoplasmonic waveguide implies different free-carrier dynamics than a bare Si waveguide. Ultrafast dynamics of the recovery tail of the pump-probe transmission graphs provide insight into the nature of the transmission recovery mechanisms. Different recovery mechanisms (or combinations of mechanisms) can be

characterized by a different time-domain function. Therefore, significant insight into the recovery mechanism can be gained through careful numerical analysis of the recovery tail. A first estimate for the time-domain function of the recovery tail can be made by considering Si photonic waveguides that operate on the principle of free-carrier excitation and recombination. It has been shown already that the recovery tail of silicon photonic waveguides follows an exponential trend with a time constant of approximately $\tau_r = 265\text{ps}$ (see Fig. 6.23(d)). Therefore, as a first estimation, an exponential function of the form $P = -Ae^{-t/\tau}$ is fit to the recovery tail of the pump-probe trace. The result of this fit is shown in Fig. 6.24(a) and fitting values of $A = 0.413$ and $\tau = 5.36\text{ps}$ are obtained. The coefficient of determination is found to be, $R^2 = 0.93$. Two main observations can be made from this fit:

1. The measured recovery occurs more quickly than the fit curve in the first $\sim 5\text{ps}$ of the recovery tail.
2. The measured recovery occurs more slowly than the fit curve in the last $\sim 10\text{ps}$ of the recovery tail shown in the figure.

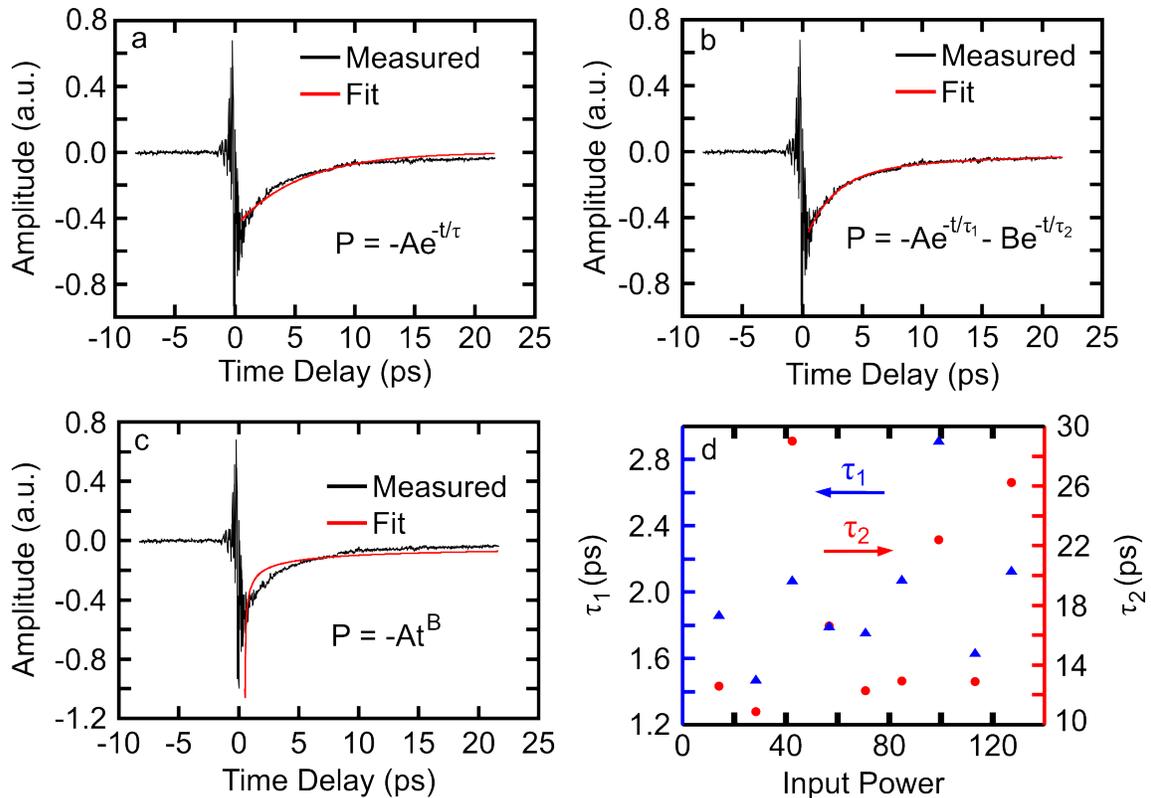


Figure 6.24: Recovery tail fit with: (a) A single exponential recovery; (b) A biexponential recovery; and (c) A power law recovery. (d) Power scaling of the two time-constants of the recovery tail, retrieved from a biexponential fit.

This strongly indicates the nanoplasmonic waveguide is recovering through two simultaneous mechanisms and can be modelled more accurately with a biexponential function

of the form: $P = -Ae^{-t/\tau_1} - Be^{-t/\tau_2}$. The result of the biexponential fit is shown in Fig. 6.24(b) and fitting values of $A = 0.368$, $\tau_1 = 2.44\text{ps}$, $B = 0.121$, $\tau_2 = 16.6\text{ps}$, and $R^2 = 0.98$ are obtained. The biexponential provides a much better fit and follows the mean path of the recovery tail virtually perfectly. In other words, the deviation of R^2 from unity arises mainly from noise in the detection setup.

Finally, since migration to surface traps can be involved, the dynamics could also be described by a power law, $P = -At^B$. The result of the power law recovery fit is shown in Fig. 6.24(c) and fitting values of $A = 0.243$, $B = -0.396$, and $R^2 = 0.74$ are obtained. It is clear from this plot that the recovery dynamics are not accurately described by a power law.

Thus, it is concluded that the nanoplasmonic waveguide recovers in a biexponential manner and that two recovery mechanisms are involved. A fit is performed to the recovery tails for various pump powers and the resulting two time constants are recorded and plotted in Fig. 6.24(d). Based on each of the traces measured, recovery times of $\tau_1 = 1.98 \pm 0.40\text{ps}$ and $\tau_2 = 17.9 \pm 6.8\text{ps}$ are determined. In this case, there is some variation in the recovery time as the input power is changed, as shown in Fig. 6.24(d).

When interacting with the intense laser field, free-carriers can be accelerated well above the saturation velocity of silicon within a half-cycle of the laser pulse, as shown in Fig. 6.19(b). Based on the energy distributions shown in Fig. 6.19(c), virtually all of the conduction electrons attain a kinetic energy greater than the saturation kinetic energy (0.0125eV) in their first 28nm of travel (i.e. statistically before their first collision). As the electrons undergo multiple collisions and continue to interact with the laser pulse, their net velocity will be limited to the saturation velocity. Therefore, for $\Delta\mathbf{r} \geq 28\text{nm}$, the majority of electrons will travel at the saturation velocity in a direction that is determined by their starting position and velocity and the series of collisions they undergo.

The fast recovery mechanism, $\tau_1 = 1.98 \pm 0.40\text{ps}$, arises from carrier sweeping out of the region of intense interaction with the probe pulse (i.e. close to the Au-Si interface). This action corresponds to the free-carriers that travel in the $-y$ direction at the saturation velocity. The probe experiences loss due to the presence of free-carriers, which is given by the following equation:

$$\frac{dI}{dz} = -\sigma_{FCA}NI, \quad (6.18)$$

where σ_{FCA} is the FCA cross-section and N is the free-carrier density. As shown in Fig. 6.6(b), the plasmonic mode profile in the silicon core is most intense at the Si-Au interface and decays exponentially normal to the interface with a length scale of $\alpha_0 = 135\text{nm}$. Therefore, as the electrons travel in the $-y$ direction, they move into regions of less intense probe radiation and contribute less to FCA of the probe. For electrons moving at the saturation velocity, a time constant of $\tau \approx 1\text{ps}$ would correspond to a displacement of α_0 . Furthermore, once the pulse has stopped interacting with the conduction band electrons, the lattice collisions gradually decrease their velocity. Based on these points,

the time required for the electrons to be swept a distance of α_0 away from the interface will be greater than 1ps, and the measured value of $\tau_1 = 1.98 \pm 0.40$ ps is a suitable value.

The slower recovery mechanism, $\tau_2 = 17.9 \pm 6.8$ ps, corresponds to non-radiative recombination of free-carriers at recombination centres and at surface traps. This occurs on a much faster timescale than the recombination time measured in silicon photonic waveguide, i.e. $\tau_p \approx 265$ ps. Due to the high intensity of the fields near the Au-Si interface, a majority of the free-carriers are generated in close proximity to the interface. Electron-hole pairs that migrate to the interface are able to recombine rapidly at surface traps. Furthermore, during nanofabrication processing, the temperature of the sample is increased up to a maximum of 300°C during certain processing steps. This temperature increase leads to diffusion of gold atoms into the silicon waveguide core. The gold atoms act as recombination centres, decreasing the free-carrier recombination time in the silicon. While many of the electrons are swept out of the high field region of the waveguide, the trajectory of some electrons is such they can remain in the high field region of the device and continue to contribute to free-carrier absorption until they encounter a recombination centre. In the present case, diffusion of the Au cap into the Si waveguide core leads to an increase the density of recombination centres and surface traps near the Si-Au interface, increasing the surface recombination velocity.

6.10 Summary

These results demonstrate several unique ultrafast nonlinear interactions in a nanoplasmonic waveguide that is largely CMOS-compatible. Perhaps the most significant aspects of this structure are its ultracompact footprint area of $0.43\mu\text{m}^2$, the strong nonlinear signatures that are generated in a very small interaction length, and its operation at telecommunication wavelengths. Third-harmonic generation with a high conversion efficiency of $\sim 10^{-5}$ and green light powers up to ~ 35 nW were measured, demonstrating the potential for compact visible light sources for integrated photonics or hybrid photonic-electronic circuitry. This would provide a means to generate coherent visible light that is synchronized to the pump excitation to deliver accurate synchronization clock pulses for CMOS electronics that are generated in the Si material and detected by the same Si material.

Electrons excited via TPA are accelerated to energies up to several eV by the ponderomotive potential that exists in the highly confined plasmonic field. Subsequent collisions with valence electrons drive impact ionization, which multiplies in an avalanche. Electrons that do not reach the threshold for impact ionization emit photons during collisions, and white light emission is observed to scale exponentially with the input power, confirming the exponential growth of avalanching electrons. Electron sweeping due to ponderomotive electron acceleration is visualized using pump-probe time domain spectroscopy and the sweeping time is measured to be $\tau_1 = 1.98 \pm 0.40$ ps, confirming that the electrons reach

saturation velocity.

Pondermotive electron acceleration has been used as a scheme for the generation of ultrashort energetic free electrons bunches from metal films into vacuum, but has not been observed in semiconductors. These findings uncover a new strong-field interaction that can be used in sensitive nanoplasmonic modulators and hybrid plasmonic-electronic transducers. This nanoplasmonic waveguide can be operated as a sensitive, ultrafast modulator that responds on a timescale of 2ps (on the order of the timescale of free-carrier diffusion in the channel of a field-effect transistor). Furthermore, this structure can be operated as an avalanche photodetector, where the avalanche can be driven either electrically, optically, or through a combination of both. This will enable a sensitive means to convert optical signals to electrical signals. Finally, the nanoplasmonic waveguide dimensions could be reduced to dimensions of $w \times L \approx 100\text{nm} \times 1000\text{nm} = 0.1\mu\text{m}^2$ with negligible change in the THG conversion efficiency or performance as an ultrafast modulator. Therefore, it could be easily integrated with low-loss Si-photonics waveguides and provide all-optical modulation in a much smaller footprint than similar Si-photonics waveguides.

References

- [1] M. Kauranen and A. V. Zayats. Nonlinear plasmonics. *Nature Photon.*, 6:737–748, 2012.
- [2] N. Rotenberg, M. Betz, and H. M. van Driel. Ultrafast all-optical coupling of light to surface plasmon polaritons on plain metal surfaces. *Phys. Rev. Lett.*, 105:017402, 2010.
- [3] A. V. Krasavin and N. I. Zheludev. Active plasmonics: controlling signals in au/ga waveguide using nanoscale structural transformations. *Appl. Phys. Lett.*, 84:1416–1418, 2004.
- [4] A. V. Krasavin, K. F. MacDonald, N. I. Zheludev, and A. V. Zayats. High-contrast modulation of light with light by control of surface plasmon polariton wave coupling. *Appl. Phys. Lett.*, 85:3369–3371, 2004.
- [5] D. Pacifici, H. I. Lezec, and H. A. Atwater. All-optical modulation by plasmonic excitation of cdse quantum dots. *Nature Photon.*, 1:402–406, 2007.
- [6] R. A. Pala, K. T. Shimizu, N. A. Melosh, and M. L. Brongersma. A nonvolatile plasmonic switch employing photochromic molecules. *Nano Letters*, 8:1506–1510, 2008.
- [7] K. F. MacDonald, Z. L. Samson, M. I. Stockman, and N. I. Zheludev. Ultrafast active plasmonics: transmission and control of femtosecond plasmon signals. *Nature Photon.*, 3:55–58, 2008.
- [8] R. W. Minck, R. W. Terhune, and C. C Wang. Nonlinear optics. *Appl. Opt.*, 5:1595–1612, 1966.
- [9] R. Dekker, N. Usechak, M. Forst, and A. Driessen. Ultrafast nonlinear all-optical processes in silicon-on-insulator waveguides. *J. Phys. D: Appl. Phys.*, 40:R249–R271, 2007.
- [10] G. P. Agrawal. *Nonlinear fiber optics*. Springer, Berlin, 2000.

- [11] H. K. Tsang, C. S. Wong, T. K. Liang, I. E. Day, S. W. Roberts, A. Harpin, J. Drake, and M. Asghari. Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength. *Appl. Phys. Lett.*, 80:416–418, 2002.
- [12] M. Dinu, F. Quochi, and H. Garcia. Third-order nonlinearities in silicon at telecom wavelengths. *App. Phys. Lett.*, 82:2954, 2003.
- [13] R. Claps, D. Dimitropoulos, Y. Han, and B. Jalali. Observation of raman emission in silicon waveguides at 1.54 μm . *Opt. Express*, 10:1305–1313, 2002.
- [14] H. Rong, A. Liu, R. Jones, O. Cohen, D. Hak, R. Nicolaescu, A. Fang, and M. Paniccia. An all-silicon raman laser. *Nature*, 433:292–294, 2005.
- [15] R. Espinola, J. Dadap, Jr. R. Osgood, S. McNab, and Y. Vlasov. C-band wavelength conversion in silicon photonic wire waveguides. *Opt. Express*, 13:4341–4349, 2005.
- [16] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, T. Tsuchizawa, T. Watanabe, J. I. Takahashi, and S. I. Itabashi. Four-wave mixing in silicon wire waveguides. *Opt. Express*, 13:4629–4637, 2005.
- [17] O. Boyraz, P. Koonath, V. Raghunathan, and B. Jalali. All optical switching and continuum generation in silicon waveguides. *Opt. Express*, 12:4094–4102, 2004.
- [18] R. Dekker, A. Driessen, T. Wahlbrink, C. Moormann, J. Niehusmann, and M. Först. Ultrafast kerr-induced all-optical wavelength conversion in silicon waveguides using 1.55 μm femtosecond pulses. *Opt. Express*, 14:8336–8346, 2006.
- [19] O. Boyraz, T. Indukuri, and B. Jalali. Self-phase-modulation induced spectral broadening in silicon waveguides. *Opt. Express*, 12:829–834, 2004.
- [20] G. W. Rieger, K. S. Virk, and J. F. Young. Nonlinear propagation of ultrafast 1.5 μm pulses in high-index-contrast silicon-on-insulator waveguides. *Applied Physics Letters*, 84:900–902, 2004.
- [21] C. Corcoran, C. Monat, C. Grillet, D. J. Moss, B. J. Eggleton, T. P. White, L. O’Faolain, and T. F. Krauss. Green light emission in silicon through slow-light enhanced third-harmonic generation in photonic-crystal waveguides. *Nature Photon.*, 3:206–210, 2009.
- [22] M. Lippitz, M. A. van Dijk, and M. Orrit. Third-harmonic generation from single gold nanoparticles. *Nano Letters*, 5:799–802, 2005.
- [23] E. M. Kim, S. S. Elovikov, T. V. Murzina, A. A. Nikulin, O. A. Aktsipetrov, M. A. Bader, and G. Marowsky. Surface-enhanced optical third-harmonic generation in ag island films. *Phys. Rev. Lett.*, 95:227402, 2005.

- [24] T. Xu, X. Jiao, and S. Blair. Third-harmonic generation from arrays of sub-wavelength metal apertures. *Opt. Express*, 17:23582–23588, 2009.
- [25] S. Sederberg, V. Van, and A. Y. Elezzabi. Monolithic integration of plasmonic waveguides into a complimentary metal-oxide-semiconductor- and photonic-compatible platform. *Appl. Phys. Lett.*, 96:121101, 2010.
- [26] J. A. Dionne, K. Diest, L. A. Sweatlock, and H. A. Atwater. Plasmostor: A metal-oxide-si field effect plasmonic modulator. *NanoLetters*, 9:897–902, 2009.
- [27] J. A. Dionne, L. A. Sweatlock, M. T. Sheldon, A. P. Alivisatos, and H. A. Atwater. Silicon-based plasmonics for on-chip photonics. *Selected Topics in Quantum Electronics, IEEE Journal of*, 16:295–306, 2010.
- [28] P. N. Butcher and D. Cotter. *The Elements of Nonlinear Optics*. University Press, Cambridge, 1990.
- [29] A. Mooradian. Photoluminescence of metals. *Phys. Rev. Lett.*, 22:185–187, 1969.
- [30] S. E. Irvine, A. Dechant, and A. Y. Elezzabi. Generation of 0.4-keV femtosecond electron pulses using impulsively excited surface plasmons. *Phys. Rev. Lett.*, 93:184801, 2004.
- [31] S. E. Irvine and A. Y. Elezzabi. Ponderomotive electron acceleration using surface plasmon waves excited with femtosecond laser pulses. *Appl. Phys. Lett.*, 86:264102, 2005.
- [32] P. Dombi, S. E. Irvine, P. Racz, M. Lenner, N. Kroo, G. Farkas, A. Mitrofanov, A. Baltuska, T. Fuji, F. Krausz, and A. Y. Elezzabi. Observation of few-cycle, strong-field phenomena in surface plasmon fields. *Opt. Express*, 18:24206–24212, 2010.
- [33] R. N. Dexter, H. J. Zeiger, and B. Lax. Cyclotron resonance experiments in silicon and germanium. *Phys. Rev.*, 104:637–644, 1956.
- [34] C. Canali, G. Majni, R. Minder, and G. Ottaviani. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. *Electron Devices, IEEE Transactions on*, 22:1045–1047, 1975.
- [35] S. E. Irvine and A. Y. Elezzabi. Surface-plasmon-based electron acceleration. *Phys. Rev. A*, 73:013815, 2006.
- [36] P. A. Wolf. Theory of electron multiplication in silicon and germanium. *Phys. Rev.*, 95:1415–1420, 1954.
- [37] R. Newman. Visible light from silicon p-n junctions. *Phys. Rev.*, 100:700–703, 1955.

- [38] A. G. Chynoweth and K. G. McKay. Photon emission from avalanche breakdown in silicon. *Phys. Rev.*, 102:369–376, 1956.
- [39] J. A. Burton. Electron emission from avalanche breakdown in silicon. *Phys. Rev.*, 108:1342–1343, 1957.
- [40] A. G. Chynoweth and G. L. Pearson. Effect of dislocations on breakdown in silicon p-n junctions. *J. Appl. Phys.*, 29:1103–1110, 1958.
- [41] M. Kikuchi and K. Tachikawa. Visible light emission and microplasma phenomena in silicon p-n junction, i. *J. Phys. Soc. Jpn.*, 15:835–848, 1959.
- [42] M. Kikuchi. Visible light emission and microplasma phenomena in silicon p-n junction, ii. classification of weak spots in diffused p-n junctions. *J. Phys. Soc. Jpn.*, 15:1822–1831, 1960.
- [43] R. P. Gupta, W. S. Khokle, H. Adachi, and H. L. Hartnagel. A new avalanche-plasma light emitter and its physical analysis. *J. Phys. D: Appl. Phys.*, 14:L31–L35, 1985.
- [44] A. Toriumi, M. Yoshimi, M. Iwase, Y. Akiyama, and K. Taniguchi. A study of photon emission from n-channel mosfet's. *IEEE T. Electron Dev.*, ED-34:1501–1508, 1987.
- [45] D. K. Gautam, W. S. Khokle, and K. B. Garg. Effect of absorption on photon emission from reverse-biased silicon p-n junctions. *Solid State Electron.*, 31:1119–1121, 1988.
- [46] D. K. Gautam, W. S. Khokle, and K. B. Garg. Photon emission from reverse-biased silicon p-n junctions. *Solid State Electron.*, 31:219–222, 1988.
- [47] E. S. Snow, S. W. Kirchoefer, P. M. Campbell, and O. J. Glembocki. Efficient light emission by impact ionization in single barrier tunnelling devices. *Appl. Phys. Lett.*, 54:2124–2126, 1989.
- [48] B. M. Arora and N. C. Das. Luminescence spectra of an n-channel metal-oxide-semiconductor field-effect transistor at breakdown. *Appl. Phys. Lett.*, 56:1152–1153, 1990.
- [49] B. P. van Drienuizen and R. F. Wolffenbuttel. Optocoupler based on the avalanche light emission in silicon. *Sensor. Actuat. A-Phys.*, 31:229–240, 1992.
- [50] J. Bude, N. Sano, and A. Yoshii. Hot-carrier luminescence in si. *Phys. Rev. B*, 45:5848–5856, 1992.
- [51] S. Yamada and M. Kitao. Recombination radiation as a possible mechanism of light emission from reverse-biased p-n junctions under breakdown condition. *Jpn. J. Appl. Phys.*, 32:4555–4559, 1993.

- [52] J. Kramer, P. Seitz, E. F. Steigmeir, H. Auderset, and B. Delley. Light-emitting devices in industrial cmos technology. *Sensor. Actuat. A-Phys.*, 37-38:527–533, 1993.
- [53] A. L. Lacaita, F. Zappa, S. Bigliardi, and M. Manfredi. On the bremsstrahlung origin of hot-carrier-induced photons in silicon devices. *IEEE T. Electron. Dev.*, 40:577–582, 1993.
- [54] A. T. Obeidat, A. Kalayjian, A. G. Andreou, and J. B. Khurgin. A model for visible photon emission from reverse-biased silicon p-n junctions. *Appl. Phys. Lett.*, 70:470–471, 1997.
- [55] N. Akil, S. E. Kerns, D. V. Hoffman, and J. P. Charles. Photon generation by silicon diodes in avalanche breakdown. *Appl. Phys. Lett.*, 73:871–872, 1998.
- [56] M. de la Bardonnie, D. Jiang, S. E. Kerns, D. V. Kerns Jr., P. Mialhe, J. P. Charles, and A. Hoffman. On the aging of avalanche light emission from silicon junctions. *IEEE T. Electron Dev.*, 46:1234–1239, 1999.
- [57] A. T. Fiory and N. M. Ravindra. Light emission from silicon: some perspectives and applications. *IEEE T. Electron Dev.*, 32:1043–1051, 2003.
- [58] N. Akil, S. E. Kerns, D. V. Kerns Jr., A. Hoffman, and J. P. Charles. A multimechanism model for photon generation by silicon junctions in avalanche breakdown. *IEEE T. Electron Dev.*, 46:1022–1028, 1999.
- [59] E. D. Palik. *Handbook of optical constants of solids*. Academic Press, San Diego, 1998.

Chapter 7

Nanoplasmonic Devices: Active and Passive

7.1 Overview

As outlined in Chapters 2 and 6, one of the key functionalities of nanoplasmonic structures is ultrafast all-optical modulation. The ability to modulate an optical signal on an ultrafast timescale in a subwavelength structure that is compatible with complementary metal-oxide-semiconductor (CMOS) processing techniques would enable the development of high-density all-optical information processing circuitry. It was also demonstrated in Chapter 2 that special measures had to be taken to couple free-space radiation into subwavelength nanoplasmonic waveguides. Nanoplasmonic antennas have the ability to confine free-space radiation to a $\sim 10\text{nm}$ gap [1] and can be used to couple this radiation to a nanoplasmonic waveguide [2]. This chapter presents five classes of nanoplasmonic devices that are designed to achieve ultrafast all-optical modulation or nanoscale confinement of electromagnetic energy via nanoplasmonic antennas.

7.2 Ultrafast All-Optical Switching and Modulation in Silicon-Based Nanoplasmonic Devices

7.2.1 Introduction

¹The last several decades have witnessed the growth and maturing of complex, high-bandwidth fibre optic networks and photonic devices. Unlike electronic devices that rely on conventional charge transport and operate at limited bandwidth, their photonic counterparts exhibit significantly higher operating speeds and perform at much lower power

¹The content of this section has been published. (1) A. Y. Elezzabi, Z. Han, S. Sederberg, and V. Van, *Opt. Express* 17, 11045-11056 (2009). (2) S. Sederberg, D. Driedger, M. Nielsen, and A. Y. Elezzabi, *Opt. Express* 19, 23494-23503 (2011).

consumption. Replacing conventional electronic circuit components with reliable, high-speed optical circuits has been one of the thrusts behind the emergence of the field of nanophotonics. However, displacing the mature electronic technology requires alternatives that possess not only better performance, but also ease of integration with the existing devices and systems. Although early nanophotonic devices were analogous to bulk and integrated photonic devices, this field has continued to evolve, and a major contender for the next-generation nanophotonic systems has emerged [3].

Nanoscale waveguides including metallic features can support plasmonic modes and can be designed to have sub-diffraction dimensions [4]. Based on demonstrations of high integration density and sharp bending radii, it has been proposed that plasmonic waveguides could form a basis for the next generation chip-scale optical technology [5]. Moreover, the high modal confinement offered by plasmonic waveguides enhances incident electric fields and allows for efficient access to nonlinear optical effects. Since the device architecture is fundamentally metal-based, it is envisioned that hybrid plasmonic-electronic devices can be integrated within the current electronic device architecture where the nanoscale metallic-dielectric waveguides allow for the transport of both electrical and optical signals simultaneously. The latter is of particularly great importance because as electronic components in a CMOS processor approach several nanometers in size, nanoplasmonic waveguides are considered an alternative route for chip interconnects. While progress has been realized in the development of nanoplasmonic optical elements, to date the majority of such components are passive elements. To realize advanced active circuit architectures, there is a need for active modulators and switching components operating at ultrahigh bandwidths and low power consumption.

Early approaches to plasmonic signal modulation involved thermo-optic effects [6, 7], plasmon amplification [8–11], charge redistribution in n^+ -Si via electric field application [12], radiation-induced interband transitions in aluminum [13], and magnetic field modulation of surface magneto-plasmon waves [14]. Although modulation of plasmonic amplitude has been shown, the effects upon which these modulations were induced restrict their operation bandwidth. Since the response of the charge modulation in silicon (Si) to an applied electric field is, in principle, similar to that of CMOS devices, its inherent capacitance limits operation to a few gigahertz. Similarly, the response of the thermo-optic effect is constrained by thermal diffusivity where switching times have been in the millisecond regime, and the bandwidth of the magneto-surface plasmon modulator is constrained by the strength of the biasing magnetic field.

7.2.2 Ultrafast Free-Carrier Time-Dynamics in Silicon

By considering Si-loaded nanoplasmonic waveguides, above-bandgap pulses can be coupled into the waveguides, exciting free-carriers in the Si features. This will alter the complex permittivity of the Si features, due to free-carrier absorption (FCA) and free-

carrier plasma dispersion (FCD), and can be used to modulate the amplitude of the plasmonic mode. The above-bandgap pulses will excite a time-dependent permittivity, $\varepsilon(t)$, in the Si features as free-carriers are excited and recombine, which is modeled by a frequency-dependent Drude-Lorentz model for a semiconductor [15]:

$$\varepsilon(t) = \varepsilon'_b \left(1 - \frac{n(t)e^2\langle\tau\rangle^2}{\varepsilon'_b\varepsilon_0m^*(1 + \omega^2\langle\tau\rangle^2)} \right) - i\varepsilon'_b \left(\frac{\varepsilon''_b}{\varepsilon'_b} + \frac{n(t)e^2\langle\tau\rangle^2}{\varepsilon'_b\varepsilon_0m^*\omega(1 + \omega^2\langle\tau\rangle^2)} \right), \quad (7.1)$$

where $\varepsilon'_b = 11.8$ and $\varepsilon''_b = 9.2 \times 10^{-4}$ are the intrinsic real and imaginary components of the background permittivity for Si which include the contribution of bound charges, ε_0 is the free-space permittivity, ω is the signal radiation frequency, e is the electronic charge, m^* is the free-carrier effective mass, where $m^{*-1} = m_e^{*-1} + m_h^{*-1}$, with $m_e^* = 0.259m_0$ as the electron and $m_h^* = 0.38m_0$ as the hole effective masses, respectively, m_0 is the electron mass, $\langle\tau\rangle = 10.6\text{fs}$ is the average carrier-carrier relaxation time, and $n(t)$ is the photogenerated carrier density in the conduction band. Both the real and imaginary parts of the permittivity and, therefore, the refractive index of the Si layers, are related to the density of the photogenerated e-h pairs (i.e. $n(t)$). For photoinjection of carriers at a density above a critical e-h density, n_c , corresponding to:

$$n_c = \frac{\varepsilon_0\varepsilon'_bm^*}{e^2} \frac{1 + \omega^2\langle\tau\rangle^2}{\langle\tau\rangle^2}, \quad (7.2)$$

the real part of the permittivity becomes negative, and the Si features behave like metals at the frequency ω . For example, for Si at $\lambda = 1550\text{nm}$, $n_c = 8.5 \times 10^{20}\text{cm}^{-3}$. From hereon in, all pump powers are referred to in terms of the maximum carrier concentration relative to the critical carrier density, i.e. n_0/n_c .

The free-carriers are generated by above-bandgap control pulses having a duration, $\tau_p = 10\text{fs}$, which are assumed to have a sech^2 temporal envelope. The dominant free-carrier decay process is assumed to be recombination of e-h pairs at recombination centres in the Si layers at a characteristic time, τ_r . Therefore, the instantaneous free-carrier density can be obtained from the equation:

$$\frac{\partial n}{\partial t} - \frac{n_0}{\tau_p} \text{sech}^2\left(\frac{t - t_0}{\tau_p}\right) - \frac{n(t)}{\tau_r}, \quad (7.3)$$

where n_0 is the maximum carrier concentration at the peak of the control pulse when $t = t_0$, τ_p is the duration of the control pulse, and $\tau_r = 1\text{ps}$ is the free-carrier recombination time in Si, which is assumed to be ion-implanted Si (II-Si). Implantation of 10^{13}cm^{-2} O^+ ions at both 200 keV and 100 keV into a Si layer produces ultrafast recombination centres that reduce the carrier lifetime from $\tau_r \sim 100\text{ns}$ to $\tau_r \sim 1\text{ps}$ [16].

7.2.3 The Five-Layer Plasmonic Slot Waveguide

Geometry

²We consider a five-layer planar plasmonic waveguide structure, as depicted in Fig. 7.1(a). The structure consists of a central free-space region of thickness, a , sandwiched between two semiconducting layers each of thickness, $d = (b - a)/2$. Unlike a conventional metal-insulator-metal plasmonic waveguide where there are only two metal-dielectric interfaces, the five-layer structure provides two additional air-semiconductor interfaces. This effectively introduces an extra field mode confinement parameter that can be used to vary the refractive index and the losses within the semiconductor layer to achieve enhanced on/off modulation ratios. Capping each semiconductor layer with a metal cladding layer allows for plasmonic modes to be guided through the device. To achieve an ultrahigh bandwidth operation, the choice of the semiconductor material is restricted to those possessing ultrafast carrier lifetimes. An excellent candidate for the material is II-Si which not only exhibits an ultrafast recombination lifetime for the photogenerated carriers, but also allows for ease of integration with metallic thin films.

Passive Operation

This analysis uses a guided-wave formalism to study the propagation characteristics in the plasmonic waveguide. While this structure can support a number of modes, only the lowest-order TM even (symmetric) mode will be analysed. As will be shown later, the TM mode provides almost twice the power confinement in the Si layers when compared to the TE mode, allowing for more efficient modulation of the field amplitude. Furthermore, the lowest-order mode is chosen because practically, it is the easiest to excite and is the least lossy. Thus, we seek to derive an analytical expression for the characteristic dispersion relation of this mode.

For the two dimensional planar waveguide with its propagation axis for a TM-polarized laser signal radiation along the x - direction, the temporal and spatial evolution of the light wave electric and magnetic fields are described by $e^{-i(\omega t - \beta x)}$, where ω is the frequency of the signal radiation and β is the complex propagation factor ($\beta = \beta_r + i\beta_i$). The light wave is assumed to be independent of the y -direction and thus for the TM polarization the magnetic field, H_y , can accordingly be described as: $H_y(x, y, z, t) = H_y(z)e^{-i(\omega t - \beta x)}$.

We commence the analysis by considering the five-layer waveguide, where the layers (i.e free-space, Si, and metal) are characterized by their intrinsic frequency-dependent complex permittivities and constant magnetic permeabilities. Hereafter, the following subscripts f , s , and m , refer to the free-space, Si, and metal regions, respectively. The

²These simulations were performed with assistance from Zhanghua Han.

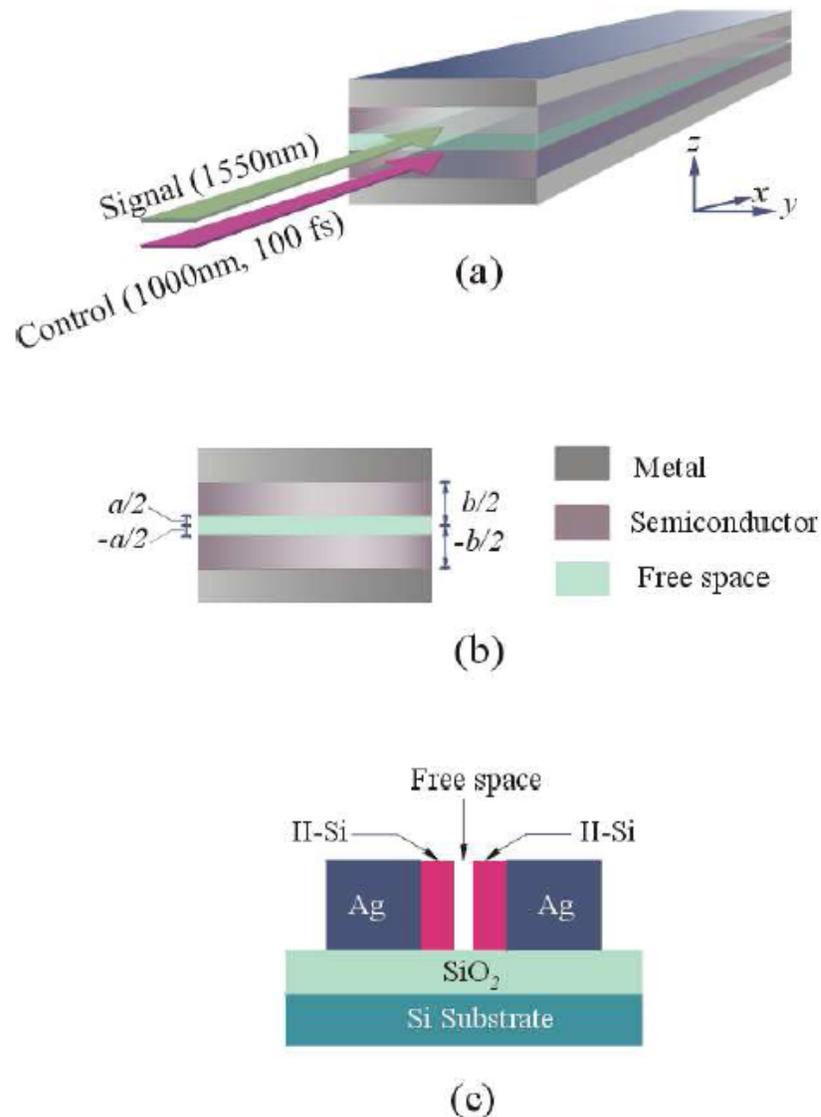


Figure 7.1: (a) A schematic diagram showing the nanoplasmonic waveguide utilized as an ultrafast optical modulator and/or switch for infrared radiation. (b) Depicts an embodiment of the device realized using II-Si and silver (Ag) layers fabricated on top of a silicon dioxide (SiO₂) layer and a Si layers. The device can be fabricated using conventional metal and Si deposition and etching techniques, whereas the free-space region is fabricated via focused ion beam milling.

characteristic TM mode wave equation describing the magnetic field is given by:

$$\frac{\partial H_y^2}{\partial z^2} + k_i^2 H_y = 0 \quad (i = f, s, m) \quad (7.4)$$

where $k_f^2 = \varepsilon_f k_0^2 - \beta^2$, $k_s^2 = \varepsilon_s k_0^2 - \beta^2$, and $k_m^2 = \beta^2 - \varepsilon_m k_0^2$ are the transverse wavevectors in the free-space, Si, and metallic regions, respectively; ε_f , ε_s , and ε_m are the relative permittivities of the free-space, Si, and metallic layers, respectively; and k_0 is the radiation wavevector in vacuum.

For a propagating even TM mode, we seek a general solution to the wave equation in the upper half space ($z \geq 0$) where the fields in both the free-space gap and the Si regions are of an oscillatory nature. Accordingly, we choose the following solution:

$$H_y(z) = \begin{cases} \cos(k_f z), & z \in (0, \frac{a}{2}) \\ A \cos(k_s z) + B \sin(k_s z), & z \in (\frac{a}{2}, \frac{b}{2}) \\ C e^{-k_m z}, & z \in (\frac{b}{2}, +\infty) \end{cases} \quad (7.5)$$

where A , B , and C are magnetic field amplitude coefficients, and a and b define the bounds of the mode in the z -direction according to Fig. 7.1(a). It should be noted that the field distribution in the lower half space is symmetric and that this solution is general even in the case where $\beta_r^2 > \varepsilon_i k_0^2$ ($i = f, s$) or k_i ($i = f, s$) becomes imaginary. In this situation the mode distribution takes an exponential form describing surface modes.

Substituting the above solution of $H_y(z)$ into Maxwell's equation yields the following TM mode electric field components:

$$E_x(z) = \frac{-i}{\omega \varepsilon_0 \varepsilon_i} \frac{\partial H_y}{\partial z} = \begin{cases} \frac{ik_f}{\omega \varepsilon_0 \varepsilon_f} \sin(k_f z), & z \in (0, \frac{a}{2}) \\ \frac{ik_s}{\omega \varepsilon_0 \varepsilon_s} (A \sin(k_s z) - B \cos(k_s z)), & z \in (\frac{a}{2}, \frac{b}{2}) \\ \frac{ik_m}{\omega \varepsilon_0 \varepsilon_m} C e^{-k_m z}, & z \in (\frac{b}{2}, +\infty) \end{cases} \quad (7.6)$$

Next, by invoking the continuity of the tangential field components (i.e. H_y and E_x) at the two interfaces $z = a/2$ and $z = b/2$, we obtain the characteristic equation for the even TM mode:

$$\tan\left(k_f \frac{a}{2}\right) = \frac{k_m \varepsilon_f}{k_f \varepsilon_m} \frac{1 - \frac{k_s \varepsilon_m}{k_m \varepsilon_s} \tan\left(k_s \frac{b-a}{2}\right)}{1 + \frac{k_m \varepsilon_s}{k_s \varepsilon_m} \tan\left(k_s \frac{b-a}{2}\right)}. \quad (7.7)$$

In the present case, since the lossy metal and Si layers are described by complex permittivities, Eq. 7.7, has a complex transcendental form, whose solutions give discrete values for the propagation factor, β , or the wavevectors k_f , k_s and k_m in the complex plane. Consecutive solution of the complex transcendental equation yields the propagation characteristics for the TM even mode. Notably, when the Si layer thickness fills the free-space

gap (i.e. $a = 0$), the above dispersion relation reduces to:

$$\tan\left(k_s \frac{b}{2}\right) = \frac{k_m \varepsilon_s}{k_s \varepsilon_m}, \quad (7.8)$$

as for a conventional metal-dielectric-metal waveguide [17].

Ultrafast Active Operation

The utilization of the aforementioned plasmonic waveguide structure as an ultrafast amplitude modulator or as a switch is as follows: When a TM-polarized laser signal radiation, whose photon energy, $h\nu_{signal}$, is less than the bandgap energy (E_g) of the Si, is launched into the plasmonic waveguide, it propagates and only suffers losses due to the metallic layers as the Si layers are transparent to the signal radiation. However, the optical properties, and particularly the losses, of the waveguide can be altered on an ultrafast time scale by the simultaneous co-propagation of a TE-polarized ultrashort laser pulse (hereafter referred to as the control pulse) whose photon energy, $h\nu_{control} \geq E_g$. The choice of TE-polarization of the control pulse ensures long propagation distance, whereas the choice $h\nu_{control} \geq E_g$ ensures that free charge carriers (electron-hole (e-h) pairs) are photogenerated in the Si layers. The transient photoinjection of the e-h pairs effectively modifies both the real and the imaginary refractive indices of the Si layers. Once a high enough carrier density is produced, the plasmonic waveguide becomes too lossy to support the propagation of the signal radiation, resulting in significant attenuation at this wavelength. However, while the turn-off process is limited only by the duration of the ultrafast control pulse, which can be as short as a few femtoseconds, the turn-on mechanism, or the recovery time of the device, is limited by the recombination lifetime of the e-h pairs in the Si layers. Nevertheless, the use of ultrafast carrier lifetime semiconductors, such as LT-GaAs [18–20], LT-In_{0.52}Al_{0.48} [21], LT-InP [22], LT-InGaAs [23], H⁺-implanted InP [24], H⁺-implanted GaAs [25, 26], or O⁺ implanted Si [16], alleviates such a constraint.

In the following investigation, we employ a signal laser radiation at $\lambda = 1550\text{nm}$ and a $\tau_p = 100\text{fs}$ ultrafast control laser pulse at a wavelength of $\lambda = 1000\text{nm}$. For the metallic layers in the cladding regions, we chose silver metal with $\varepsilon_m = -125.735 + 3.233i$ due to its low loss relative to other metals at $\lambda = 1550\text{nm}$. Low losses are an important property of “good” plasmonic metals, and must meet the following condition: $\text{Im}\{\varepsilon_m\} \ll \text{Re}\{\varepsilon_m\}$ [27], which is satisfied better by silver than the other noble metals.

Although the mode profile of the control pulse dictates the transverse distribution of free-carriers in the Si layers, the analysis can be simplified by assigning an average induced complex refractive index across the thin silicon layer. This assumption is consistent with the fact that the Si layer thickness is much less than the control pulse wavelength and carrier diffusion occurs on a nanosecond timescale, such that during the switching process it can be taken as a constant. In determining the propagation of the control pulse along the x -direction, the spatial extent of the control radiation is evaluated from the following

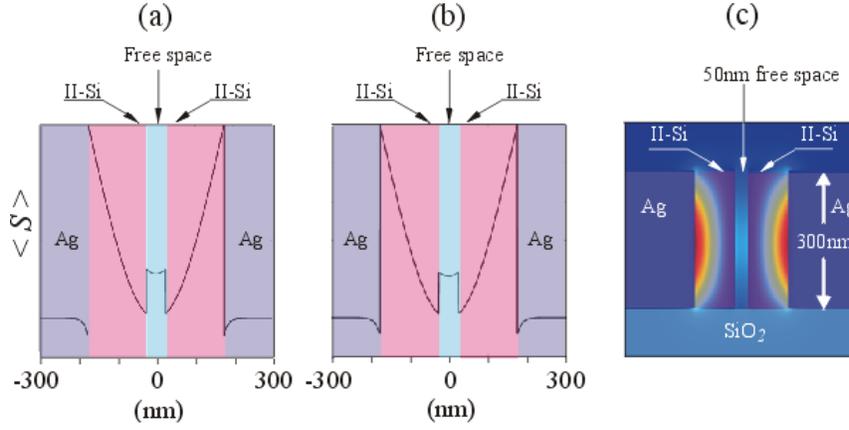


Figure 7.2: (a) Spatial distribution of time-averaged power, $\langle S \rangle$ normalized to 1W per unit width in the y -direction for 1550nm radiation within a five layer waveguide consisting of Ag(100nm) : Si(150nm) : free-space(50nm) : Si(150nm) : Ag(100nm) calculated using (a) the analytical model and (b) a finite element solution to Maxwell equations. The cross section is taken at the centre of the waveguide in panel (c). (c) The corresponding mode profile intensity (I) distribution inside the waveguide calculated using finite element method numerical solution to Maxwell's equations.

dispersion relation:

$$\tan\left(k_f \frac{a}{2}\right) = \frac{k_m}{k_f} \frac{1 - \frac{k_s}{k_m} \tan\left(k_s \frac{b-a}{2}\right)}{1 + \frac{k_m}{k_s} \tan\left(k_s \frac{b-a}{2}\right)} \quad (7.9)$$

for TE-polarized light. Equation 7.9 is obtained using a similar analysis employed in deriving the characteristic equation for the even TM mode (i.e. Eq. 7.7).

Results and Discussion

Since the nanoplasmonic waveguide explored is a planar structure having an infinite extent in one direction, it is necessary to investigate the influence of field confinement along the y -direction on the mode within the waveguide. Such analysis is important for realizing and fabricating a functional nanoplasmonic device (shown in Fig. 7.1(b)) as it provides minimum waveguide width, w , where the analytical expressions for the fields and the characteristic equations are applicable. To ascertain such analysis, it is necessary to evaluate the spatial distribution of time-averaged power, $\langle S \rangle$, for a finite width nanoplasmonic waveguide using the following expression

$$P_x = \int_{-\infty}^{\infty} \mathbf{S} \cdot d\mathbf{A} \quad (7.10)$$

$$\langle S \rangle = \frac{w}{2} \int_{-\infty}^{\infty} \text{Re}\{E_z H_y^*\} dz \quad (7.11)$$

where E_z and H_y^* are obtained from Eqs. 7.5 and 7.6, and compare it with the time-averaged power calculated from a finite element method numerical solution to Maxwell equations for a waveguide of finite width, w . Figures 7.2(a) and 7.2(b) depict the spa-

tial distributions of $\langle S \rangle$ for $\lambda = 1550\text{nm}$ radiation within a five-layer waveguide consisting of Ag(100nm):Si(150nm):free-space(50nm):Si(150nm):Ag(100nm) and a width of $w = 300\text{nm}$ calculated using the two methods. The numerical solution is in good agreement with that obtained from Eq. 7.11 indicating that due to strong confinement of the radiation in the Si layers, the model is also applicable for small waveguide widths. Figure 7.2(c) shows the 2D intensity distribution of the waveguide, from which the strong radiation confinement of the $\lambda = 1550\text{nm}$ radiation within the II-Si layers is evident.

When compared to the TM mode profile, it is found that the TE mode provides only 53% of the power confinement to the Si layers. Since optical excitation modifies the waveguide attenuation and the TM mode confinement, it is necessary to examine the temporal evolution of both the real, $\eta_s(t)$, and imaginary, $\kappa_s(t)$, components of the refractive index of the Si layers first. Notably, significant photoinjection of e-h pairs at a density $n(t) \geq n_c$ influences the TM mode propagation in a similar fashion as altering the Si layers' thickness. Figure 7.3(a) presents the real and imaginary parts of the refractive index in the first 6ps after excitation at various photoexcitation carrier densities ranging from $n_0/n_c = 0.05$ to 0.2 (i.e. $(0.43 - 1.7) \times 10^{20}\text{cm}^{-3}$). The strong dependence on n_0/n_c , and thus on the control pulse intensity, is evident in both $\eta_s(t)$ and $\kappa_s(t)$. For example, at a representative value of excitation of $n_0/n_c = 0.15$, the peak of $\eta_s(t)$ has decreased by 13%, while $\kappa_s(t)$ has increased dramatically by two orders of magnitude from 10^{-4} to 3.76×10^{-2} . Both $\eta_s(t)$ and $\kappa_s(t)$ recover back to their initial values within $\Delta t = 5\text{ps}$.

Of particular interest are the phase, $\beta_r(t)$, and the attenuation, $\beta_i(t)$, constants of the waveguide at signal wavelength of $\lambda = 1550\text{nm}$. The temporal evolution of these constants is determined by solving the dispersion relation (Eq. 7.7) at the time instant during the evolution dynamics of $n(t)$. However, since both $\beta_r(t)$ and $\beta_i(t)$ are strongly dependent on the waveguide dimensions, a and d , and on the density of the photoinjected carriers, n_0 , in the following analysis, $\beta_r(t)$ and $\beta_i(t)$ are evaluated for the following representative sets of a and d values: $(a, d) = (0\text{nm}, 150\text{nm})$, $(50\text{nm}, 150\text{nm})$, and $(100\text{nm}, 150\text{nm})$. The configuration $(a, d) = (0\text{nm}, 150\text{nm})$ represents a three-layer metal-Si-metal plasmonic waveguide. Figure 7.3(b) illustrates $\beta_r(t)$ and $\beta_i(t)$ in the first 6ps at the same photoexcitation carrier density range from $n_0/n_c = 0.05$ to 0.2 for the case $(a, d) = (50\text{nm}, 150\text{nm})$. Generally, the time evolution of $\beta_r(t)$ and $\beta_i(t)$ exactly follows that of $\eta_s(t)$ and $\kappa_s(t)$, indicating that by using an II-Si having $\tau_r = 1\text{ps}$, the recovery time of the device is $\Delta t = 5\text{ps}$. A small perturbation to the II-Si layer permittivity manifests itself as an enhanced attenuation. Thus, by operating the plasmonic waveguide in transmission mode, the dominant effect on the modulated $\lambda = 1550\text{nm}$ radiation is amplitude attenuation resulting from FCA due to $\beta_i(t)$.

While the plasmonic waveguide loss is activated with the application of an ultrafast control pulse, it is important to ensure that the device operates in the linear regime over a wide range of excitation intensities. With the aid of Fig. 7.4, this is better illustrated by plotting the minimum phase, $\beta_r^{\min}(t)$, and maximum attenuation, $\beta_i^{\max}(t)$

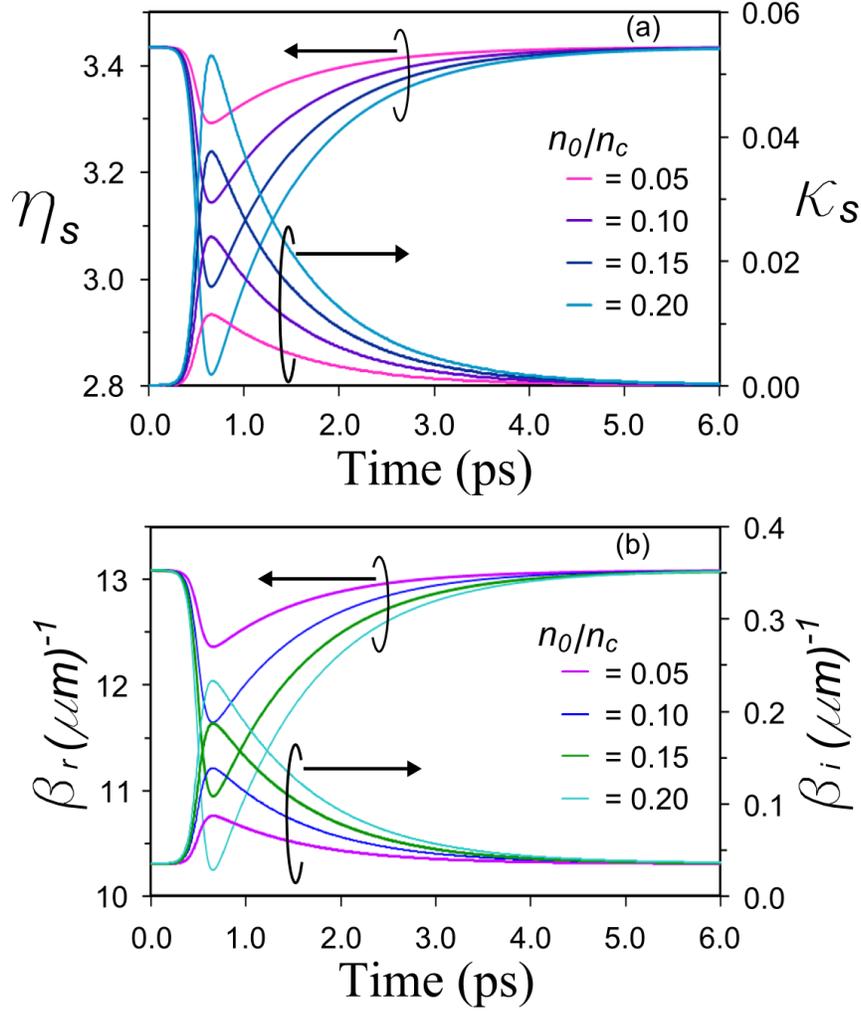


Figure 7.3: (a) Real, $\eta_s(t)$, and imaginary, $\kappa_s(t)$, values of the refractive index in the first 6ps after excitation at various photoexcitation carrier densities ranging from $n_0/n_c = 0.05$ to 0.2 . (b) $\beta_r(t)$ and $\beta_i(t)$ at the same photoexcitation carrier density range from $n_0/n_c = 0.05$ to 0.2 calculated for the $(a, d) = (50\text{nm}, 150\text{nm})$ waveguide configuration.

parameters for each set of (a, d) dimensions as a function of n_0/n_c . For waveguides $(a, d) = (0\text{nm}, 150\text{nm})$, $(50\text{nm}, 150\text{nm})$, and $(100\text{nm}, 150\text{nm})$ and at a low photoexcited carrier density, $n_0/n_c \leq 0.1$, $\beta_r^{min}(t)$ and $\beta_i^{max}(t)$ vary linearly with n_0/n_c . Within $0.1 \leq n_0/n_c \leq 0.2$ the deviation from linearity is less than 12%. Hence, the range of operation is set to be below $n_0/n_c \leq 0.2$.

Notably, prior to turning on the $\lambda = 1000\text{nm}$ control pulse (i.e. $n_0/n_c = 0$), using the values for $\beta_i^{max}(t = 0)$, long propagation distances of $21.7\mu\text{m}$, $14.2\mu\text{m}$, and $13.2\mu\text{m}$ corresponding to $(a, d) = (0\text{nm}, 150\text{nm})$, $(50\text{nm}, 150\text{nm})$, and $(100\text{nm}, 150\text{nm})$ plasmonic waveguides, respectively, are obtained. For the $\lambda = 1000\text{nm}$ control pulse wavelength, the propagation lengths are found to be $16.5\mu\text{m}$, $16.9\mu\text{m}$, $16.3\mu\text{m}$ for the same order of the waveguides. Thus, to ensure that sufficient free-carriers are excited throughout the device, one must choose the waveguide's length, $L \leq 16.5\mu\text{m}$, $16.9\mu\text{m}$, $16.3\mu\text{m}$ for each

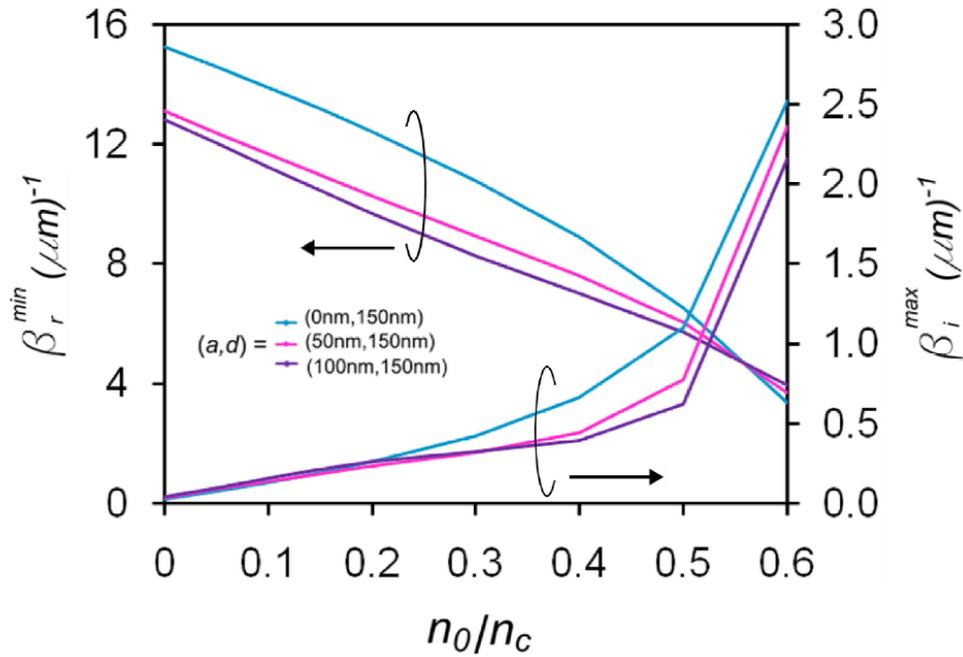


Figure 7.4: Minimum phase, β_r^{min} , and maximum attenuation, β_i^{max} , parameters for each set of (a, d) dimensions as a function of n_0/n_c from 0 to 0.2, for the waveguides having $(a, d) = (0\text{nm}, 150\text{nm})$, $(50\text{nm}, 150\text{nm})$, and $(100\text{nm}, 150\text{nm})$.

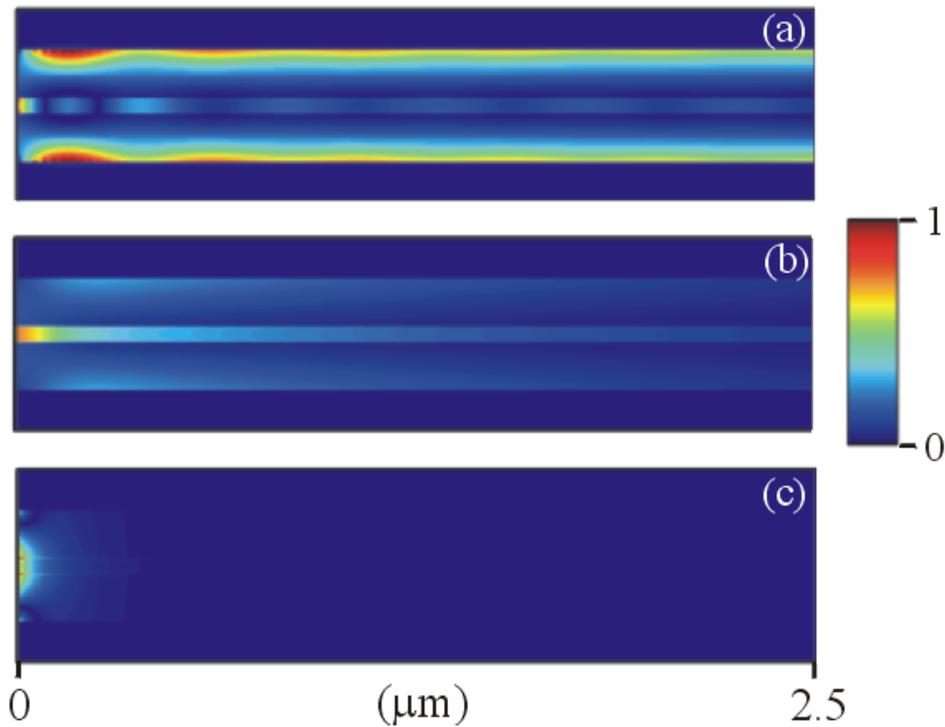


Figure 7.5: FDTD intensity (I) distribution of $\lambda = 1550\text{nm}$ signal laser at (a) $n_0/n_c = 0$, (b) $n_0/n_c = 0.5$, and (c) $n_0/n_c = 1.1$ for a $\text{Ag}(100\text{nm})\text{:Si}(150\text{nm})\text{: free-space}(50\text{nm})\text{:Si}(150\text{nm})\text{:Ag}(100\text{nm})$ waveguide.

corresponding (a, d) waveguide set.

As the control pulse intensity increases beyond a critical value, the change in the intensity distribution of the mode along the transverse direction also changes. For illustrative purposes, we present the calculations using 2D-FDTD for the $(a, d)=(50\text{nm}, 150\text{nm})$ plasmonic waveguide. Shown in Fig. 7.5 is the intensity distribution of the $\lambda = 1550\text{nm}$ light signal at three excitation electron densities of $n_0/n_c = 0, 0.5,$ and 1.1 . It should be noted that prior to excitation ($n_0/n_c = 0$), the $\lambda = 1550\text{nm}$ radiation propagates predominately in the II-Si layers as depicted in Fig. 7.5(a); however, for higher excitations ($n_0/n_c \geq 1.1$) the II-Si permittivity becomes metal-like and the mode propagates in the free-space region only as shown in Fig. 7.5(c). In this situation, the Si layers effectively ‘disappear’ and the waveguide properties mimic that of the metal-free-space-metal waveguide described by Eq. 7.8. At the highest excitation density, $n_0/n_c = 1.1$, the attenuation constant at the peak of the control pulse (i.e at $t = t_0$) is $3.642\mu\text{m}^{-1}$, yielding a propagation distance of only 137nm at $\lambda = 1550\text{nm}$.

To demonstrate the ultrafast operation of the plasmonic waveguide as an ultrawide bandwidth amplitude modulator, we investigate three devices each of length $L = 15\mu\text{m}$: (i) a three-layer Ag(100nm):Si(300nm):Ag(100nm) waveguide and (ii) two configurations of a five-layer waveguide: Ag(100nm) : Si(150nm) : free-space(50nm) : Si(150nm) : Ag(100nm) and Ag(100nm) : Si(150nm) : free-space(100nm) : Si(150nm) : Ag(100nm). Figs. 7.6(a)-6(c) illustrate the normalized output intensity of the ultrafast plasmonic switch as a function of time for these three modulators. The modulators’ responses are calculated at photoinjected carrier densities $n_0/n_c = 0$ (‘on’ state), 0.05, 0.1, 0.15, and 0.2. Assuming a 100% photon to electron generation quantum efficiency when the device is excited and waveguide dimensions of $300\text{nm} \times 300\text{nm} \times 15\mu\text{m}$, the required energy per pulse in each case corresponds to 16.6pJ, 33.2pJ, 49.8pJ and 66.4pJ, respectively. It is evident from the figure that in response to the $\lambda = 1000\text{nm}$ control pulse excitation, the $\lambda = 1550\text{nm}$ signal intensity drops sharply from the ‘on’ to the ‘off’ state in $\sim 100\text{fs}$ (following the rise time of the control pulse) and recovers back to the initial ‘on’ state within $\sim 5\text{ps}$. To assess the devices’ performance, the ‘on-off’ extinction ratios are calculated as a function of the photoinjected carrier densities as shown in Fig. 7.6(d). As expected from the increasing loss at high excitation, the extinction ratio increases with increasing photoinjected density. However, it is imperative to realize that the optimum operating conditions should combine the practicality of a low photoinjection carrier density (i.e. low control pulse energy) with the benefits of a high extinction ratio. At a maximum photoinjected density of $n_0/n_c = 0.2$, an extinction ratio of 30.3dB, is achieved using the three layer Ag(100nm):Si(300nm):Ag(100nm) waveguide. However, the Ag(100nm):Si(150nm):free-space(50nm):Si(150nm):Ag(100nm) and Ag(100nm):Si(150nm):free-space(100nm):Si(150nm):Ag(100nm) waveguides exhibit extinction ratios of 30.6dB and 35.3dB, respectively. Higher extinction ratios can be obtained by increasing the density of the photogenerated e-h density at the expense of a

higher energy budget. These extinction levels can be achieved in realistic devices fabricated via deposition and etching of II-Si and silver layers on top of a SiO₂ layer. Finally, the free-space region can be fabricated via focussed ion beam milling.

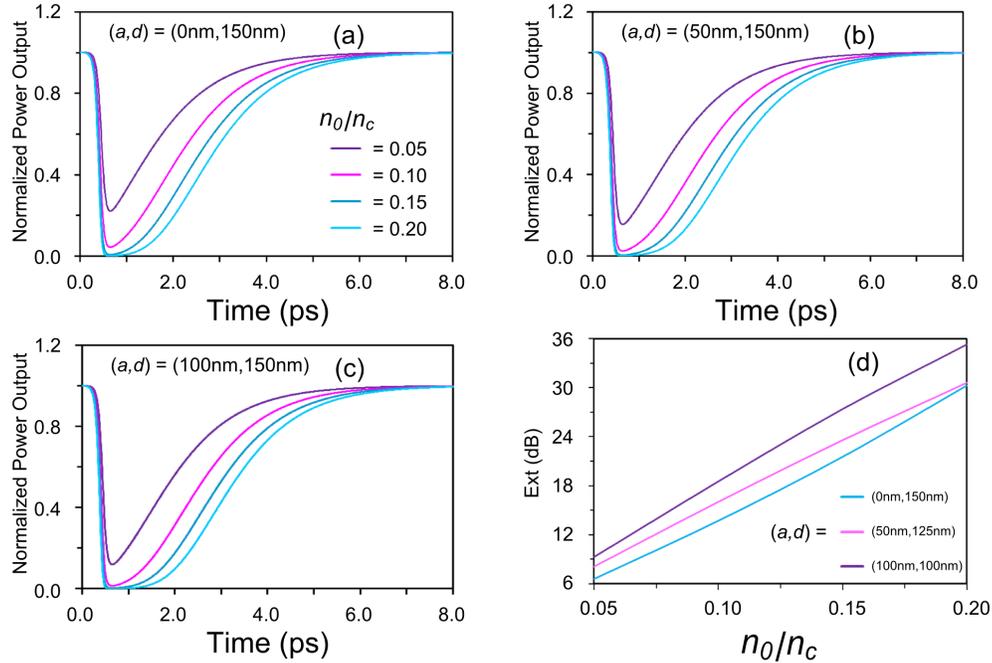


Figure 7.6: (a) The normalized output intensity of the ultrafast plasmonic switch as a function of time for three modulators. (a) Ag(100nm) : Si(300nm) : Ag(100nm) waveguide, (b) Ag(100nm) : Si(150nm) : free-space(50nm) : Si(150nm) : Ag(100nm), and (c) Ag(100nm) : Si(150nm) : free-space(100nm) : Si(150nm) : Ag(100nm). Each waveguide modulator configuration is calculated for photoinjected carrier densities of $n_0/n_c = 0$ (on state), 0.05, 0.1, 0.15, and 0.2. (d) On-off extinction ratios are calculated as function of the photoinjected carrier densities for the same nanoplasmonic waveguides.

7.2.4 The Plasmonic Nanoring Resonator

Geometry

³The nanoring resonator and associated circuitry are based on dielectric-loaded plasmonic waveguides. Signals in the S-band are transmitted through Si-loaded plasmonic waveguides, whereas above-bandgap ultrafast optical pulses are routed by SiO₂-loaded plasmonic waveguides. A schematic of the device is shown in Fig. 7.7. The plasmonic waveguides are formed on a silver film having a thickness, $t_{Ag} = 100\text{nm}$, and the width of the waveguides is uniform, $w_{Si} = w_{SiO_2} = 100\text{nm}$, throughout the device. Based on iterative design discussed in a subsequent section, the nanoring is given a radius, $r = 560\text{nm}$, which corresponds to a footprint area of $1.00\mu\text{m}^2$. An input bus plasmonic waveguide carrying S-band signals couples radiation to the resonator and is separated from the resonator

³These simulations were performed with assistance from David Driedger.

by a gap, $g_{Si} = 25\text{nm}$. Above-bandgap $\tau_p = 10\text{fs}$ optical pulses centred at $\lambda = 800\text{nm}$ are coupled into the resonator with a waveguide that is concentric to the resonator and offset from it by a gap, $g_{SiO_2} = 20\text{nm}$.

These devices can be fabricated by depositing a silver film onto a SiO_2 substrate. The Si waveguides can be created by lithographically defining a mask, dry etching through the silver film, depositing $\alpha\text{-Si}$, and removing the deposited film in the masked areas. A second layer of lithography can be performed to define the SiO_2 -loaded plasmonic waveguides in the same manner. Notably, any misalignment in this layer will alter the coupling efficiency of the above-bandgap pump pulses. Nanofabrication techniques for such small feature sizes are possible using electron beam lithography and systematic process development [28–30].

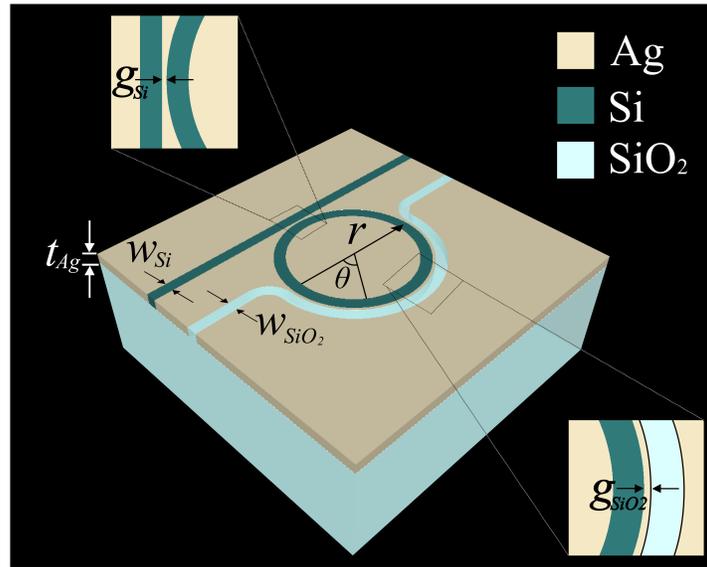


Figure 7.7: Schematic depiction of the device geometry. The nanoring is designed to have a radius, $r = 560\text{nm}$, silver film thickness, $t_{Ag} = 100\text{nm}$, input coupler separation, $g_{Si} = 25\text{nm}$, modulation coupler separation, $g_{SiO_2} = 20\text{nm}$, and uniform waveguide widths, $w_{Si} = w_{SiO_2} = 100\text{nm}$.

Passive Operation

Due to the sub-diffraction dimensions of this nanoplasmonic device and its complex geometry, full-wave numerical simulations are used as the primary design tool. Three-dimensional finite-difference time-domain (FDTD) simulations with broadband dielectric functions fit to experimental data were used to model the behaviour of the resonator. To understand the passive operation of the nanoring structure, the Si bus plasmonic waveguide was excited by a broadband pulse centred at $\lambda = 1550\text{nm}$, a full-width at half-maximum, $FWHM = 390\text{nm}$, and the plasmonic mode profile shown in Fig. 7.8(a). This mode has an effective refractive index of $n_{eff,Si} = 3.382$ at $\lambda = 1515\text{nm}$ and undergoes a propagation loss of $4.371\text{dB}/\mu\text{m}$. Although the loss of this mode is quite high, it has been shown that plasmonic devices can be coupled together with low-loss photonic waveguides

with coupling efficiencies up to 80% [31–35]. This scheme can be used to propagate signals over long distances where plasmonic waveguides may be too lossy. The transmission spectrum of the nanoring resonator is calculated by recording the electric fields in the Si bus plasmonic waveguide at a distance of $1.0\mu\text{m}$ before the nanoring resonator and $0.9\mu\text{m}$ beyond the nanoring resonator. By Fourier transforming the time domain signals, and normalizing the spectrum transmitted beyond the nanoring resonator to the spectrum that excites the nanoring resonator, the broadband transmission spectrum is obtained and is shown in Fig. 7.8(b). By iterative design of the nanoring resonator radius, r , and coupling gaps, g_{Si} and g_{SiO_2} , a transmission minimum is obtained at a wavelength of $\lambda = 1515\text{nm}$ for a ring radius of $r = 560\text{nm}$.

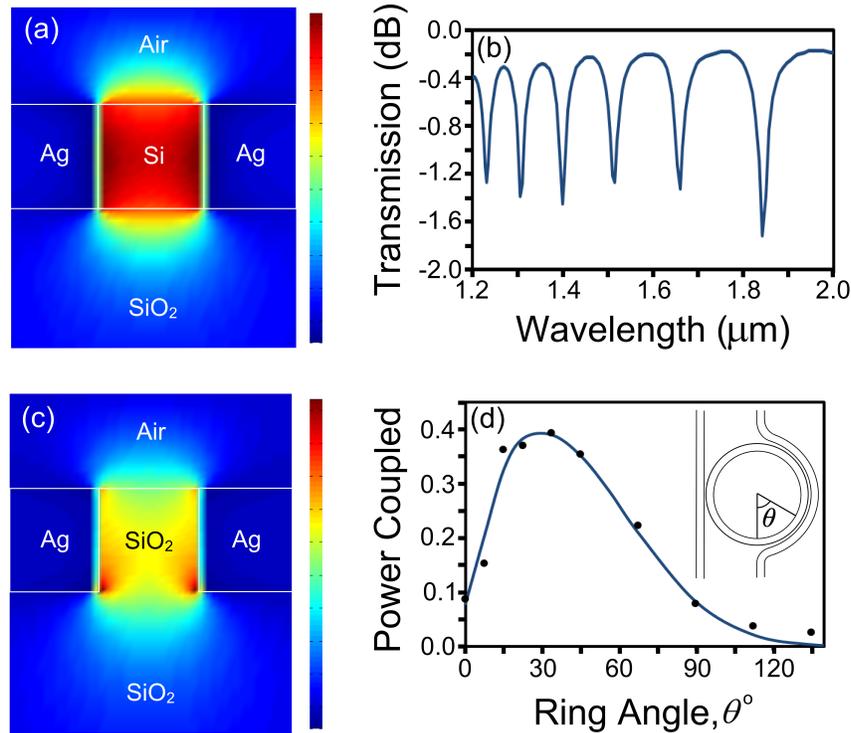


Figure 7.8: (a) Mode profile intensity (I) distribution of the excited mode in a Si-loaded plasmonic waveguide at $\lambda = 1515\text{nm}$. (b) Broadband transmission through Si bus plasmonic waveguide coupled to nanoring resonator. (c) Mode profile intensity (I) distribution of the excited mode in a SiO_2 -loaded plasmonic waveguide at $\lambda = 800\text{nm}$. (d) Pump power ($\lambda = 800\text{nm}$) coupled to nanoring versus nanoring angle, θ (see inset). A skewed Gaussian function is fitted to the recorded points.

The performance of the SiO_2 coupling plasmonic waveguide is investigated by exciting the SiO_2 plasmonic waveguide with a broadband pulse centred at $\lambda = 800\text{nm}$, a full-width at half-maximum, $FWHM = 94\text{nm}$, and the plasmonic mode profile shown in Fig. 7.8(c). This mode has an effective index of $n_{eff, \text{SiO}_2} = 1.742$ at a pump wavelength of $\lambda = 800\text{nm}$ and undergoes a propagation loss of $2.292\text{dB}/\mu\text{m}$. The electric fields are recorded at ten ring angles in the range $0^\circ \leq \theta \leq 135^\circ$ and the power, P , coupled to the ring at each angle is calculated, as shown in Fig. 7.8(d). These data points are fit with a skewed Gaussian

function of the form,

$$P = \frac{a}{\sqrt{2\pi}} \exp\left[-\frac{b}{2}(\theta - c)^2\right] \times \left[1 + \operatorname{erf}\left[\frac{d}{\sqrt{2}}(\theta - e)\right]\right], \quad (7.12)$$

where $\{a, b, c, d, e\} = \{0.12, 28, 0.029, 1.26, 5\}$ and an R-squared fit, $R^2 = 0.93$. It is important to note that the power is significantly reduced beyond $\theta = 135^\circ$, and minimal power remains to propagate around the nanoring beyond the SiO₂ bus waveguide. For $0^\circ \leq \theta \leq 35^\circ$ the power coupled to the nanoring increases with increasing θ due to the increasing coupling length between the SiO₂ waveguide and the Si nanoring. However, the above-bandgap radiation undergoes substantial loss as it continues to propagate around the Si nanoring and gradually decays by $\theta = 135^\circ$.

Ultrafast Active Operation

As discussed above, the nanoring is designed with a transmission minimum at $\lambda = 1515\text{nm}$. By coupling an ultrafast above-bandgap radiation pulse into the nanoring, free-carriers are excited and the complex refractive index in the nanoring is altered, and thus, the transmission minimum is shifted to a new wavelength. To model the refractive index at each nanoring angle, θ , in the resonator at each instant in time, it is first necessary to determine the free-carrier density as a function of nanoring angle. The free-carrier density at each point in the nanoring is proportional to the intensity of above-bandgap radiation at that point. By selecting a maximum pump strength, n_0/n_c , coupled to the nanoring at $\theta = 35^\circ$, the carrier density at each nanoring angle, θ , can be calculated from Eq. 7.12. The time-dynamics of the free-carrier density at each angle can then be determined from Eq. 7.3, and finally, the time-dynamics of the permittivity (and therefore, refractive index) at each angle can be obtained from Eq. 7.1. Surface plots relating the real component of the refractive index to the nanoring angle, θ , and time, t , for pump strengths of $n_0/n_c = 0.05, 0.10, 0.15$, and 0.20 are shown in Figs. 7.9(a)-(d), respectively. Surface plots relating the imaginary component of the refractive index to the nanoring angle, θ , and time, t , for pump strengths of $n_0/n_c = 0.05, 0.10, 0.15$, and 0.20 are shown in Figs. 7.9(e)-(h), respectively. The time required to generate free-carriers and alter the refractive index is very short and is limited primarily by the duration of the excitation pulse. However, the time required for the material to recover to its non-excited state is longer and is limited by the e-h recombination time. After a recovery time of $\Delta t = 3\text{ps}$, the material properties have returned to their non-excited state.

By applying these results to the II-Si of the nanoring resonator, it is possible to calculate the change in power transmission as the pumping strength is increased for signal radiation of a fixed wavelength, $\lambda = 1515\text{nm}$. Substantial pump radiation is coupled to the nanoring for angles in the range, $0^\circ \leq \theta \leq 168.75^\circ$, and consequently, the refractive index of the nanoring will only undergo changes in these regions. The Si of the nanoring resonator that falls in the range, $0^\circ \leq \theta \leq 68.75^\circ$, is divided into twelve

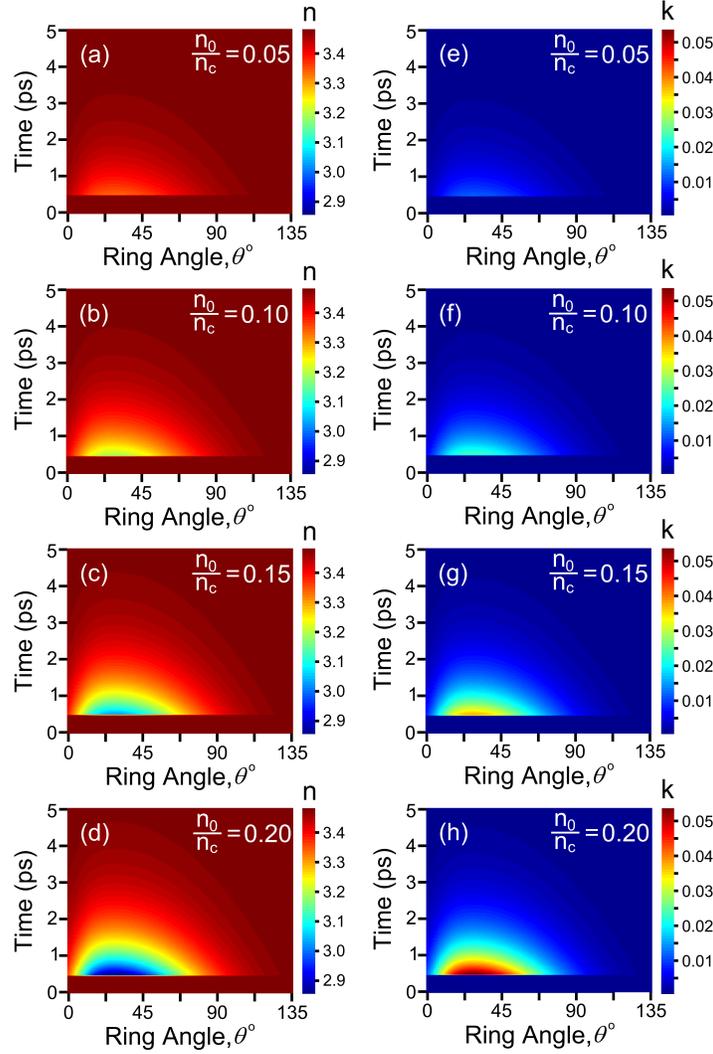


Figure 7.9: Refractive index of Si as a function of time and nanoring angle when excited by ultrafast above-bandgap pulses of $\tau_p = 10\text{fs}$ duration at $\lambda = 800\text{nm}$. The real component of the refractive index of II-Si in the ring is modelled for pump strengths of $n_0/n_c=0.05$, 0.10, 0.15, and 0.20 in (a), (b), (c), and (d), respectively. The imaginary component of the refractive index of II-Si in the nanoring is modelled for pump strengths of $n_0/n_c=0.05$, 0.10, 0.15, and 0.20 in (e), (f), (g), and (h), respectively.

segments, each of which is given its own optical properties based on data taken from Fig. 7.9. Finite-difference time-domain simulations are performed on ten different nanoring resonators, each with an angular-dependent refractive index representing a different pumping strength. Intensity distributions of the nanoring resonator and adjacent bus waveguides for pumping strengths of $n_0/n_c = 0.00$, 0.10, and 0.22 are shown in Figs. 7.10(a)-(c), respectively. In the absence of pumping, the electric fields add destructively in the II-Si bus plasmonic waveguide and a minimal signal is transmitted. However, as the pump strength increases, fields begin to add constructively in the II-Si bus plasmonic waveguide and a strong transmitted signal is observed for the case of $n_0/n_c = 0.22$. Figure 7.10(d) shows the influence of the pumping strength on the position of the transmission minimum. As the free-carrier density increases, the optical path length of the nanoring resonator decreases and the transmission minimum shifts to the blue part of the spectrum. As a result, the adjacent transmission maximum shifts closer to the signal wavelength of $\lambda = 1515\text{nm}$ and the transmitted power increases.

The dependence of power transmitted on the maximum free-carrier density in the nanoring is shown in Fig. 7.11(a), and is fit by a fifth-order polynomial with $R^2 = 1$. While a weak signal of 4.95% of the input signal strength is observed in the absence of pumping, the signal strength increases to 49% of the input signal strength for a pumping strength of $n_0/n_c = 0.22$. With knowledge of both the time-dynamics of free-carriers (from Eq. 7.3) and the power transmitted through the II-Si bus plasmonic waveguide as a function of the maximum free-carrier density in the nanoring, the time-dynamics of the power transmitted through the Si bus waveguide can be calculated by multiplying the data of Fig. 7.11(a) by the solution to Eq. 7.3. These dynamics are shown for pumping strengths of $n_0/n_c = 0.05$, 0.10, 0.15, and 0.22 in Fig. 7.11(b). Notably, the device is switched to the “on” state on a femtosecond timescale. After a time of $\Delta t = 3\text{ps}$, the device has completely recovered to the “off” state. Based on these characteristic rise and fall times, a modulation frequency of 0.33THz is possible using this structure.

As a conservative estimate, we can assume that angles in the range, $0^\circ \leq \theta \leq 180^\circ$, are pumped uniformly, which corresponds to a pumping volume, $V = 1.76 \times 10^{-14}\text{cm}^3$. Furthermore, a pumping photon with $\lambda_p = 800\text{nm}$ has an energy $E_p = h\lambda_p = (6.626 \times 10^{-34}\text{m}^2\text{kg/s})(800\text{nm}) = 2.48 \times 10^{-19}\text{J}/\text{photon}$. Therefore, the total energy required to photogenerate a free-carrier density of n_0 can be obtained from the equation,

$$E_{pump} = \eta E_p n_0 V, \quad (7.13)$$

where η is the quantum efficiency of the photogeneration process, which we assume to be unity, i.e. $\eta = 1$. An internal quantum efficiency of $\eta = 1$ has been demonstrated in bulk Si at $\lambda = 800\text{nm}$ and therefore, this is a reasonable assumption for above-bandgap radiation [36]. From Eq. 7.13, a pumping power of $P_{pump} = 0.88\text{pJ}$ must be coupled to the nanoring resonator to switch the device to the “on” state ($n_0/n_c = 0.22$).

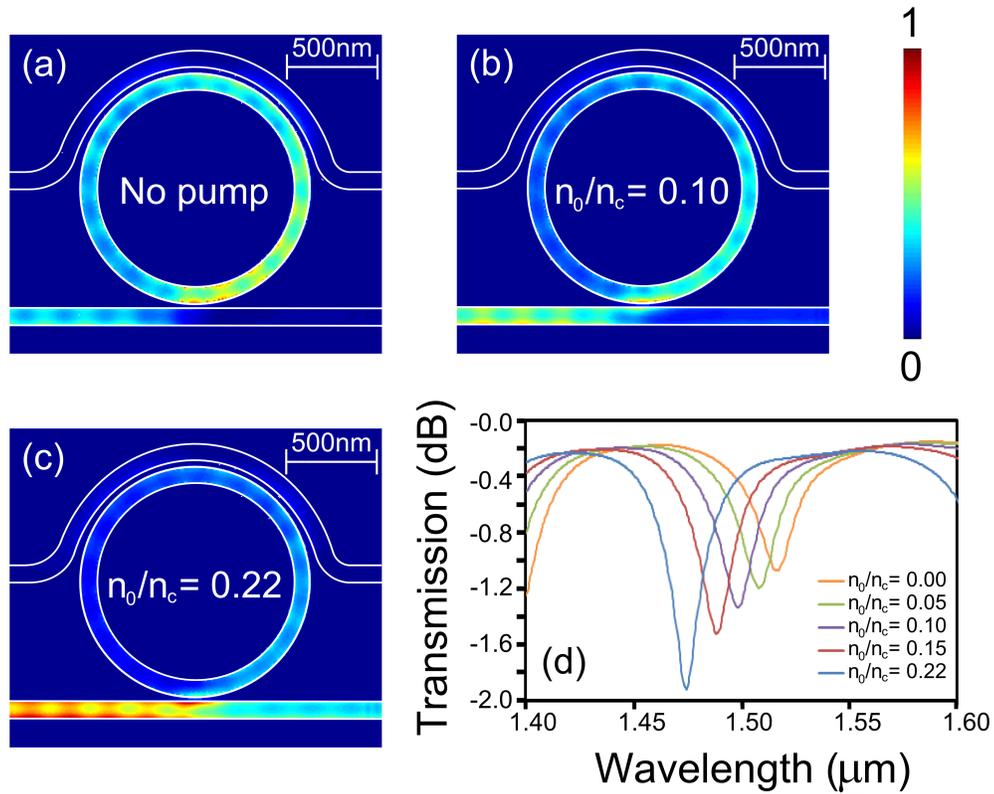


Figure 7.10: (a) Intensity (I) distribution of the nanoring resonator in the “off” state without any pump. (b) Intensity (I) distribution of the ring resonator for a pump strength of $n_0/n_c=0.10$. (c) Intensity (I) distribution of the nanoring resonator in the “on” state, with a pump strength of $n_0/n_c = 0.22$. (d) Effect of the pump strength on the position of the transmission minimum.

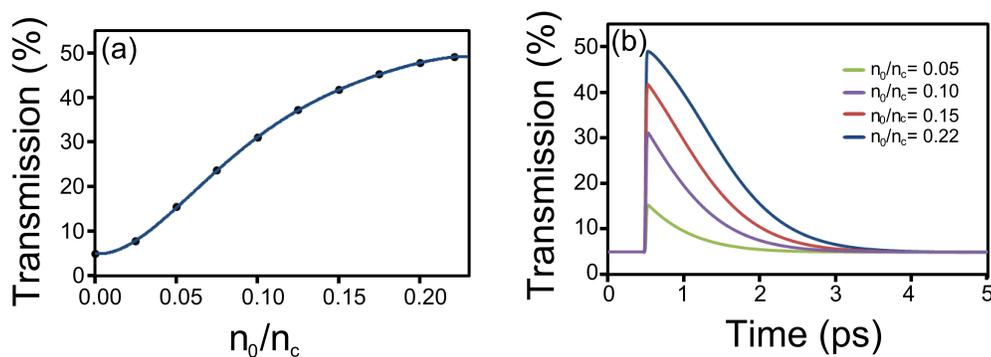


Figure 7.11: (a) Dependence of the power transmission through II-Si bus waveguide on the pump strength. (b) Power transmission through the II-Si bus waveguide as a function of time for pump strengths of $n_0/n_c = \{0.05, 0.10, 0.15, 0.22\}$.

7.2.5 Summary

We investigated a novel platform for ultrafast all-optical active control of nanoplasmonic signals based on optical free-carrier generation, absorption, and dispersion in Si-based nanoplasmonic waveguide structures. The waveguide structures can be fabricated with standard thin-film waveguide fabrication processes and can be integrated with existing nanophotonic device architectures. High field confinement at the Si-metal interfaces together with ultrafast photoexcitation of the Si layers results in enhanced single-photon absorption within the nanoplasmonic waveguide that leads to strong and ultrafast modulation of the signal beam. For applications to ultrafast switching and modulation, we have shown that the ultrahigh operation bandwidth is limited only by the duration of the control pulse and the carrier lifetime of the Si layers. Furthermore, the off-on extinction ratio or modulation depth is sufficiently high at low optical excitation, enabling the generation of ultrashort laser pulses and modulation with an extremely high signal-to-background ratio. In the case of a 5-layer slab waveguide with a footprint area of $A = 8.25\mu\text{m}^2$, FCA was used to achieve a 35dB on-off contrast ratio with a pulse energy of $E_p = 66.4\text{pJ}$. By considering a nanoring resonator, phase-sensitive modulation was achieved by FCD. This allowed the footprint area to be reduced to $A = 1.00\mu\text{m}^2$, and a 10dB on-off contrast ratio was demonstrated with a pulse energy of $E_p = 0.88\text{pJ}$.

7.3 Nanostructured Nanoplasmonic Antennas Operating in the Near- to Mid-Infrared

7.3.1 Introduction

⁴ Plasmonic antennas can be used to confine electromagnetic energy to a sub-wavelength scale in all three spatial dimensions [1, 37–40]. In general, these antennas are composed of one or more noble metal nanostructures that support a geometry-dependent particle plasmon resonance. When illuminated with electromagnetic radiation of preferential wavelengths, the noble metal nanostructures that compose the antenna polarize strongly and incident electric fields can be enhanced by two orders of magnitude in the antenna gap [41]. This enables access to sensitive light-matter interaction processes, with applications in sensing, microscopy, spectroscopy and nonlinear optics experiments [41–47].

In the case of nonlinear optics, the electric field enhancement enables more efficient access to nonlinear optical effects of materials in the immediate vicinity of the antenna gap. Whereas surface plasmon polaritons excited on thin films can enhance electric fields by a factor of ≈ 10 , nanoplasmonic antennas can enhance electric fields by a factor of

⁴The content of this section has been published. (1) S. Sederberg and A. Y. Elezzabi, *Opt. Express* 19, 10456-10461 (2011). (2) S. Sederberg and A. Y. Elezzabi, *Appl. Phys. Lett.* 98, 261105 (2011). (3) S. Sederberg and A. Y. Elezzabi, *Opt. Express* 19, 15532-15537 (2011).

$\approx 100 - 1000$ [48]. For $\chi^{(2)}$ nonlinear materials, the nonlinear polarization is proportional to $I^2 = E^4$. Therefore, an electric field enhancement of $100\times$ in the antenna gap would correspond to a $10^8\times$ stronger nonlinear interaction than what could be achieved with the free-space radiation alone. In the case of $\chi^{(3)}$ nonlinear materials, the nonlinear polarization is proportional to $I^3 = E^6$, and an electric field enhancement of $100\times$ in the antenna gap would correspond to a $10^{12}\times$ stronger nonlinear interaction. Therefore, nanoplasmonic antennas offer distinct advantages over plasmons excited on thin films or in waveguides, and will undoubtedly find applications in nonlinear nanoplasmonic circuitry. To allow for high-density optical circuitry, it is particularly important to reduce the antenna dimensions as much as possible, while maintaining strong electric field enhancement in the gap.

Typical antenna geometries consist of rods, spheres or triangles. Interestingly, the broadband spectral response of radio-frequency, microwave, and terahertz antennas has been shown to depend strongly on the antenna geometry [49, 50]. As such, the operating frequency of the antenna and other important antenna properties can be controlled by carefully designing its shape. In contrast to the vast development of microwave antennas, research in optical antennas is relatively at its infancy. This is largely because the fabrication and characterization technologies that are necessary for these nanometer length scale structures have only recently become accessible. Furthermore, metals behave differently at optical and near-infrared frequencies than at microwave frequencies, and most microwave design techniques are not valid for plasmonics antennas [51–53]. At microwave frequencies, metals are normally modelled as perfect electric conductors, where the negligible losses are analogous to the DC resistivity. In contrast, metal losses at optical or near-infrared frequencies result from damped electron oscillations and possibly interband transitions, and cannot be ignored. Furthermore, dispersion in the metals can no longer be neglected, particularly for visible frequencies in the vicinity of plasma resonances or interband transitions. These factors invalidate microwave analysis techniques, such as the perfectly conducting wire model and surface-impedance integral equations, and require more computationally intensive techniques, such as volume integral equations or the FDTD method.

The nanoscale antenna dimensions make these devices suitable for integration into a sophisticated standalone device that can be applied to spectroscopy [41, 43, 44], clinical diagnosis [41, 43, 44], or lithography [45]. Challenges arise when the experimenter wishes to observe a process in the near- or mid-infrared, such as molecular vibration modes, and antenna dimensions must be increased. For example, shifting the resonance from 800nm to $2.5\mu\text{m}$ increases the device footprint by an order of magnitude, hindering integration onto a chip. In the interest of harnessing light for efficient nanometer-resolution vibrational spectroscopy of molecules in the near- and mid-infrared, it would be advantageous to develop antennas to operate in the mid-infrared spectral range, yet maintain a nanoscale footprint and imaging resolution. Thus, it would be desirable to red-shift the

resonance of an antenna in a continuous and controllable manner over a wide range of wavelengths, while maintaining a constant antenna footprint. Rather than modifying the dielectric environment by introducing new materials (at the expense of more complicated nanofabrication), it is more attractive to change the geometry of the gold present to tune the antenna resonance conditions.

7.3.2 Simulation Technique

A home-made, multiprocessor, three-dimensional FDTD code (Appendix A) is used to simulate the plasmonic antennas. In each simulation, the antenna of interest is given the properties of gold and is placed on a SiO₂ substrate having a refractive index, $n_s = 1.5$. The Drude model is used to represent the permittivity function of gold over a wide spectral range. Drude parameters of $\varepsilon_\infty = 9.069$, $\tau_c = 8.669\text{fs}$, and $\omega_p = 1.354 \times 10^{16}\text{rad/s}$ are based on an optimal fit to well-established experimental data and ensure that the optical properties of gold are well-represented for the wavelengths of interest [54, 55]. The minimum feature size of gold is maintained at or above dimensions of experimentally-confirmed structures for this model [1]. Provided that the density of a thin gold film is the same as its bulk value, the optical properties of the film converge to the bulk properties for film thicknesses greater than approximately $f = 25\text{nm}$ [56]. Therefore, a gold film thickness of $f = 35\text{nm}$ is used for these investigations.

The antenna gap is held constant at $30\text{nm} \times 30\text{nm}$ for each structure and a spatial resolution of $\Delta x = \Delta y = \Delta z = 3\text{nm}$ is used to ensure that the electromagnetic fields in the antenna gap are accurately resolved. A time step of $\Delta t = 4.75\text{as}$ ensures stability at this resolution. Each triangle is connected to its neighbours by a thin strip to avoid geometric singularities, which also ensures excitation of a collective response from the entire structure (i.e. the triangles are all connected). The flare angle of the antenna is held constant at 90° , so that the length and width of the antenna are equal.

To obtain broadband characteristics of the antennas in a single simulation, each antenna is excited by a plane wave that has been temporally modulated by a Gaussian function, permitting access to wavelengths in the range, $550\text{nm} \leq \lambda \leq 4.0\mu\text{m}$. The antennas are excited from below the SiO₂ substrate and the electric fields are measured 3nm above the top surface of the antenna. The broadband enhancement factor of the antenna is determined by calculating the Fourier Transform of the time-domain signal and normalizing it to the spectrum of the input pulse. A perfectly-matched layer with a geometrically-graded conductivity is used to minimize non-physical reflections at the simulation space boundaries.

7.3.3 Fractal Geometry

Despite the prominent role that Euclidean geometry plays in mathematics, science, and engineering, there are very few objects in nature that it accurately describes. Many

naturally-occurring structures are composed of features with numerous length scales, where each length scale appears similar to the others when magnified by an appropriate factor. Common examples include a snowflake, a tree and a coastline. A mathematical abstraction of Euclidean geometry that acts as a much better descriptor of these self-similar or self-affine objects is fractal geometry [57]. Rather than being described by elementary shapes such as polygons and arcs, a fractal is defined by a basis shape and an operator that acts iteratively on that shape. Applying the operator to the basis shape produces the first-iteration shape; applying it again to the first-iteration shape produces the second-iteration shape; and so on. When carried out an infinite number of times, the resulting shape can exhibit interesting properties, such as zero area and infinite perimeter simultaneously. Moreover, as a fractal structure is magnified further and further, the new details that appear have the same form as those in the parent structure. In this respect, fractals are said to be self-similar on multiple length scales.

Fractals exist in a space that is characterized by a non-integer, or fractional dimension, known as the Hausdorff dimension [58]. Hausdorff dimension describes how the length scale decreases as the iteration increases and how the number of features composing the fractal multiplies as the iteration increases. Hausdorff dimension is defined as [57, 58] :

$$D = \frac{\log(N)}{\log(S)}, \quad (7.14)$$

where N is the factor by which the number of triangles increases after each iteration, and S is the length scaling factor by which the triangles are reduced.

One well-known fractal is the Sierpiński triangle, shown in Fig. 7.12(a)-(d), whose basis shape is a triangle. In going from the basis shape in Fig. 7.12(a) to the first iteration shape in Fig. 7.12(b), three triangles, each one-quarter of the area of the original triangle, are arranged to fill the outline of the original triangle. In subsequent iterations, this same operation is applied to each triangle that composes the fractal. The common bowtie antenna is shown in Fig. 7.16(a). Placing two Sierpiński triangles next to one another produces a structure that is similar to the bowtie antenna, but contains fractal features, as shown in Fig. 7.16(b).

In electromagnetic theory, the manner in which an object scatters an incident photon is strongly dependent on the size of that object relative to the wavelength spectrum that composes the photon. The presence of numerous length scales in fractal structures makes their interaction with electromagnetic radiation intriguing and suggests that the object can act preferentially in multiple bands of wavelengths. Furthermore, a fractal can have a much larger perimeter-to-footprint ratio than its basis shape, leading to red-shifted resonance conditions and compact devices. Numerous microwave antenna designs incorporating fractal geometry have been investigated both theoretically and experimentally and appealing properties including a compact footprint and multiband operation have been successfully demonstrated [59]. Fractal antennas and apertures have also been in-

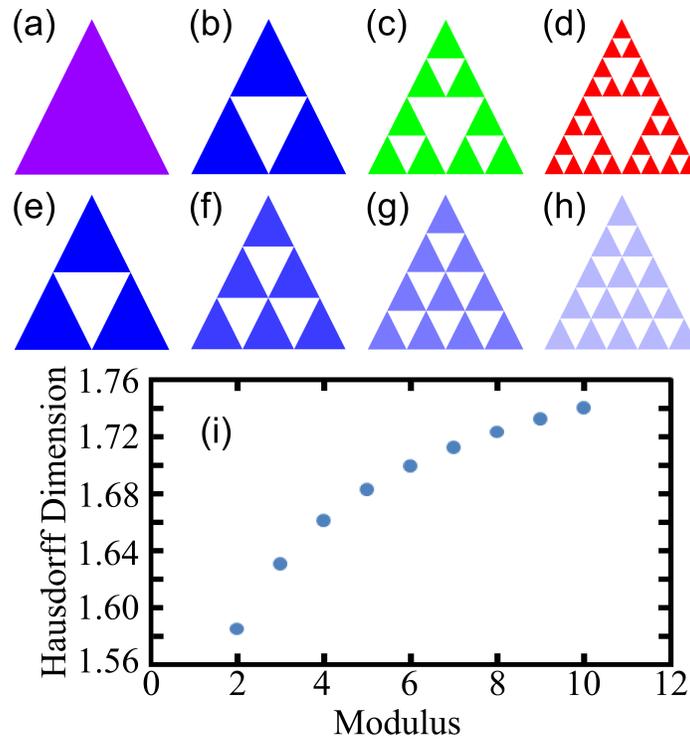


Figure 7.12: (a) Sierpiński triangle basis shape; and (b)-(d) first three iterations of Sierpiński triangle. (e)-(h) Modulus $M = 2, 3, 4$ and 5 Pascal triangles, respectively. (i) The relationship between modulus and Hausdorff dimension for a Pascal triangle.

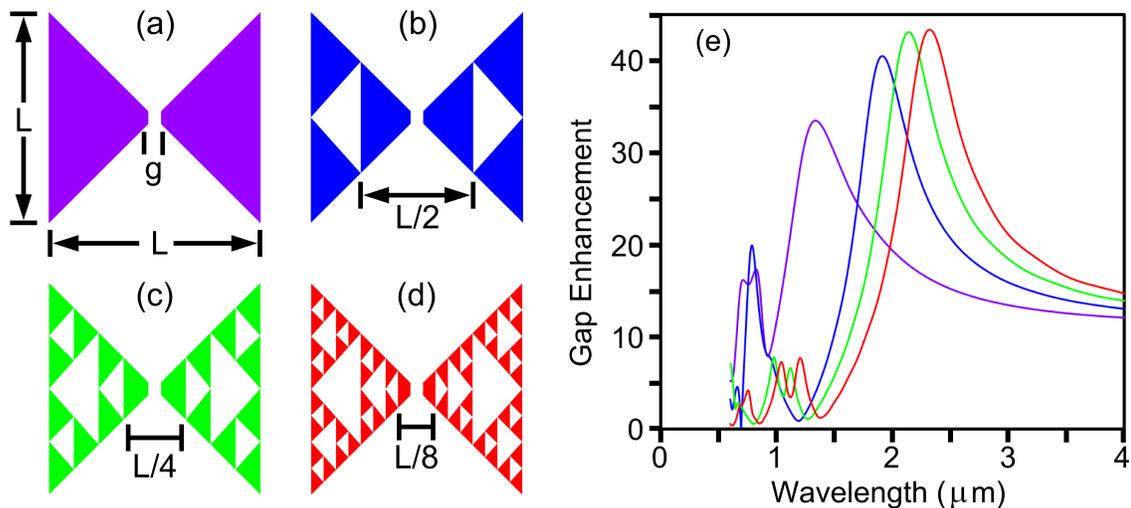


Figure 7.13: Schematic representations of the simulation geometries: (a) bowtie antenna and (b) first-, (c) second- and (d) third-iteration Sierpiński fractal plasmonic antennas. The broadband electric field enhancement factor is shown in (e). The antenna colour in (a)-(d) corresponds to the line colour in (e).

investigated at terahertz frequencies with similar results [60–63]. It has been demonstrated that nanoscale fractal apertures allow for enhanced transmission and resolution for near- to mid-infrared wavelengths, but multiband behaviour is less pronounced at this spectral range, due to higher metal losses [64, 65].

7.3.4 Sierpiński Fractal Nanoplasmonic Antenna

An antenna design that has undergone intensive research and applications at optical and near-infrared wavelengths is the bowtie antenna, which is composed of two triangles and is shown in Fig. 7.13(a). The Sierpiński triangle is a class of fractal that uses triangles as a basis shape. Placing two Sierpiński triangles together forms a modified bowtie antenna, which will be referred to as the Sierpiński Fractal Plasmonic Antenna (SFPA). The first three iterations of SFPA are shown in Figs. 7.13(a)-(c), respectively. Clearly, the size and number of triangles present and their arrangement changes as the fractal iteration is increased. Undoubtedly, the ability of the antenna to polarize will depend strongly on how the constituent triangles are arranged. Hence, it is reasonable to state that as the fractal iteration is varied, the broadband antenna impedance will change and the resonant wavelength will shift. The remainder of this investigation is devoted to understanding this phenomenon.

The broadband electric field enhancement factor of each antenna iteration with a constant length, $L = 475\text{nm}$, is shown in Fig. 7.13(e). There are several observations that can be drawn from these spectral response plots. As the iteration of SFPA increases from zero to three, the wavelength of the main resonance is red-shifted from $1.40\mu\text{m}$ to $2.44\mu\text{m}$ and the enhancement factor increases from 37 to 49. In addition, there is a rearrangement of the minor resonances that occur at shorter wavelengths. As shown in Fig. 7.13(e), the bowtie antenna has minor resonances at $\lambda = \{560, 822\}\text{nm}$, whereas the first iteration SFPA has minor resonances at $\lambda = \{554, 660, 787\}\text{nm}$. It should be pointed out that these resonances are much weaker than the main resonance and do not effectively confine incident radiation to the gap. Therefore only the main resonance will be discussed in the remainder of this section.

In the interest of developing empirical design rules describing the operation of these antennas, the resonant wavelength and resonant enhancement factor have been plotted versus antenna length for each of the four structures under consideration in Fig. 7.14. From these plots, it is evident that the resonant wavelength scales linearly with the antenna length and that increasing the Sierpiński fractal iteration red-shifts the resonance regardless of the antenna length. One observation that can be drawn from these plots is that the factor by which the resonant wavelength of the first-iteration antenna is red-shifted when compared to a bowtie antenna of the same length is virtually constant and is approximately 1.41. A similar observation can be made for the second- and third-iteration structures, which are red-shifted by factors of 1.57 and 1.68, respectively, relative to a

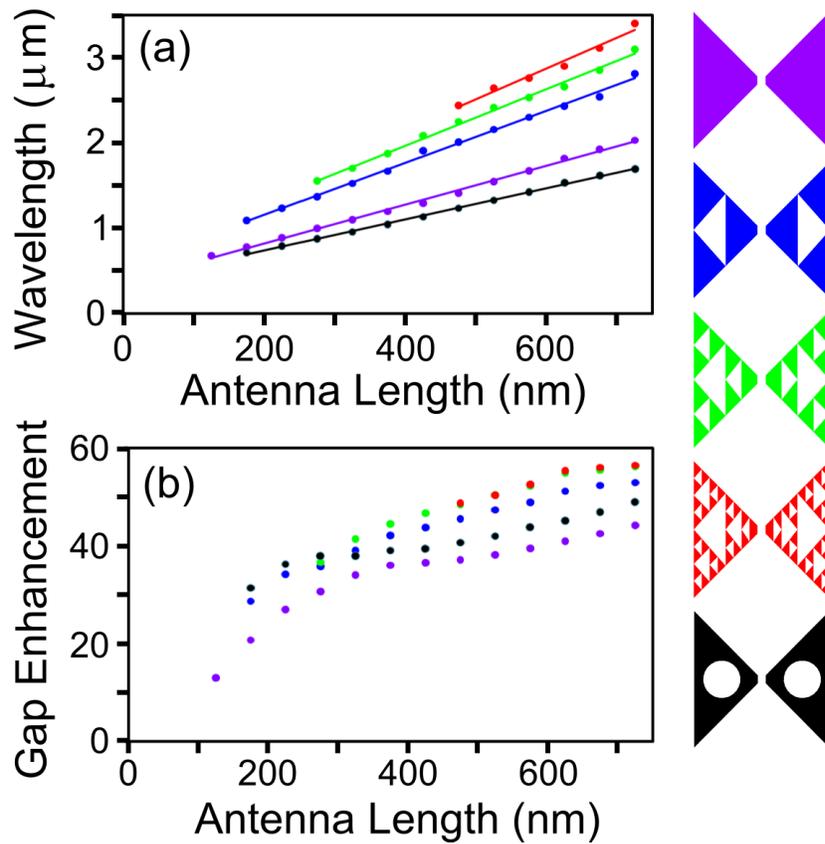


Figure 7.14: Empirical relationships between antenna length and (a) resonant wavelength and (b) resonant enhancement factor for each antenna geometry under consideration. The colour of antennas on the right corresponds to the line colours in (a) and (b).

bowtie antenna of the same length. To verify that fields are confined to the antenna gap and not at other locations of the antenna, the intensity distribution at resonance is calculated for each antenna structure and is shown in Fig. 7.15. For each SFPA iteration, the incident radiation is confined to the antenna gap and transmission through other triangular holes in the antenna is negligible. The full-width half-maximum is calculated along the vertical direction at the centre of each antenna gap and is found to be 37.86nm, 37.95nm, 37.90nm, and 37.86nm for the bowtie, first, second, and third iteration antennas with a constant length, $L = 475\text{nm}$, respectively. Remarkably, as the fractal iteration increases, the resonant wavelength of the antenna is red-shifted, yet allows for confinement to the same nanoscale dimensions. In the case of the bowtie antenna with $L = 475\text{nm}$, incident fields are confined to $\lambda/47$ and enhanced by a factor of 37. In comparison, the third iteration SFPA confines incident fields to $\lambda/80$ and enhances them by a factor of 49.

Geometric Investigation

In an effort to understand the source of the observed red-shift, it is natural to examine the antenna geometry. Table 7.1 summarizes some of the key geometric properties of each antenna. As the iteration increases, the total volume of gold in the antenna decreases and

the total perimeter of the gold features increases. Besides these simple geometric parameters, it is important to consider the possibility that it is the self-similar arrangement of triangles and their ability to polarize that may be the cause of this behaviour. To understand the importance of each of these parameters (perimeter, area and polarizability), a subsequent series of simulations was performed. A non-fractal test-case was designed to compare with the first iteration antenna. Rather than removing a triangle from the bowtie antenna, a circle with the same area as the triangle was removed. The geometric properties of this structure are shown in Table 1. Several lengths of this structure were simulated, and plots of the resonant wavelength and enhancement factor versus antenna length are shown in Fig. 7.14, along with the previous results for comparison. Interestingly, the circular structure consistently produces a blue-shifted resonance when compared to the bowtie antenna. The resonant enhancement factor is typically 10% less than for the first iteration SFPA. Even though the area and perimeter properties of the circular structure are comparable to the fractal structure, the performance is inferior, demonstrating the unique ability of the fractal geometry to produce red-shifted resonances and large confinement factors.

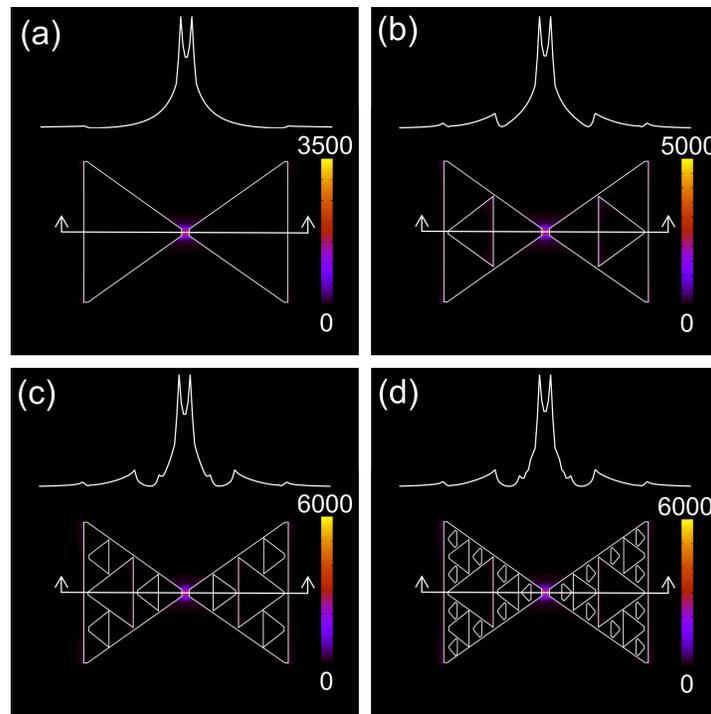


Figure 7.15: Intensity (I) distributions at resonance for (a) bowtie antenna and (b) first-, (c) second- and (d) third-iteration Sierpiński fractal plasmonic antennas. The line section shows the intensity (I) distribution along the x -axis of the antenna and within the antenna gap.

Antenna	Gold Area	Gold Perimeter	Red-Shift Factor
Bowtie	A	P	1.00
First	$0.750 \times A$	$1.50 \times P$	1.41
Second	$0.563 \times A$	$2.25 \times P$	1.57
Third	$0.422 \times A$	$3.38 \times P$	1.68
Circle	$0.750 \times A$	$1.37 \times P$	0.87

Table 7.1: Geometric Properties of Antennas under Investigation

7.3.5 Pascal Triangle Nanoplasmonic Antenna

In performing an investigation into the relationship between Hausdorff dimension and antenna performance, it would be ideal to keep the antenna outline constant while modifying only the internal structure of the antenna to vary the Hausdorff dimension. A geometric abstraction of the Sierpiński triangle that satisfies this requirement is a class of triangles known as the Pascal triangles. The first four moduli of Pascal triangles are shown in Fig. 7.12(e)-(h). As the modulus, M , increases, the size of triangle that makes up the structure decreases and the number of triangles composing the structure increases. The Hausdorff dimension can be related to the modulus by the following formula:

$$D = \frac{\log(\sum_{m=1}^M m)}{\log(M)}, \quad (7.15)$$

which is plotted in Fig. 7.12(i). Notably, the Sierpiński triangle is a modulus two ($M = 2$) Pascal triangle. Schematic depictions of antenna geometries for the second ($M = 2$), third ($M = 3$) and fourth ($M = 4$) moduli of the Pascal triangle are shown in Fig. 7.16(b)-(d). The primary resonance of an $M = 3$ antenna with length, $L = 475\text{nm}$, is at $\lambda = 2.01\mu\text{m}$,

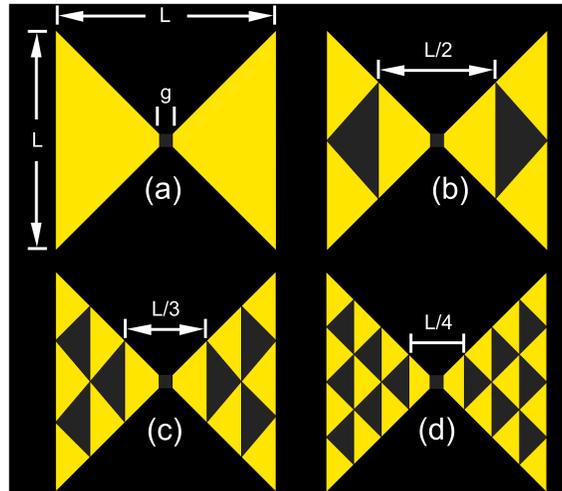


Figure 7.16: Schematic representation of (a) bowtie antenna, and bowtie antennas with Pascal triangle modulus (b) $M = 2$, (c) $M = 3$, and (d) $M = 4$ geometry.

as shown in Fig. 7.17. Increasing the modulus to $M = 4$, 5, and 6 produces resonances at $\lambda = 1.98\mu\text{m}$, $\lambda = 1.93\mu\text{m}$, and $\lambda = 1.91\mu\text{m}$, respectively, also shown in Fig. 7.17. A

simple bowtie antenna with $L = 475\text{nm}$ is found to resonate at $\lambda = 1.41\mu\text{m}$. Therefore, bowtie antennas with Pascal triangle geometry produce resonances that are substantially red-shifted relative to a standard bowtie antenna. In a similar manner as the SFPAs, the observed 50% change in the resonance frequency allows these antennas to operate at longer wavelengths without increasing the size of the antenna. This unique feature allows the antenna to confine light of a longer wavelength to the same nanoscale dimensions as a bowtie antenna.

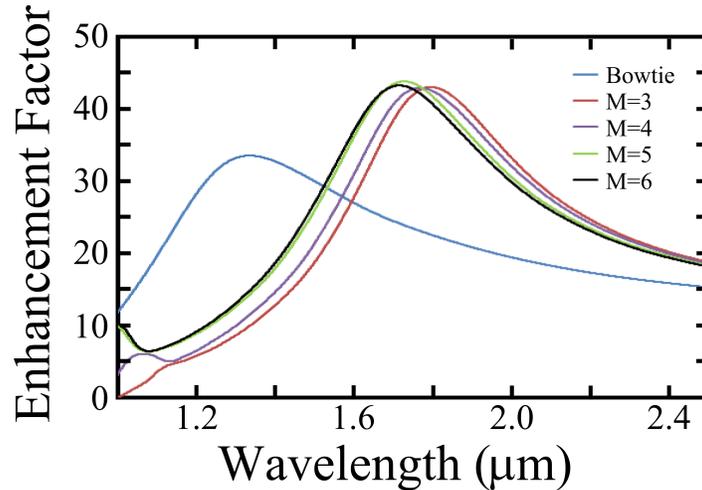


Figure 7.17: Broadband enhancement factor for a bowtie antenna with $L = 475\text{nm}$, along with $M = \{3, 4, 5, 6\}$ antennas with $L = 475\text{nm}$.

Logarithmic plots of the near-field intensity distribution of modulus $M = 3, 4,$ and 5 antennas at their resonances are shown in Figs. 7.18(a), (b), and (c), respectively. These distributions are normalized to the amplitude of the input excitation. Regardless of the value of M , the incident electromagnetic energy is confined to the antenna gap and not to other regions of the antenna.

A total of fifty-five antenna structures with moduli in the range, $2 \leq M \leq 10$, and lengths in the range, $175 \leq L \leq 725\text{nm}$, are used to investigate the influence of Hausdorff dimension on the antenna behaviour. As a typical example, the resonant wavelength is plotted versus the Hausdorff dimension for a constant antenna length, $L = 475\text{nm}$, and it is observed that as the Hausdorff dimension increases, the resonant wavelength decreases in a manner that is approximately parabolic, as shown in Fig. 7.19(a). Considering the converse, we plot the resonant wavelength versus antenna length for a constant modulus, $M = 4$, as shown in Fig. 7.19(b). In the same manner as a standard bowtie antenna (also shown in Fig. 7.19(b)), the resonant wavelength scales linearly with the antenna length. A paraboloid of the form $\lambda = c_1 L^2 + c_2 L + c_3 D^2 + c_4 D + c_5 LD + c_6$ is fit to data from the 55 antennas with $\{c_1, c_2, c_3, c_4, c_5, c_6\} = \{-1.43 \times 10^{-5}, 7.92, -3.53 \times 10^3, 1.22 \times 10^5, -3.03, -1.00 \times 10^5\}$ and $R^2 = 0.994$, where D is the Hausdorff dimension. This surface is shown in Fig. 7.19(c). Examining this plot, it can be observed that the antenna length is a sensitive parameter for adjusting the resonant wavelength, regardless

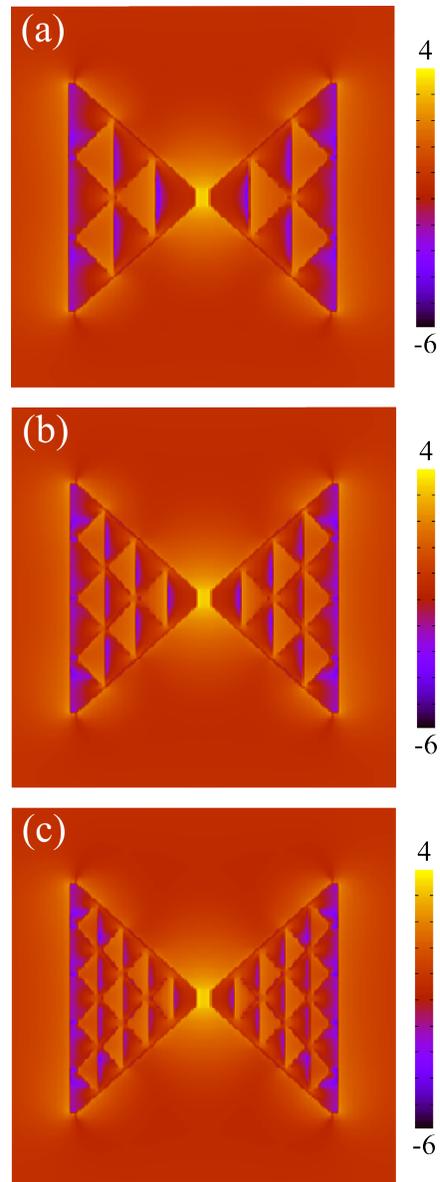


Figure 7.18: Logarithmic scale intensity (I) distributions for (a) $M = 3$, (b) $M = 4$, and (c) $M = 5$ antennas. Intensity distributions are normalized to the input excitation.

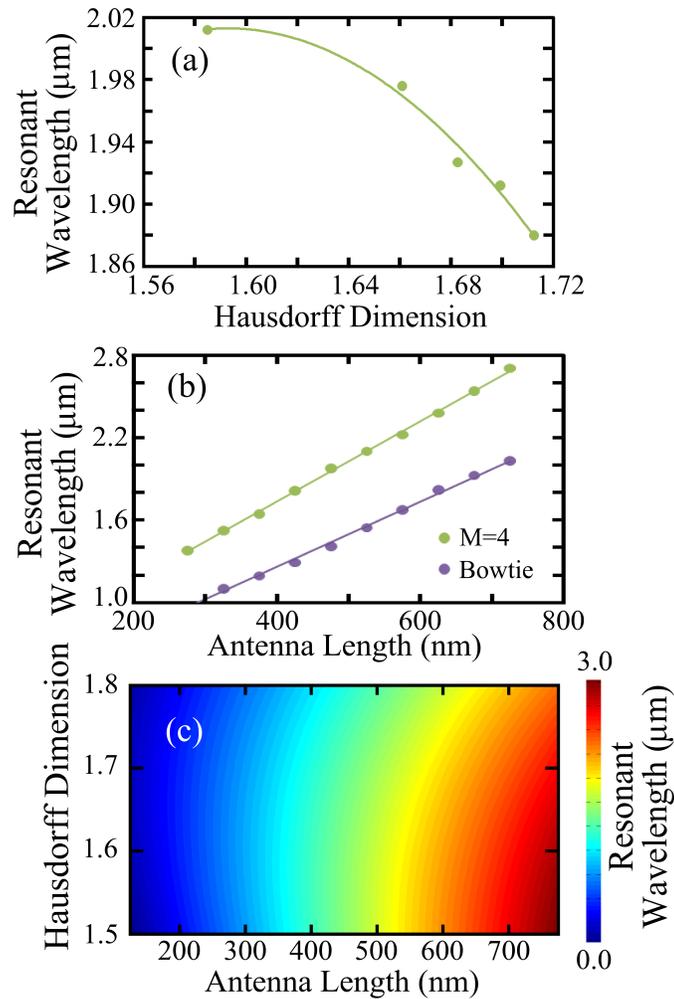


Figure 7.19: (a) Resonant wavelength versus modulus for $L = 475\text{nm}$ (constant). A parabolic trend line is included. (b) Resonant wavelength versus antenna length for $M = 4$ (constant). The trend for a bowtie antenna is shown for comparison. (c) Paraboloid surface-fit to resonant wavelengths of 55 antennas that were simulated.

of the Hausdorff dimension. Interestingly, the Hausdorff dimension can equally be used to achieve fine tuning of the antenna resonance.

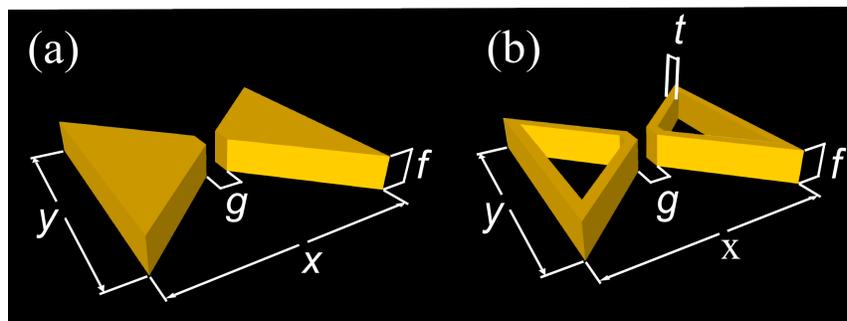


Figure 7.20: Schematic representations of the simulation geometries: (a) bowtie antenna and (b) contour bowtie antenna.

7.3.6 The Nanoplasmonic Contour Bowtie Antenna

Contour Thickness and Antenna Length

The geometry of a typical bowtie antenna can be characterized by four parameters, as shown in Fig. 7.20(a): the gap width, g ; the antenna x -dimension, x ; the antenna y -dimension, y , and the film thickness, f . In this study, we consider two metallic triangular contours, as shown in Fig. 7.20(b) and use the contour thickness, t , to design the antenna resonance, while the antenna dimensions are kept constant. Notably, the regular bowtie antenna is a special case of the contour bowtie antenna that occurs when the contour completely fills the centre region.

The broadband enhancement factor for antennas with a constant length, $x = 475\text{nm}$, and contour thicknesses, $t = \{40, 60, 80\}\text{nm}$ along with a bowtie antenna with $x = 475\text{nm}$ is shown in Fig. 7.21(a). From these plots, it is evident that regardless of the thickness, t , λ_{res} of the contour bowtie antenna is always at a longer wavelength than a bowtie antenna of the same length. Furthermore, as t decreases, λ_{res} shifts to longer wavelengths and the enhancement factor increases. As the contour thickness decreases from $t = 100\text{nm}$ to $t = 30\text{nm}$, λ_{res} shifts from $1.69\mu\text{m}$ to $2.42\mu\text{m}$ and the enhancement factor increases from 33.5 to 42.9. Notably, a bowtie antenna with the same dimensions has a resonance at $\lambda_{res} = 1.40\mu\text{m}$ and an enhancement factor of 37.1. Further simulations demonstrate that a standard bowtie antenna with a length, $x = 902\text{nm}$, would be necessary to resonate at $\lambda_{res} = 2.42\mu\text{m}$. This demonstrates that the contour bowtie antenna can be used to decrease x by a factor of 1.90, decrease the footprint by a factor of 3.60, and increase the enhancement by 28%.

The ring-like shape of the antenna warrants an investigation into the polarization purity of radiation that couples to the antenna. First, simulations are performed for a bowtie antenna with $L = 475\text{nm}$. In the first simulation, the antenna is excited with x -polarized radiation, and the wavelength of the primary resonance is obtained, $\lambda = 1409\text{nm}$. In the next simulation, the antenna is excited with y -polarized radiation, and the gap enhancement at $\lambda = 1409\text{nm}$ is found to be only 7.91% of the value obtained for x -polarized excitation. This procedure is repeated for a contour bowtie antenna with the same length, and it is found that the gap enhancement for y -polarized excitation is 14.3% of the enhancement observed for x -polarization. This demonstrates that the polarization selectivity of the contour bowtie antenna is diminished compared to the standard bowtie antenna.

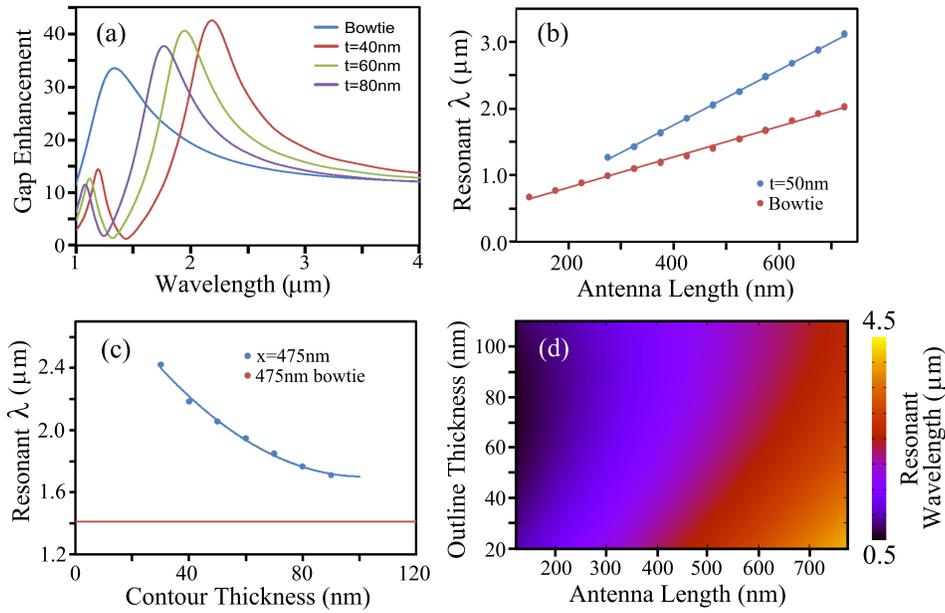


Figure 7.21: (a) Broadband enhancement factors for a bowtie antenna and contour bowtie antennas with $t = \{40, 60, 80\}$ nm and a length, $x = 475$ nm. (b) λ_{res} versus antenna length for a bowtie antenna and a $t = 50$ nm contour bowtie antenna, each with a length, $x = 475$ nm. (c) λ_{res} versus contour thickness, t , for an antenna length, $x = 475$ nm. λ_{res} for a bowtie antenna with $x = 475$ nm is shown for reference. (d) Paraboloid fit to 70 simulated data points relating λ_{res} to $\{x, t\}$.

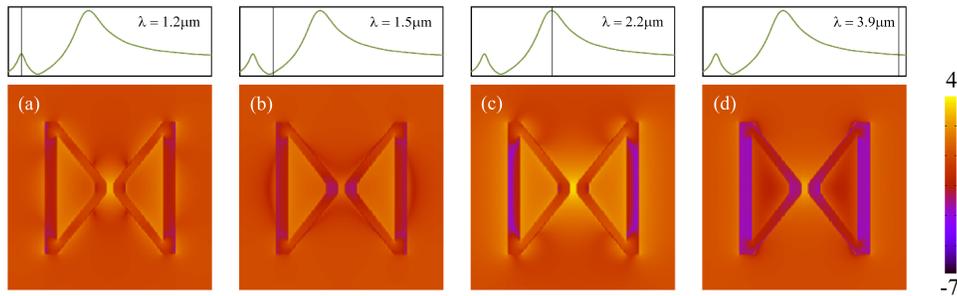


Figure 7.22: Logarithmic scale intensity (I) distributions at spectral points of interest: (a) Secondary resonance at $\lambda = 1.2\mu\text{m}$, (b) anti-resonance at $\lambda = 1.5\mu\text{m}$, (c) primary resonance at $\lambda = 2.2\mu\text{m}$, and (d) above resonance at $\lambda = 3.9\mu\text{m}$.

Regardless, these results demonstrate that the contour bowtie antenna behaves in the same way as a plasmonic nanoshell. In the case of a nanoshell, it has been shown that the plasmon resonance depends on the radius of the core material and on the thickness of the gold nanoshell. Increasing the radius of the core material (analogous to increasing the antenna dimensions) red-shifts the resonance, whereas increasing the nanoshell thickness (analogous to increasing t) blue-shifts the resonance [66]. Furthermore, it has been demonstrated in previous works that decreasing the antenna gap dimension yields greater electric field enhancement [1]. Decreasing the thickness, t , increases λ_{res} of the antenna and the gap dimension relative to λ_{res} decreases. As a result, decreasing t pro-

duces greater enhancement in the antenna gap. Such a significant increase in λ_{res} and enhancement factor would be unattainable by loading a nanoshell bowtie antenna with a dielectric [67].

The ability to design the resonant wavelength of the antenna is demonstrated by simulating a wide range of antenna lengths, x , and contour thicknesses, t . The wavelength of the primary resonance is calculated for 70 antennas with lengths spanning $175\text{nm} \leq x \leq 725\text{nm}$ in 50nm steps and contour thicknesses spanning $30\text{nm} \leq t \leq 100\text{nm}$ in 10nm steps. Fig. 7.21(b) shows a typical plot of λ_{res} versus x for $t = 50\text{nm}$. The length trend for a bowtie antenna is shown for comparison. Evidently, the relationship is linear and the operating wavelength of the contour bowtie antenna is higher than the bowtie antenna for each of the antenna lengths simulated. Interestingly, as the antenna length decreases, the empty space in the centre region of the antenna begins to fill and it starts to resemble a bowtie antenna. These two lines intersect at $x=141\text{nm}$, which is the point where the centre of the contour bowtie antenna is completely filled and its behaviour converges to that of the bowtie antenna. Similarly, the dependence between contour thickness, t , and λ_{res} is plotted in Fig. 7.21(c). This trend is parabolic, which shows that λ_{res} is directly proportional to the area removed from the centre of the antenna. Building on Figs. 7.21(b)-(c), a paraboloid of the form $\lambda = c_1x^2 + c_2x + c_3t^2 + c_4t + c_5xt + c_6$ is fit to 70 data points for various values of x and t with $\{c_1, c_2, c_3, c_4, c_5, c_6\} = \{6.75 \times 10^{-4}, 4.55, 0.129, -17.2, -2.01 \times 10^{-2}, 765\}$ and $R^2 = 0.999$. A surface plot of this fit is shown in Fig. 7.21(d).

To understand the behaviour of the antenna, it is useful to plot the intensity distribution at points of interest in its spectrum. The intensity distribution for a contour antenna with $x = 475\text{nm}$ and $t = 40\text{nm}$ is plotted at its secondary resonance ($\lambda = 1.2\mu\text{m}$) in Fig. 7.22(a), its anti-resonance ($\lambda = 1.5\mu\text{m}$) in Fig. 7.22(b), its primary resonance ($\lambda = 2.2\mu\text{m}$) in Fig. 7.22(c), and above λ_{res} ($\lambda = 3.9\mu\text{m}$) in Fig. 7.22(d). When excited on-resonance, the intensity distributions show that the antenna is strongly polarized and incident energy is confined to the gap. By controlling the incident electric field intensity, all sensitive processes can be made to originate from the antenna gap, making this design suitable for probing a small number of atoms or molecules localized to the gap region. When excited at its anti-resonance, fields are slightly enhanced in the hollow region of the antenna, while fields in the gap experience only minimal enhancement. Although enhancement is observed in the gap when the antenna is excited above λ_{res} , there is minimal interaction with the remainder of the antenna. Examination of Fig. 7.22(c) reveals that the interaction of the radiation with the vertical edge on-resonance is notably different than with the other edges. Additionally, there is moderate electric field enhancement in the hollow region of the antenna. Therefore, it would be interesting to isolate the vertical edge and examine its influence on the resonance conditions of the antenna. This is done by two separate studies: (1) Varying only the thickness of the vertical edge while maintaining the thickness of the other edges constant; and (2) Displacing the vertical edge

towards the centre of the antenna.

Vertical Edge Thickness

To investigate the influence of the vertical edge thickness, t_v , the thickness of the two diagonal edges is maintained constant at $t = 30\text{nm}$, the antenna length is kept constant at $x = 475\text{nm}$, and the thickness of the vertical edge is varied between $0\text{nm} \leq t_v \leq 150\text{nm}$. A schematic depiction of this geometry is shown in Fig. 7.23(a). The effect of the edge thickness on λ_{res} is presented in Fig. 7.23(b). In a similar manner as increasing the contour thickness (shown in Fig. 7.21(c)) in a uniform manner decreases λ_{res} , so does increasing only the vertical edge thickness. However, in this case λ_{res} decreases linearly with the vertical edge thickness. For a uniform edge thickness, each infinitesimal increase in t produces a greater change to the geometry than increasing t_v . The convergence of the geometry to that of the bowtie antenna is approximately linear when only t_v is increased, yielding the linear trend in Fig. 7.23(b).

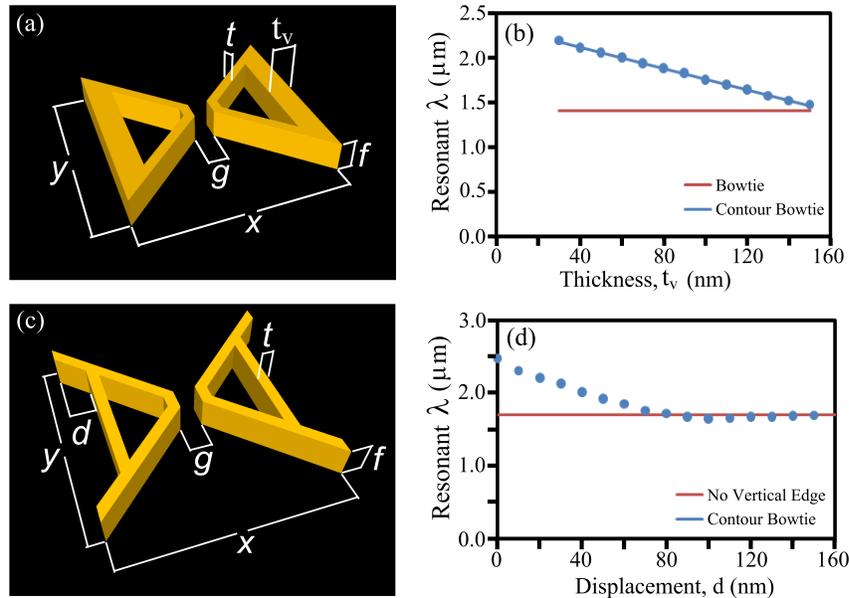


Figure 7.23: (a) Schematic representation of a contour bowtie antenna with $t_v \neq t$. (b) λ_{res} versus t_v for $\{x, t\} = \{475, 30\}\text{nm}$. λ_{res} for a bowtie antenna is shown for reference. (c) Schematic representation of a contour bowtie antenna with a shifted vertical edge. (d) λ_{res} versus edge displacement, d . The case when the vertical edge is removed altogether is shown for reference.

Vertical Edge Displacement

A typical example of this geometry is shown in Fig. 7.23(c). The antenna length is kept constant at $x = 475\text{nm}$, while the vertical edge is shifted towards the centre of the antenna by a distance, d . The relationship between d and λ_{res} is shown in Fig. 7.23(d). Interestingly, as d increases, the useful area of the antenna effectively shrinks and λ_{res}

decreases up to $d = 80\text{nm}$. To illustrate this, we consider $d = 50\text{nm}$. The vertical edges of both the left and right contours are shifted inwards, so their combined shift is $2d = 100\text{nm}$. Interestingly, λ_{res} of a normal contour bowtie antenna with $t = 30\text{nm}$ and $x = 375\text{nm}$ is almost identical at $\lambda_{res} = 1940\text{nm}$, indicating that the corner segments extending beyond the vertical edge do not interact with incident radiation at resonance. The situation changes as the vertical edge is shifted beyond 80nm and λ_{res} saturates and approaches $\lambda_{res} = 1692\text{nm}$ when $d = 150\text{nm}$. When the vertical edge is removed altogether, $\lambda_{res} = 1696\text{nm}$ and the gap enhancement decreases only slightly to 41.2, demonstrating that the antenna is still somewhat functional without the vertical edge and that as the vertical edge is shifted beyond $d = 80\text{nm}$, it interacts only minimally with the incident radiation. Based on these observations, the placement of the vertical edge as shown in Fig. 7.20(b) is the optimal case for this antenna design.

7.3.7 Summary

The above results provide theoretical insight into the operation of gold plasmonic bowtie antennas that include nanoscale features in each of the constituent triangles. We have simulated structures that offer a compact antenna footprint, a large confinement factor, resonant wavelengths in the range, $700\text{nm} \leq \lambda \leq 4.5\mu\text{m}$, and electric field enhancement factors of up to 56 in the centre of the antenna gap. In the case of bowtie antennas with Sierpiński triangle geometry, an increase in the fractal iteration produced a red-shift in the antenna resonance, enabling more compact antenna designs. Through simulation of non-fractal geometries, it was demonstrated that this behaviour is related to the ability of the fractal geometry to become polarized. In contrast, increasing the Hausdorff dimension of bowtie antennas with Pascal triangle geometry produced a slight blue-shift in the antenna resonance.

In the case of a contour bowtie antenna, the behaviour is analogous to plasmonic nanoshells, and increasing the antenna length and decreasing the contour thickness both produce a red-shift the antenna resonance conditions. These results demonstrate that these antennas offer more compact dimensions when compared to a bowtie antenna and that the operating wavelength of the antenna can be tuned over a wide range by changing the contour thickness, t , while maintaining a constant antenna footprint. Both the thickness of the vertical edge of the antenna and its position are found to be very important for optimizing the antenna performance.

It is anticipated that these results will act as a starting point for both theoretical and experimental investigations into plasmonic antennas with fractal or other complex geometries. In addition, these designs may be used to develop compact plasmonic antennas for sensitive investigation of light-matter interaction processes in the mid-infrared, such as molecular vibration modes, or extreme nonlinearities in dielectrics or wide bandgap semiconductors.

References

- [1] D. P. Fromm, A. Sundaramurthy, P. J. Schuck, G. Kino, and W. E. Moerner. Gap-dependent optical coupling of single “bowtie” nanoantennas resonant in the visible. *Nano Letters*, 4:957–961, 2004.
- [2] J. S. Huang, T. Feichtner, P. Biagioni, and B. Hecht. Impedance matching and emission properties of nanoantennas in an optical nanocircuit. *Nano Letters*, 9:1897–1902, 2009.
- [3] M. I. Brongersma and P. G. Kik. *Surface Plasmon Nanophotonics*. Springer, Dordrecht, 2007.
- [4] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J. Y. Laluet, and T. W. Ebbesen. Channel plasmon subwavelength waveguide components including interferometers and ring resonators. *Nature*, 440:508–511, 2006.
- [5] E. Ozbay. Plasmonics: Merging photonics and electronics at nanoscale dimensions. *Science*, 311:189–193, 2006.
- [6] T. Nikolajsen, K. Leosson, and S. I. Bozhevolnyi. Surface plasmon polariton based modulators and switches operating at telecom wavelengths. *Applied Physics Letters*, 85:5833–5835, 2004.
- [7] J. Gosciniak, S. I. Bozhevolnyi, T. B. Andersen, V. S. Volkov, J. Kjelstrup-Hansen, L. Markey, and A. Dereux. Thermo-optic control of dielectric-loaded plasmonic waveguide components. *Opt. Express*, 18:1207–1216, 2010.
- [8] D. J. Bergman and M. I. Stockman. Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems. *Phys. Rev. Lett.*, 90:027402, 2003.
- [9] M. Ambati, S. H. Nam, E. Ulin-Avila, D. A. Genov, G. Bartal, and X. Zhang. Observation of stimulated emission of surface plasmon polaritons. *Nano Letters*, 8:3998–4001, 2008.
- [10] M. T. Hill, M. Marell, E. S. P. Leong, B. Smalbrugge, Y. Zhu, M. Sun, P. J. van Veldhoven, E. J. Geluk, F. Karouta, Y. S. Oei, R. Nötzel, C. Z. Ning, and M. K.

- Smit. Lasing in metal-insulator-metal sub-wavelength plasmonic waveguides. *Opt. Express*, 17:11107–11112, 2009.
- [11] I. De Leon and P. Berini. Amplification of long-range surface plasmons by a dipolar gain medium. *Nature Photon.*, 4:382–387, 2010.
- [12] J. A. Dionne, K. Diest, L. A. Sweatlock, and H. A. Atwater. Plasmostor: A metal-oxide-si field effect plasmonic modulator. *NanoLetters*, 9:897–902, 2009.
- [13] K. F. MacDonald, Z. L. Samson, M. I. Stockman, and N. I. Zheludev. Ultrafast active plasmonics: transmission and control of femtosecond plasmon signals. *Nature Photon.*, 3:55–58, 2008.
- [14] K. J. Chau, S. E. Irvine, and A. Y. Elezzabi. A gigahertz surface magneto-plasmon optical modulator. *IEEE J. Quantum Electron.*, 40:571–579, 2004.
- [15] Jr. P. A. Schumann and R. P. Phillips. Comparison of classical approximateions to free-carrier absorption in semiconductors. *Solid-State Electron.*, 10:943–948, 1967.
- [16] F. E. Doany, D. Grischkowsky, and C. C. Chi. Carrier lifetime versus ion-implanted dose in silicon on sapphire. *Appl. Phys. Lett.*, 50:460, 1987.
- [17] J. A. Dionne, L. A. Sweatlock, H. A. Atwater, and A. Polman. Plasmon slot waveguides: Towards chip-scale propagation with subwavelength-scale localization. *Phys. Rev. B*, 73:035407, 2006.
- [18] S. Gupta, M. Y. Frankel, J. A. Valdmanis, J. F. Whitaker, G. A. Mourou, F. W. Smith, and A. R. Calawa. Subpicosecond carrier lifetime in gaas grown by molecular beam epitaxy at low temperatures. *Appl. Phys. Lett.*, 59:3276, 1991.
- [19] S. Gupta, J. F. Whitaker, and G. A. Mourou. Ultrafast carrier dynamics in iii-v semiconductors grown by molecular beam epitaxy at very low substrate temperatures. *IEEE J. Quantum Electron.*, 28:2464–2472, 1992.
- [20] A. Y. Elezzabi, J. Meyer, M. K. Y. Hughes, and S. R. Johnson. Generation of 1-ps infrared pulses at $10.6\mu\text{m}$ by use of low-temperature-grown gaas as an optical semiconductor switch. *Opt. Lett.*, 19:898–900, 1994.
- [21] S. Gupta, P. K. Bhattacharya, J. Pamulapati, and G. Mourou. Subpicosecond photoreponse of carriers in low-temperature molecular beam epitaxial $\text{in}_{0.52}\text{al}_{0.48}\text{as}/\text{inp}$. *Appl. Phys. Lett.*, 57:1543–1545, 1990.
- [22] Y. Kostoulas, L. J. Waxer, I. A. Walmsley, G. W. Wicks, and P. M. Fauchet. Femtosecond carrier dynamics in low-temperature-grown indium phosphide. *Appl. Phys. Lett.*, 66:1821–1823, 1995.

- [23] L.F. Lester, K.C. Hwang, P. Ho, J. Mazurowski, J.M. Ballingall, J. Sutliff, S. Gupta, J. Whitaker, and S.L. Williamson. Ultrafast long-wavelength photodetectors fabricated on low-temperature ingaas on gaas. *Photonics Technology Letters, IEEE*, 5:511–514, 1993.
- [24] K. F. Lamprecht, S. Juen, L. Palmetshofer, and R. A. Höpfel. Ultrashort carrier lifetimes in H⁺ bombarded inp. *Appl. Phys. Lett.*, 59:926–928, 1991.
- [25] M. B. Johnson, T. C. McGill, and N. G. Paulter. Carrier lifetimes in ion-damaged gaas. *Appl. Phys. Lett.*, 54:2424–2426, 1989.
- [26] A. Y. Elezzabi, J. Meyer, and M. K. Y. Hughes. 600 fs 10.6 μ m infrared pulse generation with radiation-damaged gaas reflection switch. *Appl. Phys. Lett.*, 66:402–404, 1995.
- [27] M. I. Stockman. Nanoplasmonics: past, present, and glimpse into future. *Opt. Express*, 19:22029–22106, 2011.
- [28] Y. Feng, M. Feser, A. Lyon, S. Rishton, X. Zeng, S. Chen, S. Sassolini, and W. Yun. Nanofabrication of high aspect ratio 24 nm x-ray zone plates for x-ray imaging applications. *J. Vac. Sci. Technol. B*, 25:2004, 2007.
- [29] M. A. Mohammad, T. Fito, J. Chen, M. Aktary, M. Stepanova, and S. K. Dew. Interdependence of optimum exposure dose regimes and the kinetics of resist dissolution for electron beam nanolithography of polymethylmethacrylate. *J. Vac. Sci. Technol. B*, 28:L1, 2010.
- [30] M.A. Mohammad, S.K. Dew, S. Evoy, and M. Stepanova. Fabrication of sub-10 nm silicon carbon nitride resonators using a hydrogen silsesquioxane mask patterned by electron beam lithography. *Microelectronic Engineering*, 88:2338–2341, 2011.
- [31] S. Sederberg, V. Van, and A. Y. Elezzabi. Monolithic integration of plasmonic waveguides into a complimentary metal-oxide-semiconductor- and photonic-compatible platform. *Appl. Phys. Lett.*, 96:121101, 2010.
- [32] S. Zhu, T. Y. Liow, G. Q. Lo, and D. L. Kwong. Silicon-based horizontal nanoplasmonic slot waveguides for on-chip integration. *Opt. Express*, 19:8888–8902, 2011.
- [33] S. Zhu, T. Y. Liow, G. Q. Lo, and D. L. Kwong. Full complementary metal-oxide-semiconductor compatible nanoplasmonic slot waveguides for silicon electronic photonic integrated circuits. *Appl. Phys. Lett.*, 98:021107, 2011.
- [34] Y. Song, J. Wang, Q. Li, M. Yan, and M. Qiu. Broadband coupler between silicon waveguide and hybrid plasmonic waveguide. *Opt. Express*, 18:13173–13179, 2010.

- [35] Z. Han, A. Y. Elezzabi, and V. Van. Experimental realization of subwavelength plasmonic slot waveguides on a silicon platform. *Opt. Lett.*, 35:502–504, 2010.
- [36] J. Meitzner, F. G. Moore, B. M. Tillotson, S. D. Kevan, and G. L. Richmond. Time-resolved measurement of free carrier absorption, diffusivity, and internal quantum efficiency in silicon. *Appl. Phys. Lett.*, 103:092101, 2013.
- [37] P. Mühlischlegel, H.-J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl. Resonant optical antennas. *Science*, 308:1607–1609, 2005.
- [38] M. Schnell, A. Garcia-Extarri, A. J. Huber, K. Crozier, J. Aizpurua, and R. Hillenbrand. Controlling the near-field oscillations of loaded plasmonic nanoantennas. *Nature Photon.*, 3:287–291, 2009.
- [39] H. Fischer and O. J. F. Martin. Engineering the optical response of plasmonic nanoantennas. *Opt. Express*, 16:9144–9154, 2008.
- [40] K. Crozier, A. Sundaramurthy, G. S. Kino, and C. F. Quate. Optical antennas: Resonators for local field enhancement. *J. Appl. Phys.*, 94:4632, 2003.
- [41] F. Neubrech, A. Pucci, T. W. Cornelius, S. Karim, A. Garcia-Etxarri, and J. Aizpurua. Resonant plasmonic and vibrational coupling in a tailored nanoantenna for infrared detection. *Phys. Rev. Lett.*, 101:157403, 2008.
- [42] L. Wang and X. Xu. High transmission nanoscale bowtie-shaped aperture probe for near-field optical imaging. *Appl. Phys. Lett.*, 90:261105, 2007.
- [43] S. Aksu, A. A. Yanik, R. Adato, A. Artar, M. Huang, and H. Altug. High-throughput nanofabrication of infrared plasmonic nanoantenna arrays for vibrational nanospectroscopy. *Nano Letters*, 10:2511–2518, 2010.
- [44] R. Adato, A. A. Yanik, J. J. Amsden, D. L. Kaplan, F. G. Omenetto, M. K. Hong, S. Erramilli, and H. Altug. Ultra-sensitive vibrational spectroscopy of protein monolayers with plasmonic nanoantenna arrays. *Proceedings of the National Academy of Sciences*, 106:19227–19232, 2009.
- [45] S. Kim, J. Jin, Y. J. Kim, I. Y. Park, Y. Kim, and S. W. Kim. High-harmonic generation by resonant plasmon field enhancement. *Nature*, 453:757–760, 2008.
- [46] N. Yu, E. Cubukcu, L. Diehl, D. Bour, S. Corzine, J. Zhu, Gloria Höfler, K. B. Crozier, and F. Capasso. Bowtie plasmonic quantum cascade laser antenna. *Opt. Express*, 15:13272–13281, 2007.
- [47] K. D. Ko, A. Kumar, K. H. Fung, R. Ambekar, G. L. Liu, N. X. Fang, and K. C. Toussaint. Nonlinear optical response from arrays of au bowtie nanoantennas. *Nano Letters*, 11:61–65, 2011.

- [48] S. A. Maier. *Plasmonics: Fundamentals and Applications*. Springer, New York, 2007.
- [49] J. D. Kraus. *Antennas*. McGraw-Hill, New York, 1988.
- [50] H. Jasik. *Antenna Engineering Handbook*. McGraw-Hill, New York, 1961.
- [51] L. Novotny. Effective wavelength scaling for optical antennas. *Phys. Rev. Lett.*, 98:266802, 2007.
- [52] A. Alù and N. Engheta. Input impedance, nanocircuit loading, and radiation tuning of optical nanoantennas. *Phys. Rev. Lett.*, 101:043901, 2008.
- [53] G.W. Hanson. On the applicability of the surface impedance integral equation for optical and near infrared copper dipole antennas. *Antennas and Propagation, IEEE Transactions on*, 54:3677–3685, 2006.
- [54] P. B. Johnson and R. W. Christy. Optical constants of the noble metals. *Phys. Rev. B*, 6:4370–4379, 1972.
- [55] A. Vial, A. S. Grimault, D. Macias, D. Barchiesi, and M. L. de la Chapelle. Improved analytical fit of gold dispersion: Application to the modeling of extinction spectra with a finite-difference time-domain method. *Phys. Rev. B*, 71:085416, 2005.
- [56] C. Reale. Optical constants of vacuum deposited thin metal films in the near infrared. *Infrared Phys.*, 10:175–181, 1970.
- [57] B. B. Mandelbrot. *The Fractal Geometry of Nature*. W. H, Freeman, 1982.
- [58] K. J. Falconer. *Fractal Geometry: Mathematical Foundation and Applications*. Wiley, Chichester, 2003.
- [59] D.H. Werner and S. Ganguly. An overview of fractal antenna engineering research. *Antennas and Propagation Magazine, IEEE*, 45:38–57, 2003.
- [60] C. Gaubert, L. Chusseau, A. Giani, D. Gasquet, F. Garet, F. Aquistapace, L. Duvillelet, J. L. Coutaz, and W. Knap. Thz fractal antennas for electrical and optical semiconductor emitters and receptors. *Phys. Status Solidi C*, 1:1439–1444, 2004.
- [61] Y. J. Bao, B. Zhang, Z. Wu, J. W. Si, M. Wang, R. W. Peng, X. Lu, J. Shao, Z. F. Li, X. P. Hao, and N. B. Ming. Surface-plasmon-enhanced transmission through metallic film perforated with fractal-featured aperture array. *Appl. Phys. Lett.*, 90:251914, 2007.
- [62] F. Miyamaru, Y. Saito, M. W. Takeda, B. Hou, L. Liu, W. Wen, and P. Sheng. Terahertz electric response of fractal metamaterial structures. *Phys. Rev. B*, 77:045124, 2008.

- [63] A. Agrawal, T. Matsui, W. Zhu, A. Nahata, and Z. V. Vardeny. Terahertz spectroscopy of plasmonic fractals. *Phys. Rev. Lett.*, 102:113901, 2009.
- [64] J. Matteo and L. Hesselink. Fractal extensions of near-field aperture shapes for enhanced transmission and resolution. *Opt. Express*, 13:636–647, 2005.
- [65] B. Hou, X. Q. Liao, and J. K. S. Poon. Resonant infrared transmission and effective medium response of subwavelength h-fractal apertures. *Opt. Express*, 18:3946–3951, 2010.
- [66] R. D. Averitt, D. Sarkar, and N. J. Halas. Plasmon resonance shifts of au-coated au_2s nanoshells: Insight into multicomponent nanoparticle growth. *Phys. Rev. Lett.*, 78:4217–4220, 1997.
- [67] Y.-F. Chau, H.-H. Yeh, and D. P. Tsai. A new type of optical antenna: Plasmonics nanoshell bowtie antenna with dielectric hole. *Journal of Electromagnetic Waves and Applications*, 24:1621–1632, 2010.

Chapter 8

Conclusions

The results presented in this thesis demonstrate several important steps towards realizing all-optical nanoplasmonic circuitry on a complementary metal-oxide-semiconductor (CMOS)-compatible fabrication platform. Extensive nanofabrication process development was undertaken to achieve high-yield processes for fabricating silicon (Si) photonic waveguides, silicon-based nanoplasmonic waveguides interfaced to silicon photonic waveguides, and silicon-based nanoplasmonic waveguides integrated onto a micron-scale characterization beam. With the exception of the gold (Au) components, all of the materials and processing techniques were CMOS-compatible. Notably, the Au features could easily be replaced with a CMOS-compatible plasmonic material, such as TiN, Cu, or Al, permitting monolithic integration of electronic and nanoplasmonic components.

Preliminary investigations into nonlinearities in silicon-on-insulator (SOI) photonic waveguides were presented in Chapter 3 to gain a detailed understanding of the interactions and to develop a performance baseline for experiments with Si-based nanoplasmonic waveguides. Two-photon absorption (TPA), free-carrier absorption (FCA), self-phase modulation, free-carrier plasma dispersion, and the thermo-optic effect were each investigated, along with the interplay between these effects. Of particular interest were the nonlinear loss mechanisms, TPA and FCA, for use in all-optical modulation. The free-carrier recovery time of Si was measured to be approximately $\tau_r = 265\text{ps}$, and was found to increase with increasing pump power. It was observed that Si photonic waveguides shorter than $L = 25\mu\text{m}$ yielded a weak modulation depth.

The generation of strong third-harmonic generation (THG) from SOI photonic waveguides was presented in Chapter 4. In longer Si photonic waveguides, it was found that as the waveguide core dimensions were decreased, the peak intensity was greater, and produced a higher free-carrier density near the pulse peak. This, in turn, produced a rapidly decreasing refractive index, which blue-shifted and broadened the third-harmonic (TH) spectrum. By decreasing the length of the Si photonic waveguide, the blue-shifting was eliminated, and TH conversion efficiencies up to $\eta_{THG} = 2.8 \times 10^{-5}$ were demonstrated.

Subsequent experiments were performed in SOI nanoplasmonic waveguides. Initial experiments demonstrated interfacing between SOI photonic waveguides and SOI nanoplas-

monic waveguides, with a coupling efficiency of 37%. The propagation length of the nanoplasmonic mode was measured to be $L_{prop} = 2.0\mu\text{m}$. The short propagation length arises from the high refractive index of Si, which produces extremely strong electromagnetic confinement at the Au-Si interface and high Ohmic losses. Although this is undesirable for passive routing, such confinement is highly desirable for achieving strong nonlinear signatures in a very short length scale.

Studies of ultrafast nonlinear interactions were performed on silicon-based nanoplasmonic waveguides integrated onto a micron-scale characterization beam, eliminating any inaccuracies introduced by the on-chip coupler. This enabled several unique ultrafast nonlinear effects to be investigated in detail. When excited with femtosecond $\lambda = 1550\text{nm}$ pulses, bright visible light emission from the nanoplasmonic waveguides was easily observed with the naked eye under normal room lighting. When imaged with a microscope, strong white light emission was observed from the input facet of the nanoplasmonic waveguide and strong green THG was observed from the output facet.

By analyzing the spectrum of the visible light emitted from the output facet of the nanoplasmonic waveguide, it was determined that the green light spanned a spectrum of $500\text{nm} \leq \lambda \leq 530\text{nm}$, while the white light spanned a spectrum of $375\text{nm} \leq \lambda \leq 725\text{nm}$. Power scaling trends revealed that the green light power scaled with the cube of the input intensity, confirming the THG process, while the white light power grew exponentially with the input power, signifying a different and unique light emission process.

Significant insight into the interactions responsible for the white light emission was gained through numerical simulations. Electrons excited via TPA were accelerated to energies up to several eV by the ponderomotive potential that exists in the highly confined plasmonic field. Subsequent collisions with valence electrons drive impact ionization, which multiplied in an avalanche. Electrons that did not reach the threshold for impact ionization emit photons during collisions, explaining the exponential growth of the white light power.

Subsequently, electron sweeping away from the Au-Si interface due to ponderomotive electron acceleration was visualized using pump-probe time domain spectroscopy and the sweeping time was measured to be $\tau_2 = 1.98 \pm 0.40\text{ps}$, confirming that the electrons travel away from the interface at the saturation velocity. A second timescale was also observed, $\tau_2 = 17.9 \pm 6.8\text{ps}$, which was attributed to free-carrier recombination at diffused Au recombination centres and at surface traps.

The final chapter presented two Si-based ultrafast all-optical nanoplasmonic modulator structures: a nanoplasmonic slab waveguide and a nanoplasmonic ring resonator, which achieved signal modulation on a timescale of 3-5ps. As well, three new classes of nanoplasmonic antennas were presented that operated at red-shifted resonant wavelengths with respect to bowtie antennas, enabling higher integration densities of the antennas.

Future Directions

A key factor to consider when studying any nanoscale system is how to interface it to the macroscopic world. The most enabling nanofabrication platform for studying nonlinear interactions in Si-based nanoplasmonic devices was the integration of nanoplasmonic waveguides onto a macroscopic characterization beam. The flexibility of the fabrication steps allow for any nanoplasmonic or nanophotonic structure to be fabricated on this platform, and studied directly with ultrafast lasers. Therefore, this serves as an excellent means to perform first-principles studies on nanoplasmonic devices before they are integrated onto a chip with other structures. As with any technology, there is still much to be explored at smaller dimensions, and in particular, to match the scale of the nanoplasmonic waveguides to the scale of modern transistors. The following are several other possibilities that remain to be explored:

- Although breakthrough THG conversion efficiencies were demonstrated, it is anticipated that substantial increases are still possible. The geometric and operational simplicity of the devices studied leaves many avenues open for subsequent designs to increase the conversion efficiency, while maintaining a high operational bandwidth, permitting the generation of μW powers in a chip-scale Si structure. Further developments into integrating these visible light sources with low-loss photonic waveguides (such as SiN) can be pursued, enabling the generated light to be routed and used on a chip.
- The nanoplasmonic modulators operated on the principle of nonlinear loss due to TPA and FCA. Notably, the presence (and avalanche multiplication) of free-carriers can introduce a substantial change to the refractive index that can be used to develop phase-sensitive devices. Therefore, new classes of resonant nanoplasmonic modulators, such as ring resonators, Mach-Zehnder interferometers, and Fabry-Pérot resonators can be developed, with potentially higher operating efficiencies.
- All-optical logic gates can be developed, with operational principles in direct analogy to electronic logic gates. The potential for cascaded logic and for all-plasmonic integrated circuitry can then be investigated. Moreover, the device recovery time can be reduced further by using ion-implanted Si with a recovery time as low as $\tau \sim 600\text{fs}$, enabling terahertz modulation frequencies.
- The high sensitivity of the electron avalanche process and the resulting electrical currents also indicate possible applications as nanoscale avalanche photodetectors on a chip, where the avalanche can be driven either electrically, optically, or through a combination of both. This will enable a sensitive means to convert optical signals to electrical signals.

- The timescale of the carrier sweeping (i.e. $\tau_1 \approx 2\text{ps}$) demonstrates that terahertz radiation is generated as a bi-product of the ponderomotive electron acceleration. Therefore, it would be interesting to measure the strength of the terahertz fields and to examine schemes that could be used to both generate and route terahertz signals in a larger silicon-based waveguide.
- The broadband electric field enhancement in the near- to mid-infrared provided by the nanoplasmonic antennas make them an ideal structure for investigating nonresonant nonlinearities in insulators and wide-gap semiconductors and the timescale of the resulting electron dynamics, particularly in light of the recent breakthroughs in this area [1, 2].

This thesis primarily exploited nonlinear and strong-field interactions to achieve modulation in nanoplasmonic devices. It should be noted that other effects can be used for modulation, and could potentially enable additional functionality in nanoplasmonic modulators. One possibility that has been investigated is magneto-optical control of plasmon propagation [3–5]. In this configuration, the presence of a slowly-varying magnetic field alters the magneto-optic medium and modulates the plasmonic signal. In particular, the magneto-optic effect can be used in Faraday rotation devices, such as optical isolators. Since plasmonic waveguides include high-conductivity metal features, it is convenient to use the magnetic field associated with electrical currents to adjust the properties of an adjacent magneto-optical materials and modulate the plasmonic signal. The realization of nanoplasmonic waveguides loaded with magneto-optical materials would enable polarization rotation in integrated plasmonic structures, providing an additional degree of freedom for modulation. The switching speed of magneto-optical modulators is limited by the time it takes for the magnetization to switch orientation, which is on the order of nanoseconds. Therefore, magneto-plasmonic modulators are most suitable for circuitry that does not require ultrafast modulation.

An alternative approach for additional functionality would be through phase-change materials. Interestingly, the operation of the first reported active plasmonic waveguide was based on a temperature-induced phase change in gallium [6]. Recently, there have been several other reports of surface plasmon devices working on the principle of phase transitions, incorporating exotic materials such as $\text{Ge}_2\text{Sb}_2\text{Te}_5$ [7] and VO_2 [8]. Although phase transitions enable an additional means to modulate plasmon propagation, these transitions typically occur on the microsecond to second timescale, and would not be suitable for ultrafast modulation. However, phase transitions could provide a useful means to activate a plasmonic circuit. For example, increasing the temperature of a VO_2 -loaded nanoplasmonic waveguide changes the VO_2 from a semiconductor to a metal. Such a transition could be used to control whether the circuit would operate in the passive regime or in the active regime. Furthermore, plasmonic structures loaded with phase-change materials could be used as environment sensors in plasmonic circuitry.

Outlook

Although Si photonic waveguides are currently being implemented as optical interconnects in computer chips, their maximum integration density will largely determine the extent of their application. As shown in Section 2.4.1, even a spacing of 100nm between photonic waveguides leads to detrimental cross-talk, and therefore, their most promising applications lie optical signal transmission over long distances where integration density is not important. However, there will still be many electronic devices on the chip where the integration density is too high for Si photonic interconnects to be practical. In these regions, the devices are separated by a very small distance, making high-confinement plasmonic waveguides an excellent means to achieve optical interfacing between devices. Despite the high propagation losses of the plasmonic modes, the device separation will be small enough that strong signals can still be delivered. Notably, the plasmonic circuitry could consist of both hybrid plasmonic waveguides (for lower propagation losses) and the structures investigated in this work (for efficient switching and optical-to-electrical signal conversion).

The well-established growth and processing techniques of Si, along with the availability of high-quality telecommunications band lasers and SOI substrates have made Si photonics a readily accessible field of research. These practical factors have enabled its rapid development over the last several decades and have led to promising industrial applications. For the same reasons, Si-based nanoplasmonics has great potential that has yet to be realized by the plasmonics or photonics community. The most novel results from the thesis provide an understanding of the fundamental nonlinear and strong field interactions occurring in Si-based nanoplasmonic waveguides, and lay the groundwork for developing a number of devices. These findings offer a means to harness the potential of the emerging field of nonlinear nanoplasmonics.

References

- [1] A. Schiffrin, T. Paasch-Colberg, N. Karpowicz, V. Apalkov, D. Gerster, S. Muhlbrandt, M. Korbman, J. Riechert, M. Schultze, J. V. Holzner, S. andd Barth, R. Kienberger, R. Ernstorfer, V. S. Yakovlev, M. I. Stockman, and F. Krausz. Optical-field-induced current in dielectrics. *Nature*, 493:70–74, 2013.
- [2] M. Schulze, E. M. Bothschafter, A. Sommer, W. Schweinberger, M. Fiess, M. Hofstetter, R. Kienberger, V. Apalkov, V. S. Yakovlev, M. I. Stockman, and F. Krausz. Controlling dielectrics with the electric field of light. *Nature*, 493:75–78, 2013.
- [3] K. J. Chau, S. E. Irvine, and A. Y. Elezzabi. A gigahertz surface magneto-plasmon optical modulator. *IEEE J. Quantum Electron.*, 40:571–579, 2004.
- [4] V. V. Temnov, G. Armelles, U. Woggon, D. Guzatov, A. Cebollada, A. Garcia-Martin, J. M. Garcia-Martin, T. Thomay, A. Leitenstorfer, and R. Bratschitch. Active magneto-plasmonics in hybrid metal-ferromagnet structures. *Nature Photon.*, 4:107–111, 2010.
- [5] B. Sepulveda, L. M. Lechuga, and G. Armelles. Magneto-optic effects in surface-plasmon-polaritons slab waveguides. *J. Lightw. Technol.*, 24:945–955, 2006.
- [6] A. V. Krasavin and N. I. Zheludev. Active plasmonics: controlling signals in au/ga waveguide using nanoscale structural transformations. *Appl. Phys. Lett.*, 84:1416–1418, 2004.
- [7] Y. G. Chen, T. S. Kao, B. Ng, X. Li, X. G. Luo, B. Luk'yanchuk, S. A. Maier, and M. H. Hong. Hybrid phase-change plasmonic crystals for active tuning of lattice resonances. *Opt. Express*, 21:13691–13698, 2013.
- [8] D. W. Ferrara, J. Nag, E. R. MacQuarrie, A. B. Kaye, and R. F. Haglund. Plasmonic probe of the semiconductor to metal phase transition in vanadium dioxide. *Nano Lett.*, 13:4169–4175, 2013.

Appendix A

Finite Difference Time Domain Simulations

Theoretical Overview

The finite difference time domain (FDTD) simulation technique is a full-wave vectorial solution to Maxwell's equations. The most common implementation of FDTD is the Yee scheme [1]. By discretizing Maxwell's equations both spatially and temporally using the finite difference formulas for derivatives, a set of six equations is obtained, enabling calculation of each field component at a given point in space and time. These equations cannot be solved in a closed form, and require knowledge of the field components at the previous time step. Therefore, the simulation must be performed in a step-by-step manner. At each time step, the six field components are calculated at every mesh point and stored in memory before the program proceeds to the next temporal step. This process repeats itself in a loop until a predefined shutoff criterion is satisfied.

In this scheme, a time, t , can be referenced in its discrete form, $n\Delta t$, where $n \in \mathbb{Z} \geq 0$ and Δt is the temporal resolution. A unit cell of an arbitrary simulation space is shown in Fig. A.1. For a uniform, cartesian spatial discretization mesh, a coordinate, (x, y, z) , can be denoted in its discretized form, $(i\Delta x, j\Delta y, k\Delta z)$, where $\Delta x = \Delta y = \Delta z = \Delta$ is the separation between nearest neighbours of the spatial discretization mesh. This geometry represents the relative position of the field components. Notably, the \mathbf{E} and \mathbf{H} components are offset from one another by half a grid spacing, $\Delta/2$. Moreover, a field component at a point, $(i\Delta x, j\Delta y, k\Delta z)$, at a time $n\Delta t$, is denoted as $\Psi^n(i, j)$. For example, the E_z field component can be written as $E_z^n(i, j)$.

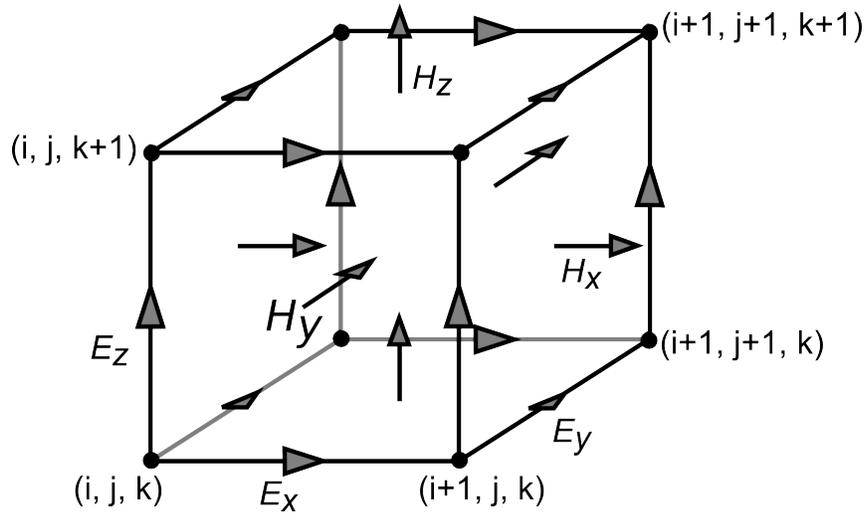


Figure A.1: A unit cell of the FDTD Yee scheme.

Finite Difference Expressions

First Order Derivatives

As a set of coupled first order differential equations, a discretization of Maxwell's equations requires an approximation for the derivative operator, e.g. $\partial\Psi/\partial x = \Psi_x$. For a field component, Ψ , at a point $(x_i, y_j, z_k) = (i\Delta x, j\Delta y, k\Delta z)$, the forward Taylor series expansion can be written as:

$$\begin{aligned} \Psi(x + \Delta x, y, z) &= \Psi(x, y, z) + \Delta x \cdot \Psi_x(x, y, z) \\ &\quad + \frac{\Delta x^2}{2!} \Psi_{xx}(x, y, z) + \dots \end{aligned} \quad (\text{A.1})$$

or

$$\begin{aligned} \Psi(i + 1, j, k) &= \Psi(i, j, k) + \Delta x \cdot \Psi_x(i, j, k) \\ &\quad + \frac{\Delta x^2}{2} \Psi_{xx}(i, j, k) + \dots \end{aligned} \quad (\text{A.2})$$

This can be rewritten as:

$$\begin{aligned} \Psi_x(i, j, k) &= \frac{\Psi(i + 1, j, k) - \Psi(i, j, k)}{\Delta x} - \frac{\Delta x^2}{2!} \\ &= \frac{\Psi(i + 1, j, k) - \Psi(i, j, k)}{\Delta x} + O(\Delta x), \end{aligned} \quad (\text{A.3})$$

which is known as the forward-difference formula. In a similar manner, the backward Taylor series expansion can be written as:

$$\begin{aligned} \Psi(i = 1, j, k) &= \Psi(i, j, k) - \Delta x \cdot \Psi_x(i, j, k) \\ &\quad + \frac{\Delta x^2}{2!} \Psi_{xx}(i, j, k) - \dots \end{aligned} \quad (\text{A.4})$$

Following the same procedure produces the backward-difference formula:

$$\Psi_x(i, j, k) = \frac{\Psi(i, j, k) - \Psi(i-1, j, k)}{\Delta x} + O(\Delta x) \quad (\text{A.5})$$

Notably, both the forward and backward difference formulas are accurate to $O(\Delta x)$. A more accurate first order derivative approximation can be obtained by subtracting (A.4) from (A.2):

$$\begin{aligned} \Psi(i+1, j, k) - \Psi(i-1, j, k) &= \left(\Psi(i, j, k) + \Delta x \cdot \Psi_x(i, j, k) \right. \\ &\quad \left. + \frac{\Delta x^2}{2} \Psi_{xx}(i, j, k) + \dots \right) \\ &\quad - \left(\Psi(i, j, k) - \Delta x \cdot \Psi_x(i, j, k) \right. \\ &\quad \left. + \frac{\Delta x^2}{2!} \Psi_{xx}(i, j, k) - \dots \right) \\ &= 2\Delta x \cdot \Psi_x(i, j, k) \\ &\quad + 2\frac{\Delta x^3}{3!} \Psi_{xxx}(i, j, k) + \dots \end{aligned} \quad (\text{A.6})$$

from which it can be written:

$$\Psi_x(i, j, k) = \frac{\Psi(i+1, j, k) - \Psi(i-1, j, k)}{2\Delta x} + O(\Delta x^2). \quad (\text{A.7})$$

This is known as the central difference formula.

Electromagnetic Field Update Equations

The central difference formula can be applied both spatially and temporally to Maxwell's curl equations, which are rewritten below for reference:

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} - \sigma^* \mathbf{H} \quad (\text{A.8})$$

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{H} \quad (\text{A.9})$$

Here, an equivalent magnetic loss term, σ^* has been included in case it is desirable to model magnetic field attenuation. The update equation for one field component, e.g. E_x , is developed as follows:

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \sigma E_x + \varepsilon \frac{\partial E_x}{\partial t}, \quad (\text{A.10})$$

can be rewritten as:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(-\sigma E_x + \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right). \quad (\text{A.11})$$

Applying the finite difference formulas at time $n + 1/2$ and mesh point i, j, k produces:

$$\begin{aligned} \frac{E_x|_{i,j,k}^{n+1} - E_x|_{i,j,k}^n}{\Delta t} &= \frac{1}{\varepsilon_{i,j,k}} \left[-\sigma_{i,j,k} \left(\frac{E_x|_{i,j,k}^{n+1} + E_x|_{i,j,k}^n}{2} \right) \right. \\ &\quad + \frac{H_z|_{i,j+1/2,k}^{n+1/2} - H_z|_{i,j-1/2,k}^{n+1/2}}{\Delta y} \\ &\quad \left. - \frac{H_y|_{i,j,k+1/2}^{n+1/2} - H_y|_{i,j,k-1/2}^{n+1/2}}{\Delta z} \right]. \end{aligned} \quad (\text{A.12})$$

Collecting like terms yields:

$$\begin{aligned} E_x|_{i,j,k}^{n+1} \left(\frac{1}{\Delta t} + \frac{\sigma_{i,j,k}}{2\varepsilon_{i,j,k}} \right) &= E_x|_{i,j,k}^n \left(\frac{1}{\Delta t} - \frac{\sigma_{i,j,k}}{2\varepsilon_{i,j,k}} \right) \\ &\quad + \frac{1}{\varepsilon_{i,j,k}} \left(\frac{H_z|_{i,j+1/2,k}^{n+1/2} - H_z|_{i,j-1/2,k}^{n+1/2}}{\Delta y} \right. \\ &\quad \left. - \frac{H_y|_{i,j,k+1/2}^{n+1/2} - H_y|_{i,j,k-1/2}^{n+1/2}}{\Delta z} \right). \end{aligned} \quad (\text{A.13})$$

Solving for $E_x|_{i,j,k}^{n+1}$ gives the update equation for E_x :

$$\begin{aligned} E_x|_{i,j,k}^{n+1} &= \left(\frac{1 - \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) E_x|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) \\ &\quad \cdot \left(\frac{H_z|_{i,j+1/2,k}^{n+1/2} - H_z|_{i,j-1/2,k}^{n+1/2}}{\Delta y} - \frac{H_y|_{i,j,k+1/2}^{n+1/2} - H_y|_{i,j,k-1/2}^{n+1/2}}{\Delta z} \right). \end{aligned} \quad (\text{A.14})$$

The remaining update equations can be obtained in the same manner and are written below for reference:

$$\begin{aligned} E_y|_{i,j,k}^{n+1} &= \left(\frac{1 - \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) E_y|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) \\ &\quad \cdot \left(\frac{H_x|_{i,j,k+1/2}^{n+1/2} - H_x|_{i,j,k-1/2}^{n+1/2}}{\Delta z} - \frac{H_z|_{i+1/2,j,k}^{n+1/2} - H_z|_{i-1/2,j,k}^{n+1/2}}{\Delta x} \right) \end{aligned} \quad (\text{A.15})$$

$$\begin{aligned} E_z|_{i,j,k}^{n+1} &= \left(\frac{1 - \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) E_z|_{i,j,k}^n + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}\Delta t}{2\varepsilon_{i,j,k}}} \right) \\ &\quad \cdot \left(\frac{H_y|_{i+1/2,j,k}^{n+1/2} - H_y|_{i-1/2,j,k}^{n+1/2}}{\Delta x} - \frac{H_x|_{i,j+1/2,k}^{n+1/2} - H_x|_{i,j-1/2,k}^{n+1/2}}{\Delta y} \right) \end{aligned} \quad (\text{A.16})$$

$$H_x|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k}^*\Delta t}{2\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^*\Delta t}{2\mu_{i,j,k}}} \right) H_x|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^*\Delta t}{2\mu_{i,j,k}}} \right) \quad (\text{A.17})$$

$$H_y|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}} \right) H_y|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}} \right) \cdot \left(\frac{E_y|_{i,j,k+1/2}^n - E_y|_{i,j,k-1/2}^n}{\Delta z} - \frac{E_z|_{i,j+1/2,k}^n - E_z|_{i,j-1/2,k}^n}{\Delta y} \right) \quad (\text{A.18})$$

$$H_z|_{i,j,k}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}} \right) H_z|_{i,j,k}^{n-1/2} + \left(\frac{\frac{\Delta t}{\mu_{i,j,k}}}{1 + \frac{\sigma_{i,j,k}^* \Delta t}{2\mu_{i,j,k}}} \right) \cdot \left(\frac{E_z|_{i+1/2,j,k}^n - E_z|_{i-1/2,j,k}^n}{\Delta x} - \frac{E_x|_{i,j,k+1/2}^n - E_x|_{i,j,k-1/2}^n}{\Delta z} \right) \quad (\text{A.19})$$

$$\cdot \left(\frac{E_x|_{i,j+1/2,k}^n - E_x|_{i,j-1/2,k}^n}{\Delta y} - \frac{E_y|_{i+1/2,j,k}^n - E_y|_{i-1/2,j,k}^n}{\Delta x} \right).$$

Metals

To model metallic media, the Drude model was incorporated into the FDTD algorithm. The permittivity of the metal can be introduced into the FDTD time-stepping scheme by adding an auxiliary differential equation relating \mathbf{E} and \mathbf{D} :

$$\gamma_0 \frac{\partial \mathbf{D}}{\partial t} + \frac{\partial^2 \mathbf{D}}{\partial t^2} = \omega_p^2 \varepsilon_0 \mathbf{E} + \gamma_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (\text{A.20})$$

where γ_0 is the damping coefficient and ω_p is the plasma frequency. By discretizing this equation and calculating \mathbf{D} in each time step, the behaviour of metals can be included.

Berenger Perfectly Matched Layer

The perfectly matched layer is an absorbing boundary condition that is designed to attenuate signals as they interact with the simulation space boundary [2]. By impedance matching the absorbing layer to the surrounding medium, reflections back into the simulation space are minimal and the simulation space is effectively infinite. The reflection due to any electromagnetic impedance mismatch can be written as:

$$\frac{E_r}{E_i} = \frac{Z_0 - Z_1}{Z_0 + Z_1}, \quad (\text{A.21})$$

where E_0 is the incident electric field, E_r is the reflected electric field, Z_0 is the impedance of the medium in which the incident wave travels, and Z_1 is the impedance of the medium upon which the incident wave impinges. The expressions for Z_0 and Z_1 can also be written as:

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \quad (\text{A.22})$$

$$Z_1 = \sqrt{\frac{\mu_1}{\varepsilon_1}}. \quad (\text{A.23})$$

Impedance matching must be satisfied to minimize reflections, i.e. $E_r = 0$:

$$Z_0 = Z_1 \quad (\text{A.24})$$

$$\frac{\mu_0}{\varepsilon_0} = \frac{\mu_1}{\varepsilon_1}. \quad (\text{A.25})$$

Furthermore, a finite electric conductivity, σ , and magnetic conductivity, σ^* , must be introduced to the boundary material to achieve attenuation. A non-physical version of Maxwell's equations can be written to include magnetic loss via magnetic currents:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} - \sigma^* \mathbf{H} \quad (\text{A.26})$$

$$\nabla \times \mathbf{H} = \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} \quad (\text{A.27})$$

To implement a perfectly matched layer, each of Maxwell's equations are split into two separate equations, for a total of twelve equations:

$$\varepsilon_0 \frac{\partial E_{xy}}{\partial t} + \sigma_y E_{xy} = \frac{\partial(H_{zx} + H_{zy})}{\partial y} \quad (\text{A.28})$$

$$\varepsilon_0 \frac{\partial E_{xz}}{\partial t} + \sigma_z E_{xz} = -\frac{\partial(H_{yx} + H_{yz})}{\partial z} \quad (\text{A.29})$$

$$\varepsilon_0 \frac{\partial E_{yz}}{\partial t} + \sigma_z E_{yz} = \frac{\partial(H_{xy} + H_{xz})}{\partial z} \quad (\text{A.30})$$

$$\varepsilon_0 \frac{\partial E_{yx}}{\partial t} + \sigma_x E_{yx} = -\frac{\partial(H_{zx} + H_{zy})}{\partial x} \quad (\text{A.31})$$

$$\varepsilon_0 \frac{\partial E_{zx}}{\partial t} + \sigma_x E_{zx} = \frac{\partial(H_{yx} + H_{yz})}{\partial x} \quad (\text{A.32})$$

$$\varepsilon_0 \frac{\partial E_{zy}}{\partial t} + \sigma_y E_{zy} = \frac{\partial(H_{xy} + H_{xz})}{\partial y} \quad (\text{A.33})$$

$$\mu_0 \frac{\partial H_{xy}}{\partial t} + \sigma_y^* H_{xy} = -\frac{\partial(E_{zx} + E_{zy})}{\partial y} \quad (\text{A.34})$$

$$\mu_0 \frac{\partial H_{xz}}{\partial t} + \sigma_z^* H_{xz} = \frac{\partial(E_{yx} + E_{yz})}{\partial z} \quad (\text{A.35})$$

$$\mu_0 \frac{\partial H_{yz}}{\partial t} + \sigma_z^* H_{yz} = -\frac{\partial(E_{xy} + E_{xz})}{\partial z} \quad (\text{A.36})$$

$$\mu_0 \frac{\partial H_{yx}}{\partial t} + \sigma_x^* H_{yx} = \frac{\partial(E_{zx} + E_{zy})}{\partial x} \quad (\text{A.37})$$

$$\mu_0 \frac{\partial H_{zx}}{\partial t} + \sigma_x^* H_{zx} = -\frac{\partial(E_{yx} + E_{yz})}{\partial x} \quad (\text{A.38})$$

$$\mu_0 \frac{\partial H_{zy}}{\partial t} + \sigma_y^* H_{zy} = \frac{\partial(E_{xy} + E_{xz})}{\partial y} \quad (\text{A.39})$$

These equations can then be discretized using an exponential time-stepping approach to obtain the PML field update equations.

Massively Parallel Implementation

Due to its widespread use, versatility, and the availability of comprehensive scientific libraries, this simulation software was implemented in C++. In its final version, this software consisted of over 11000 lines of code, and would occupy approximately 300 pages. For this reason, the code has been omitted from this appendix.

The computationally demanding nature of full wave electromagnetic simulations provided motivation to implement a simulation software that allows the problem at hand to be divided into smaller pieces and run on separate processor cores. Naturally, the different cores cannot perform the simulation in a completely independent manner; they must communicate with one another throughout the course of the simulation. MPICH2 is one of the most popular implementations of MPI (message passing interface) and provides convenient commands for synchronizing individual cores and exchanging information between them. This library formed the basis for the parallel processor implementation of this FDTD code.

Several approaches to dividing the FDTD algorithm into multiple tasks are possible. However, it would be the most desirable to divide the simulation problem into an arbitrary number of sub-problems and run them on a cluster (i.e. a massively parallel implementation). The most flexible approach for doing so is to spatially divide the simulation space. In this way, each core participating in the simulation is responsible for one section of the simulation geometry. At each time step, every core updates the fields in its section of geometry. When all of the cores have finished, they begin exchanging the tangential components of the electromagnetic fields in the planes that border neighbouring simulation spaces. Once this is complete, the program proceeds to the next time step.

While it would be simple to divide the simulation space for a fixed number of cores, such as 2, 4, 8, or 16, it is much more challenging to divide the simulation space for an unknown number of cores. The idea of a simulation “cell” was conceived to achieve this possibility. Before running the simulation, the user defines how many “cells” to divide the simulation space into in each dimension. For example, if the user had access to a 32 core cluster, they can wish to divide the simulation space into $x \times y \times z = 4 \times 4 \times 2$ cells. During the initialization stage of the simulation, the cells are created and each cell determines which section of the geometry that it is responsible for. In addition, the cells must divide up any sources or detectors that span more than one cell. The cells then systematically determine which cells it borders. The cells need this information so that they can communicate the plane of tangential fields to their neighbouring cells.

Upon completion of the initialization stages, the program moves into the time-looping. As stated earlier, each cell completes its field updates for that time step and then sends (receives) the necessary planes of tangential to (from) its neighbours before proceeding to the next time step.

References

- [1] K. Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:302–307, 1966.
- [2] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200, 1994.

Appendix B

Nanofabrication Process Steps

Wafer Cleaning

Piranha Cleaning

Piranha solution is used to remove organics from the substrate before any fabrication work is begun. Piranha is prepared by adding 30% hydrogen peroxide to 96% sulfuric acid in a 1:3 ratio. Hydrogen peroxide is a base and sulfuric acid is an acid. Each of these two chemicals will react with and remove organics from the substrate. In addition, the chemical reaction between the acid and base is exothermic and the mixture reaches boiling temperatures, which further aids in dissolving organics.

Although highly effective for cleaning substrates, there are several subtleties that must be mentioned. First, hydrogen peroxide reacts with silicon to form a thin oxide layer. If subsequent steps require silicon etching, this oxide could interfere and produce undesired results, depending on the etching technique. In general, it can be desirable to piranha clean at any intermediate step when it is possible during the process flow. However, it should be noted that piranha slowly etches metals. Therefore, if there are already metal features on the substrate, no further piranha cleaning should be performed.

Buffered Oxide Etching

As mentioned in the previous section, a piranha-cleaned silicon substrate will have a thin oxide layer on its surface. Buffered oxide etching can be used to remove this oxide layer before proceeding to subsequent steps. Buffered oxide etchant is a 10:1 mixture of 40% ammonium fluoride and 49% hydrofluoric acid. As suggested by the name, the hydrofluoric acid is buffered by the ammonium fluoride. This is done for two reasons: (1) To reduce the etch rate to a controlled 2nm/s at room temperature; and (2) To maintain a consistent etch rate over long etching durations, further ensuring precision and allowing reuse of the solution.

Electron Beam Lithography

Electron beam lithography (EBL) is a fabrication technique that can be used to fabricate structures with sub-100nm dimensions [1]. Whereas the minimum feature size of a photolithography mask is limited by the diffraction limit, sophisticated electron beam optics allow for linewidths down to 5nm, depending on the resist, materials, and exposure parameters used. Because plasmonic devices are inherently sub-diffraction, photolithography simply does not allow for development of fine enough features and EBL must be used. This is particularly true of the photolithography systems in the NanoFab, where the resolution is limited to $\sim 1\mu\text{m}$.

Components of an Electron Beam Lithography System

There are two main components to an electron beam lithography system: (1) The column, which forms and positions the electron beam; and (2) the stage and its controller, which position the sample that is being exposed.

Within the column, there must be a source of electrons. This source can be either a thermionic source or a field emission source. In a thermionic source, a conducting material is heated until the energy of its constituent electrons exceeds its work function barrier. In a field emission source, the work function barrier is overcome by applying an electric field that is strong enough to permit tunneling through the barrier. More recently, these two techniques have been combined, resulting in thermal field emission sources. In this source, a sharp tungsten needle (radius of less than $1\mu\text{m}$) is heated to around 1800 K and an electric field is applied between the tips.

A source can be characterized by three parameters known as the virtual source size, the brightness, and the energy spread of emitted electrons. The size of the source determines the factor by which the lenses in the column must demagnify the electrons to achieve focusing on the sample. The brightness of the source is related to the current in the electron beam: the brighter the electron source, the higher the current. If the beam contains electrons with a wide range of energies, it is analogous to white light. Conversely, if the beam contains electrons with a narrow line of energies, it is analogous to monochromatic light. A monochromatic beam is desirable because the electron optics will act uniformly on a beam with a single energy of electrons, allowing for higher resolution and positioning accuracy.

Electron lenses must be used to bring the emitted electrons to a focus, which can use either electrostatic or magnetic fields. These lenses suffer from worse aberrations than their optical counterparts, namely spherical aberrations (the focal length of the lens changes as you move further from the central axis) and chromatic aberrations (electrons of different energies are focused differently). These aberrations can be reduced by utilizing only electrons that are incident on the centre of the lens, resulting in higher beam quality and reduced current. A magnetic lens works on the principle of the force that is imparted

on electrons in the presence of a diverging magnetic field. A diverging magnetic field along the optical axis produces a force that brings the electrons towards the optical axis, effectively focusing the electrons. The magnetic field also causes the electrons to rotate about the optical axis, following a helical path. This helical path can introduce complications in the beam positioning which must be corrected for. Electrostatic lenses have worse aberrations than magnetic lenses and are not used as often.

Spray apertures are used in the column to stop deviant electrons without altering the beam. Blanking apertures are used to turn the beam on and off, and act to deflect the beam away from the aperture hole. Beam limiting apertures control the beam convergence angle of the electrons as they traverse through the system and also sets the beam current. Finally, the beam is scanned across the sample by deflection elements, which can be based on either magnetic or electrostatic effects.

Pattern Exposure and Development

Regardless of how finely a beam can be focused, the situation becomes more complicated as the electrons interact with a resist. As electrons traverse through the resist, some will undergo small angle scattering events, leading to a larger effective beam diameter at the bottom of the resist. These are known as the forward scattered electrons. The increase in beam diameter in nanometers has been determined empirically and can be expressed by the following formula:

$$d_f = 0.9 \left(\frac{R_t}{V_b} \right)^{1.5}, \quad (\text{B.1})$$

where R_t is the resist thickness in nanometers and V_b is the beam voltage in kilovolts.

In addition, many electrons will penetrate through the resist, interact with the substrate and undergo large angle scattering interactions. These electrons can return through the resist at an offset from the incident electron beam, resulting in unwanted resist exposure, known as the electron beam proximity effect. The fraction of electrons that are backscattered depends primarily on substrate material and is roughly 0.17 for silicon. The distance an electron can return through the resist depends on the energy of the incident electron as well as the substrate properties.

As the primary electrons lose energy, a significant portion of their remaining energy is consumed by secondary electrons with energies from 2 to 50 eV. These electrons are generated as electrons that exceed the ionization potential of the resists' constituent molecules interact with the resist. These electrons are responsible for a majority of the resist exposure. Due to their low energy, their range in the resist is only a few nanometers and they contribute very little to the proximity effect. Traditionally, it is challenging to achieve consistent features with dimensions less than 20nm and advanced techniques must be pursued in this range.

Magnetron Sputtering of Metals

Magnetron sputtering belongs to a broad class of material deposition techniques known as physical vapour deposition (PVD) [2]. As suggested by the name, this involves a physical process whereby the material to be deposited condenses from a vapourized form onto the substrate surface, forming a thin film. Due to the sensitivity of these processes, they are carried out in a high vacuum chamber. In general, deposition requires a material source, an apparatus used to obtain a vapour from that source, and a means of depositing that vapour onto the substrate of interest.

In sputter deposition, the material which is to be deposited (the target) and the substrate holder are situated in a vacuum chamber. The substrate holder is set to rotate so that the deposition material is deposited uniformly over the substrate surface. Argon gas is introduced into the chamber and a negative voltage is applied to the target, creating an electric field in the chamber. A plasma is lit when free electrons of the target material are ejected and have sufficient energy to ionize nearby argon atoms. The positively-charged argon ions are accelerated towards the target by the electric field and undergo collisions with the surface atoms of the target. Provided that the electric field in the chamber is high enough, the argon ions have sufficient energy to ablate the target. During the interaction between a particular argon ion and the target surface, the argon ion undergoes a series of collisions and several target atoms can be vapourized. Some of these target atoms are released with a momentum vector that allows for a trajectory that intersects the substrate surface, and these atoms contribute to the formation of a thin film on the substrate.

Besides target ions being freed during the ion bombardment, primary electrons and ions are also emitted. These electrons accelerate in the electric field and ionize more argon atoms, resulting in more argon ions. The free electrons produce a plasma of argon ions, which vapourize the target, and this vapour deposits on the substrate.

Despite the early success of diode sputtering, there are severe limitations that hinder the film quality. As stated earlier, the presence of primary electrons is required to maintain a plasma in the chamber. As the electron energy increases and argon pressure decreases, the mean free path of the primary electrons increases. For low argon pressures, argon atoms will be ionized further from the target on average and are less likely to collide with the target. In addition, the primary electrons are more likely to collide with the substrate, degrading the thin film quality. For such low pressures, ionization efficiencies are low, and self-sustained discharges cannot be maintained at pressures below about 10mTorr. Conversely, if pressures are too high, sputtered atoms are more likely to collide with argon atoms and never reach the substrate.

These limitations were largely reduced by the development of magnetron sputtering systems. Magnets are situated behind the cathode and set up a magnetic field that causes the electrons to follow a circuit in close proximity to the target. As a result of this, argon atoms are ionized much closer to the target, increasing the sputtering rate drastically.

In addition, high-energy electrons are kept away from the substrate, enhancing the film quality. This also allows for a stable discharge to be sustained at a reduced argon pressure. Operating at a reduced pressure decreases the probability of a target atom undergoing a scattering interaction, which in turn improves the efficiency of target atom adhesion and the deposition rate.

Electron Beam Evaporation

Electron beam physical vapour deposition, or electron beam evaporation, is performed under high vacuum ($\sim 10^{-6}$ Torr). A tungsten filament is used to obtain a beam of electrons, which are directed on a target of the material to be deposited. When incident on the target, the electron energy is dissipated very rapidly and is converted to thermal energy. For sufficient electron beam current, the surface of the target will melt, and eventually, a vapour will be produced. This vapour is emitted from the sample with a high directivity and can precipitate on the sample, coating it with a thin layer of the target material in a highly directional fashion.

Inductively-Coupled Plasma Reactive Ion Etching

Background

Two common wet etchants for silicon are potassium hydroxide (KOH) and tetramethylammonium hydroxide (TMAH). When a silicon wafer is immersed in bath of etchant, the wafer is etched anisotropically, i.e. the etching rate varies depending on the crystal plane. The (100) and (110) planes etch much more rapidly than the (111) plane. As a result, the (100) and (111) planes form a 57.4° angle with one another. In many fabrication processes, it is desirable to etch features with 90° sidewalls into a silicon wafer. The Bosch process is an elegant means of achieving this and was originally developed by Robert Bosch.

Bosch Process

The Bosch process is based on two steps that are repeated until the desired etching depth is obtained [3]. The Bosch process is a dry etching process, which means that the entire process is performed by exposing the sample to dry gases in a vacuum chamber rather than immersing the sample in a liquid. The sample can be masked with a soft mask such as a resist, or a hard mask such as metal or silicon dioxide. In the first step of the Bosch process, the substrate is etched by releasing SF_6 into the chamber. A plasma is formed, resulting in the presence of fluorine radicals. These react with the silicon substrate to produce SiF_4 gas, which is pumped out of the system by the vacuum pump. In this case, the etch is isotropic, and will produce a notch in unmasked regions of the silicon substrate. For this reason, the etch is only carried out for short time, so that the etch depth is roughly

a micron or less. This is depicted schematically in Fig. B.1(a). In the second step of the process, the passivation step, C_4F_8 is released into the chamber. Through interactions in the plasma, the C_4F_8 molecules link together and form a teflon-like polymer. This polymer deposits on the sample and prevents the etch-plasma from interacting with the silicon. During the next etch step, the polymer must be removed from the bottom of the existing etch, but maintained on the sidewalls, so that vertical etching is achieved. To do this, the ions in the plasma physically sputter the polymer, revealing the silicon surface. These ions are accelerated by the electric field in the chamber, and bombard the sample at normal incidence. As a result, the sidewalls are not sputtered, but the bottom of the etch is. This is depicted schematically in Fig. B.1(b). These etch and deposition steps are then repeated in a cyclic manner until the desired etch depth is achieved. Although vertical sidewalls are achieved, there are ripples present in the sidewalls, which can be unacceptable for some applications.

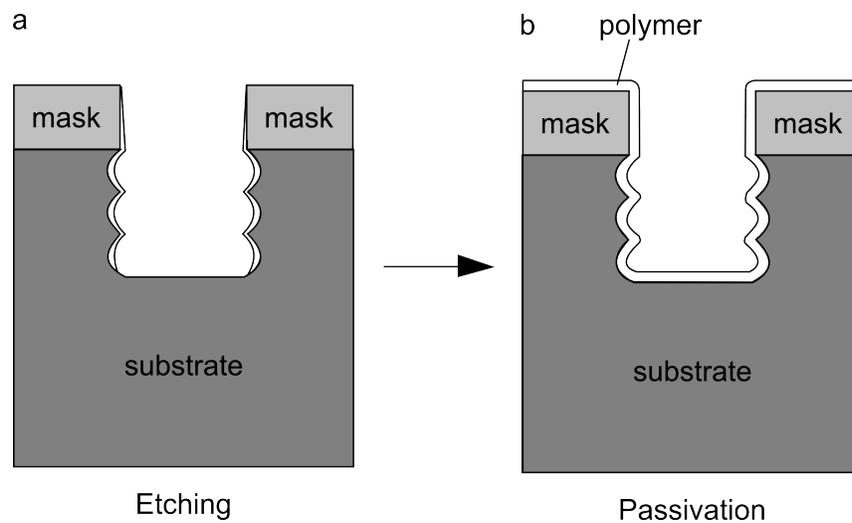


Figure B.1: Schematic depiction of the (a) passivation and (b) etching steps of the Bosch process.

References

- [1] A. A. Tseng, K. Chen, C.D. Chen, and K.J. Ma. Electron beam lithography in nanoscale fabrication: recent development. *Electronics Packaging Manufacturing, IEEE Transactions on*, 26:141–149, 2003.
- [2] P. J. Kelly and R. D. Arnell. Magnetron sputtering: a review of recent developments and applications. *Vacuum*, 56:159–172, 2000.
- [3] B. Wu, A. Kumar, and S. Pamarthy. High aspect ratio silicon etch: A review. *J. Appl. Phys.*, 108:051101, 2010.

Appendix C

Calculation of the Theoretical Third-Harmonic Generation Spectrum

The ideal third-harmonic spectrum from the laser pulses is calculated from the free-space spectrum of radiation emitted from the femtosecond laser. An exemplary free-space spectrum is shown in Fig. C.1(a). Mixing between any three constituent frequencies, $\omega = \{\omega_i, \omega_j, \omega_k\}$, of the laser radiation spectrum in the nonlinear medium can produce a green photon whose energy is given by:

$$\omega_{ijk} = \omega_i + \omega_j + \omega_k \quad (\text{C.1})$$

In general, spectrometer data can be represented as a set of discrete points by the equation:

$$I(\omega) = \sum_{i=1}^{N_i} a_i \delta[\omega - \omega_i] \quad (\text{C.2})$$

where $I(\omega)$ is the spectral intensity distribution, a_i is the relative intensity of the component at $\omega = \omega_i$, N_i is the total number of frequency components, and $\delta[\omega]$ is the Kronecker delta function, defined as:

$$\delta[\omega] = \begin{cases} 0, & \text{if } \omega \neq 0 \\ 1, & \text{if } \omega = 0 \end{cases} \quad (\text{C.3})$$

A weighted sum of any three frequency components from the laser spectrum can then be written as:

$$I(\omega = \omega_i + \omega_j + \omega_k) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \sum_{k=1}^{N_k} a_i a_j a_k \delta[\omega - (\omega_i + \omega_j + \omega_k)] \quad (\text{C.4})$$

Based on the above expression, the ideal third harmonic spectrum is calculated from the femtosecond laser pulses, and is shown in Fig. C.1(b). It is important to note that this expression does not account for dispersion or absorption in the linear or nonlinear properties of the silicon core of the nanoplasmonic waveguides under consideration. In addition, it

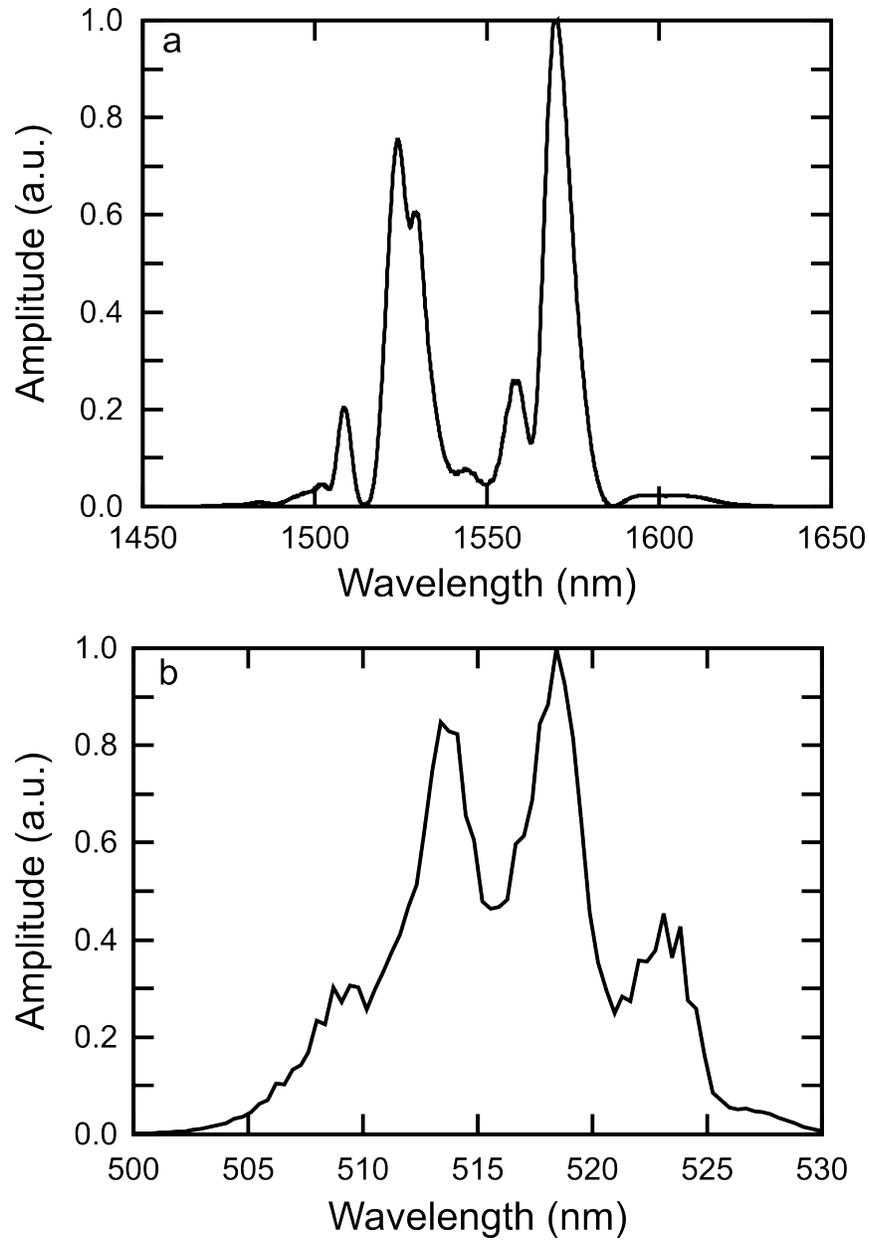


Figure C.1: (a) Femtosecond laser spectrum, measured in free-space. (b) Calculated third-harmonic spectrum.

does not account for any spectral non-uniformity in the coupling efficiency from free-space to the nanoplasmonic waveguides. The code used to perform the calculations is included below:

```

%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
%% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %% %%
function [] = calc_THG_spectra()

load 'lambda.mat'
load 'amplitude.mat'
[x,y] = size(lambda);
n = 9;
count = 0;

for a = ((n-1)/2+1):n:x-(n-1)/2
    count = count+1;
    l_fund(count) = lambda(a);
    amp_fund(count) = sum(amp((a-(n-1)/2):(a+(n-1)/2)));
end

amp_fund = amp_fund/max(amp_fund);

[x,y] = size(l_fund);
s_min = 1;
s_max = y;

l1 = 200;
l2 = 1000;
dl = .1;

[a,h] = hdrload('54.SCO');
l_THG = h(:,1);
[x1,y1] = size(l_THG);
Power = zeros(x1);

for a = s_min:s_max
    for b = s_min:s_max
        for c = s_min:s_max
            l = 1/(1/l_fund(a) + 1/l_fund(b) + 1/l_fund(c));
            P = amp_fund(a)*amp_fund(b)*amp_fund(c);
            [I,cv] = searchclosest(l_THG,l);
            Power(I) = Power(I) + P;
        end
    end
end

save power.mat Power

```

Appendix D

The Pump-Probe Time-Domain Technique

In this thesis, pump-probe measurements are performed to gain insight into time-dynamics of free-carrier recombination and free-carrier sweeping in nanophotonic and nanoplasmonic waveguides. The temporal resolution of this technique is limited only by the pulse duration, which is $\tau_p \approx 84\text{fs}$ in the experiments presented in this thesis. Therefore, this technique can be used to obtain insight into ultrafast dynamics with a much higher temporal resolution than electronic techniques.

In pump-probe measurements, the femtosecond pulses are split into two separate beams using a beam splitter. Typically, the peak pump power coupled to the waveguide of interest was in the range of $5\text{W} \leq P_{peak}^{pump} \leq 40\text{W}$ and the peak probe power was approximately $P_{peak}^{probe} = 2.75\text{W}$. The probe is directed to a retroreflector that is mounted on a PC-controlled stepper motor, which controls the arrival time of the probe to the device under test with a temporal resolution of $\Delta t = 24\text{fs}$. The pump undergoes a fixed delay to compensate for the optical path length difference accrued by the probe during its delay. The pump beam is mechanically chopped at $F_{chop} = 1.4\text{kHz}$ to allow for lock-in detection. The pump and probe are then recombined using a beam-splitter and propagated collinearly towards the microscope objective. The microscope objective end-fire couples both the pump and probe into the waveguide, and a lensed optical fibre out-couples the transmitted pump and probe signals and delivers them to a fast photodiode (Electro-optics Technology, Inc. ET-3000). A schematic depiction of the pump-probe system is shown in Fig. D.2. The signal from the fast photodiode is then delivered to a lock-in amplifier system (Stanford Research Systems SR830). The lock-in amplifier is synchronized with the mechanical chopper that was applied to the pump beam. The lock-in amplifier includes a digital low-pass filter with an adjustable integration time-constant, which enables some control over the signal-to-noise ratio (SNR) of the measured signal and the total duration of the scan. In general, there is a trade-off between the scan duration and the SNR. In most cases an integration time constant of 30ms was used for the

experiments presented in this thesis. A Matlab control program was used to automate and synchronize the stepper motor and lock-in detection setups, allowing for consistent pump-probe time-domain traces to be obtained.

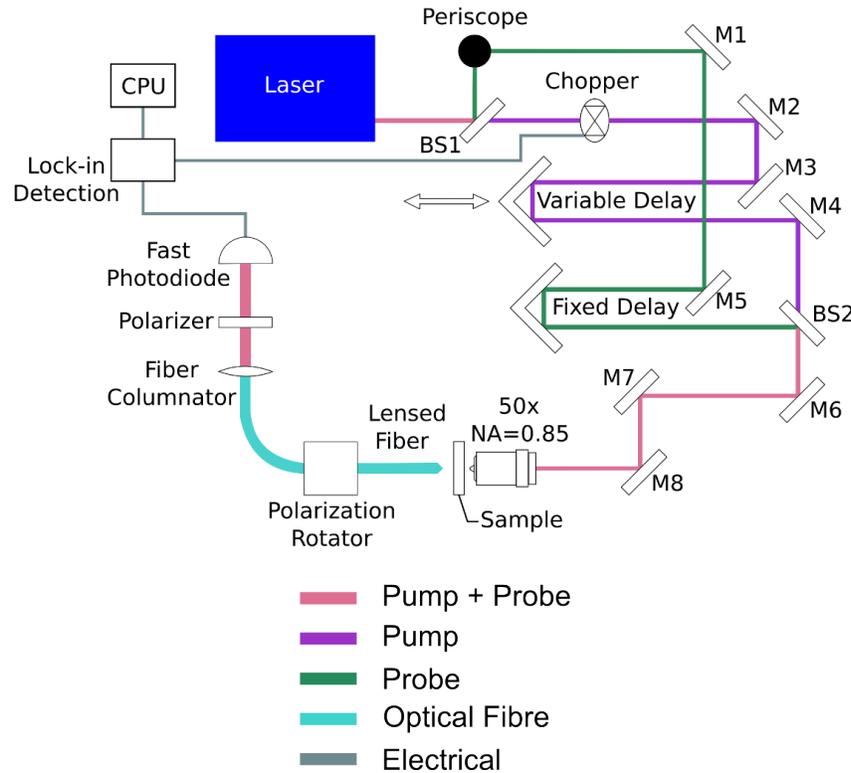


Figure D.1: Schematic representation of the pump-probe system.

In many cases, it is desirable to detect only the probe radiation. There are several common ways to achieve this, which are outlined below:

- Spatial separation - in the case of large samples, where tight focusing and sensitive alignment are not required, it is sufficient to propagate the pump and probe towards the sample at different angles. Provided that there is sufficient overlap between the pump and probe when interacting in the sample, the excitations caused by the pump will be mapped onto the probe. However, the pump and probe will be transmitted (or reflected, depending on the configuration) at different angles, and therefore, by careful placement of a photodetector and an aperture, the pump can be blocked.
- Spectral separation - in the case where a laser drives either a second- or third-harmonic generation module, or an optical parametric amplification module, the pump and probe can have different frequency components. Therefore, by passing the transmitted signal through an appropriate spectral filter, the pump can be filtered out.
- Polarization separation - the pump and probe can have orthogonal polarization states. Therefore, by passing the transmitted signal through a polarizer, only the

probe can be detected.

In the pump-probe experiments, the pump is used to excite TPA and FCA in a waveguide, while the probe is used to observe the change in transmission through the waveguide during the excitation and relaxation processes. Two different pump-probe configurations were used in this thesis to measure free-carrier dynamics. In the first, both the pump and the probe polarizations were oriented parallel to one another and to excite the TM mode of the waveguide. The loss due to TPA scales with the square of the input intensity, i.e.

$$\frac{dI}{dz} = -\beta_{TPA}I^2, \quad (\text{D.1})$$

where I is the input intensity, β_{TPA} is the TPA coefficient, and z is the distance along the waveguide axis in the direction of propagation. The loss due to FCA scales linearly with the input intensity, i.e.

$$\frac{dI}{dz} = -\sigma_{FCA}NI, \quad (\text{D.2})$$

where σ_{FCA} is the free-carrier absorption cross-section and N is the free-carrier density. Therefore, the TPA interaction will be the dominant feature present in the pump-probe trace. In specific, the dip in transmission when the pulse peaks overlap (and fields add constructively) will be the main feature and the recovery tail due to FCA will be negligible. In this case, both the pump and probe are detected simultaneously (i.e. no measures are taken to eliminate the pump), and the result is essentially an autocorrelation trace.

In the second configuration, the polarization of the pump and the probe fields are orthogonal. In this case, the transmitted pump radiation is blocked with a polarizer. In the case of the waveguides studied in this thesis, the orthogonal polarization makes the magnitude of the TPA dip less pronounced because the superposition of the pulse peaks will be greatly reduced, enabling more sensitive probing of the recovery tail (since FCA does not depend on polarization). Schematic renderings of the femtosecond laser, pump-probe system, and waveguide characterization setup are shown in Fig. D.2.

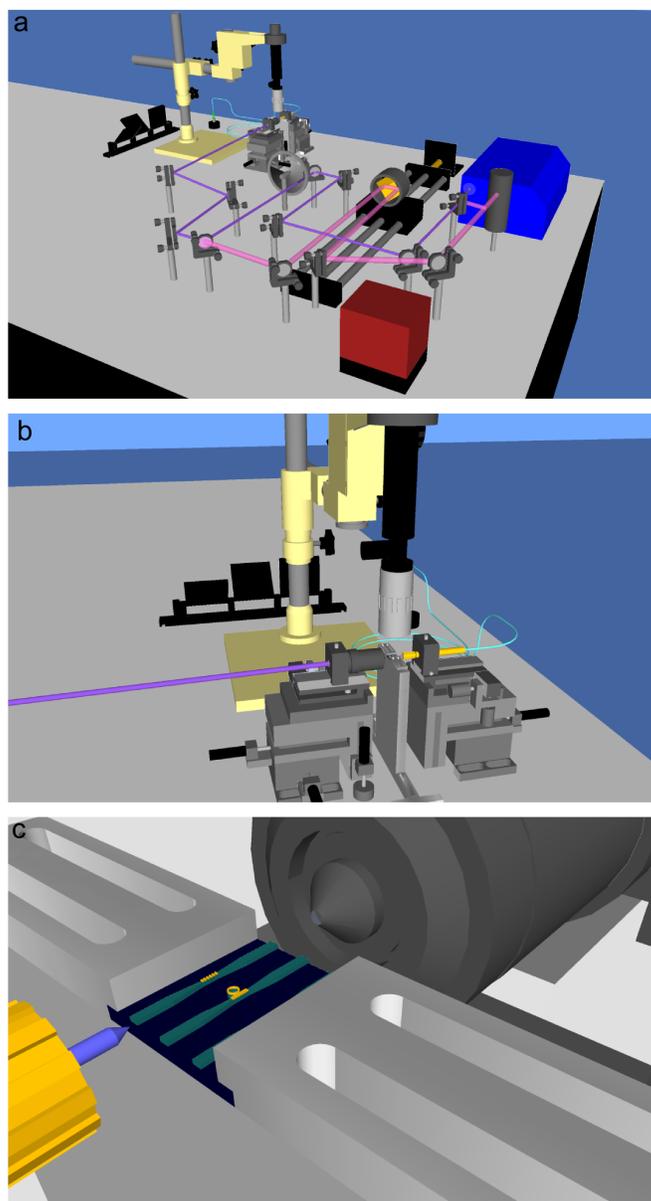


Figure D.2: Views of the pump-probe system.

Appendix E

Spectrometer Calibration

To determine the conversion efficiency of the third-harmonic generation (THG) process, it was necessary to obtain a correlation between the counts registered by the optical spectrum analyser (Ocean Optics S2000) and the third-harmonic power. The third-harmonic spectrum was centred at $\lambda = 517\text{nm}$. Therefore, coherent green radiation at $\lambda = 532\text{nm}$ was used to calibrate the spectrometer in the vicinity of the THG radiation. This radiation was generated from a commercially available solid state green laser ($\lambda = 532\text{nm}$) that used a second-harmonic generation crystal to double the frequency of a $\lambda = 1064\text{nm}$ laser diode. A bandpass filter was placed at the output aperture of the device to eliminate any emitted $\lambda = 1064\text{nm}$ radiation. The free-space power, P_{fs} , of the $\lambda = 532\text{nm}$ radiation was varied between $10\text{mW} \leq P_{fs} \leq 50\text{mW}$ using a system of calibrated neutral density filters and a precision power meter. The radiation was coupled into the tip of a lensed single-mode optical fibre (the same one that was used in the experiments) and the exact power coupled to the fibre, P_{fib} , was measured to be in the range, $190 \leq P_{fib} \leq 925\text{pW}$, as shown in Fig. E.1. For each coupled power, a spectrum was accumulated over $\tau_{int} = 293\text{ms}$ and 10 averages were taken. The total photon counts detected by the spectrometer were obtained by summing the spectral points near $\lambda = 532\text{nm}$ that were above the noise floor. In this way, a relationship between the THG light power propagating through the fibre and the number of photon counts detected by the spectrometer was obtained. The result was a linear calibration line, which is shown in Fig. E.1(b). Using a linear fit to this data, it was determined that a green light power in the fibre, $P_{fib} = 1.0\text{nW}$, produced 17740 photon counts on the spectrometer during this time interval, corresponding to a calibration factor of $59\text{fW} \cdot \text{count}^{-1} \cdot \text{s}^{-1}$.

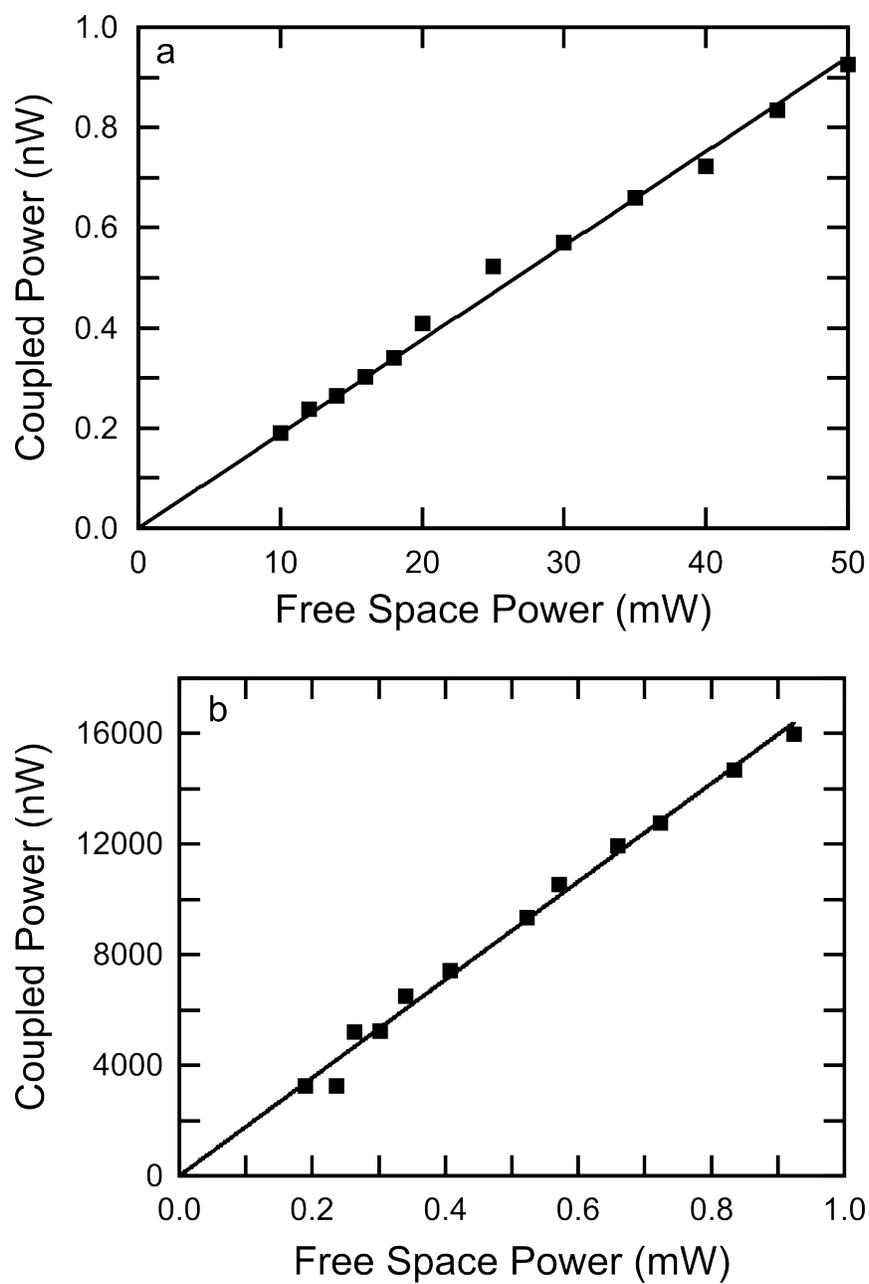


Figure E.1: (a) Plot of fibre-coupled power versus free-space laser power. (b) Plot of measured spectrometer versus fibre-coupled power.

Appendix F

Wavelength Scan Program

To characterize the broadband transmission spectra of the nanophotonic and nanoplasmonic structures, it was necessary to automate the tuning of the continuous-wave laser wavelength and the subsequent reading of the transmitted power. In order to do this, a Matlab-based graphical user interface (GUI) was created. This program permits the user to specify the starting wavelength of the scan, finishing wavelength, spectral resolution, the number of times to average the readings, and the file name to save the data to. When the scan is started, the program will tune the laser to the appropriate wavelength, read the laser power, and then read the power transmitted through the device from an optical spectrum analyser. The transmitted power is then normalized to the laser power to ensure that instabilities in the laser power are not mapped onto the results. The laser is then tuned to the next wavelength and the procedure is repeated. In this way a broadband transmission spectrum of the device is constructed one point at a time and the final data is saved to the filename specified by the user. The code used to create the GUI is included below:

```
function varargout = Laser_Scan(varargin)
% Begin initialization code
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',   gui_Singleton, ...
                  'gui_OpeningFcn',  @Laser_Scan_OpeningFcn, ...
                  'gui_OutputFcn',   @Laser_Scan_OutputFcn, ...
                  'gui_LayoutFcn',   [], ...
                  'gui_Callback',    []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code
```

```

% — Executes just before Laser_Scan is made visible.
function Laser_Scan_OpeningFcn(hObject, eventdata, handles, ...
    varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% handles    structure with handles and user data (see GUIDATA)
% varargin   command line arguments to Laser_Scan (see VARARGIN)

% Choose default command line output for Laser_Scan
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);

% UIWAIT makes Laser_Scan wait for user response (see UIRESUME)
% uiwait(handles.F);

% Outputs from this function are returned to the command line.
function varargout = Laser_Scan_OutputFcn(hObject, eventdata, ...
    handles)
% varargout  cell array for returning output args;
% hObject    handle to figure
% handles    structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;

function lambda_min_Callback(hObject, eventdata, handles)
% hObject    handle to lambda_min (see GCBO)
% handles    structure with handles and user data (see GUIDATA)

input = str2num(get(hObject, 'String'));

if (isempty(input))
    set(hObject, 'String', '1550.00');
end
guidata(hObject, handles);

% Executes during object creation, after setting all properties.
function lambda_min_CreateFcn(hObject, eventdata, handles)
% hObject    handle to lambda_min (see GCBO)
% handles    empty — handles not created until
% after all CreateFcns called

% See ISPC and COMPUTER.
if ispc && isequal(get(hObject, 'BackgroundColor'), ...
    get(0, 'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

```

```

end

function lambda_step_Callback(hObject, eventdata, handles)
% hObject    handle to lambda_step (see GCBO)
% handles    structure with handles and user data (see GUIDATA)

input = str2num(get(hObject, 'String'));

if (isempty(input))
    set(hObject, 'String', '1.0');
end
guidata(hObject, handles);

% Executes during object creation, after setting all properties.
function lambda_step_CreateFcn(hObject, eventdata, handles)
% hObject    handle to lambda_step (see GCBO)
% handles    empty - handles not created until after all
% CreateFcns called

%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject, 'BackgroundColor'), ...
    get(0, 'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

function lambda_max_Callback(hObject, eventdata, handles)
% hObject    handle to lambda_max (see GCBO)
% handles    structure with handles and user data (see GUIDATA)

input = str2num(get(hObject, 'String'));

if (isempty(input))
    set(hObject, 'String', '1560.00');
end
guidata(hObject, handles);

% Executes during object creation, after setting all properties.
function lambda_max_CreateFcn(hObject, eventdata, handles)
% hObject    handle to lambda_max (see GCBO)
% handles    empty - handles not
% created until after all CreateFcns called

%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject, 'BackgroundColor'), ...
    get(0, 'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

```

```

function N_Callback(hObject, eventdata, handles)
% hObject    handle to N (see GCBO)
% handles    structure with handles
% and user data (see GUIDATA)

input = str2num(get(hObject, 'String'));

if isempty(input)
    set(hObject, 'String', '1');
end
guidata(hObject, handles);

% — Executes during object creation, after setting
% all properties.
function N_CreateFcn(hObject, eventdata, handles)
% hObject    handle to N (see GCBO)
% handles    empty — handles not created
% until after all CreateFcns called

%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject, 'BackgroundColor'), ...
    get(0, 'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

function Filename_Callback(hObject, eventdata, handles)
% hObject    handle to Filename (see GCBO)
% handles    structure with handles and
% user data (see GUIDATA)

input = get(hObject, 'String');

if isempty(input)
    set(hObject, 'String', 'tst');
end
guidata(hObject, handles);

% Executes during object creation, after setting all properties.
function Filename_CreateFcn(hObject, eventdata, handles)
% hObject    handle to Filename (see GCBO)
% handles    empty — handles not created until
% after all CreateFcns called

%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject, 'BackgroundColor'), ...
    get(0, 'defaultUicontrolBackgroundColor'))
    set(hObject, 'BackgroundColor', 'white');
end

```

```

% — Executes on button press in execute_scan.
function execute_scan_Callback(hObject, eventdata, handles)
% hObject    handle to execute_scan (see GCBO)
% handles    structure with handles and user
% data (see GUIDATA)

instrfind;
delete(ans);

lambda_min = get(handles.lambda_min, 'String ');
lambda_step = get(handles.lambda_step, 'String ');
lambda_max = get(handles.lambda_max, 'String ');
N = get(handles.N, 'String ');
Filename = get(handles.Filename, 'String ');

lambda = str2num(lambda_min):str2num(lambda_step):...
        str2num(lambda_max);

laser=serial('COM3');
set(laser, 'BaudRate', 9600);
fopen(laser);
set(laser, 'Terminator', 'CR')
set(laser, 'Timeout', 60)

fprintf(laser, '@*IDN?')
out = fscanf(laser);

osa=serial('COM4');
set(osa, 'BaudRate', 9600);
fopen(osa);
set(osa, 'Terminator', 'CR/LF')
set(osa, 'Timeout', 60)

fprintf(osa, '*IDN?')
out = fscanf(osa);

[x,y] = size(lambda);
wavelength = str2num(lambda_max)*1e-9*ones(1,y);
amplitude = zeros(1,y);

command = [ ':SENSE:AVERAGE:COUNT ' str2num(N) ];
fprintf(osa, command)

for a = 1:y

    command = [ '@WAVE' num2str(lambda(a)) ];
    fprintf(laser, command)

```

```

out = fscanf(laser);
fprintf(laser, '@WAVELENGTH?')
out = fscanf(laser);
fprintf(laser, '@SENS:POW:FRON')
temp1 = fscanf(laser);

command = [ ':SENSE:WAVELENGTH:CENTER ' ...
           num2str(lambda(a) - .94) 'NM'];
fprintf(osa, command)
fprintf(osa, ':SENSE:WAVELENGTH:SPAN 3.0NM')
fprintf(osa, ':INITIATE')

fprintf(osa, ':CALCULATE:MARKER:MAXIMUM')
fprintf(osa, ':CALCULATE:MARKER:X? 0')
temp = fscanf(osa);
wavelength(a) = str2num(temp);

fprintf(osa, ':CALCULATE:MARKER:MAXIMUM')
fprintf(osa, ':CALCULATE:MARKER:Y? 0')
temp = fscanf(osa);
amplitude(a) = str2num(temp)/str2num(temp1);
set(handles.lambda_current, 'String', num2str(lambda(a)));
guidata(hObject, handles);

axes(handles.axes1)
plot(wavelength, amplitude)
drawnow()

end

savefile = [Filename '.mat'];
save(savefile, 'wavelength', 'amplitude')

fclose(laser)
delete(laser)
clear laser;

fclose(osa)
delete(osa)
clear osa;

```

Appendix G

Codes for Calculating Electron Trajectories

The following code is used to calculate electron trajectories from finite-difference time-domain (FDTD) data, as described in Section 6.6.1. The first file, ‘track_trajectories.m’ is responsible for specifying the simulation conditions, loading the FDTD fields, creating new electrons, and time-stepping through the FDTD fields. Most of the underlying details of the calculations are located in the second file, ‘electron.m’, which contains the ‘electron’ class. This class contains the physical constants related to each electron and the update equations used to calculate the electron trajectories. It also records the trajectory of the electrons, their velocity, and the electric field they experience throughout the interaction.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% track_trajectories.m %%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [] = track_trajectories()
clear all
close all
Emax = 3e9;           % Define peak electric field strength (V/m)

tmin = 2;             % time step to start trajectory analysis
tmax = 1500;

d = 10e-9;           % FDTD spatial resolution (m)

% Simulation spatial boundaries to consider
ymin = 0.1e-6;
ymax = 0.6e-6;
zmin = 320e-9;
zmax = 340e-9;
y1 = round(ymin/(inj_n*d));
y2 = round(ymax/(inj_n*d));
z1 = round(zmin/(inj_n*d));
```

```

z2 = round(zmax/(inj_n*d));

% number of electrons to inject
t_inj_max = 800;
% integer multiples of timesteps to inject electrons in
inj_n = 2;
t_up = 0;
% initial velocity to impart electrons with
vi = sqrt(0.5*2*1.602e-19/(0.26*9.11e-31));

if(1)

    % main time loop
    for t = 2:(tmax-tmin+2)
        t

        % load fields of new timestep
        tstep = t+tmin-1;
        load(['../Ey_t_' num2str(tstep) '.mat']);
        load(['../Ez_t_' num2str(tstep) '.mat']);

        % if at appropriate time step, inject electrons
        if ((t+tmin) <= t_inj_max) && mod(t, inj_n) == 0
            t_up = t_up + 1;

            % create pseudo-random starting positions
            y = inj_n*d*(y1:y2);
            z = inj_n*d*(z1:z2);
            [a,b] = size(y);
            y = y+inj_n*d*(0.5-rand(a,b));
            [a,b] = size(z);
            z = z+inj_n*d*(0.5-rand(a,b));

            [y,z] = meshgrid(y,z);
            [a1,b1] = size(y);

            % create a starting velocity
            v_rand = rand(a1,b1);
            vy = vi*ones(a1,b1).*cos(2*pi*v_rand);
            vz = vi*ones(a1,b1).*sin(2*pi*v_rand);

            % create new electrons with the above positions and
            % velocities
            for a = 1:a1
                for b = 1:b1
                    trajectories(t_up,a,b) = electron(y(a,b),...
                    z(a,b),vy(a,b),vz(a,b),t+tmin,tmax,Emax,...
                    Ey,Ez);
                end
            end
        end
    end
end

```



```

% flag to track whether electron has reached
% output facet
ofacet_flag = 0;
E                % electric field vector
H                % magnetic field vector
y                % y coordinate vector
z                % z coordinate vector
ymat
zmat
E_fixed
iy_fixed;
iz_fixed;
t_init;
t_fin;
Emax;            % peak electric field
tmin;            % starting timestep
tmax;            % ending timestep
g;
t;
tm1;
weight;

end

properties (Constant)

% effective mass of electron (kg)
m_e = 0.26*9.10938188*10^(-31);
% time step
dt = 23*19.0657*10^(-18);
% charge of electron (C)
q = 1.60217657*10^(-19);
% plane of the metal cap
z_metal = 340e-9;
% plane of buried oxide
z_oxide = 10e-9;
% waveguide input facet
y_input = 0;
% waveguide output facet
y_output = 5e-6;
% 1/(dy*dz);
c = 1e16;

end

methods

function obj = electron(y0,z0,vy0,vz0,tmin,tmax,...

```

E_{max}, E_y, E_z)

```

% class constructor
    if nargin > 0

        % allocate memory
        length = tmax-tmin+5;
        obj.E = zeros(length,2);
        obj.v = zeros(length,2);
        obj.d = zeros(length,2);
        obj.d0(1,1) = y0;
        obj.d0(1,2) = z0;
        obj.d(1,1) = y0;
        obj.d(1,2) = z0;
        obj.d(2,1) = y0;
        obj.d(2,2) = z0;
        obj.v0(1,1) = vy0;
        obj.v0(1,2) = vz0;
        obj.v(1,1) = vy0;
        obj.v(1,2) = vz0;
        obj.v(2,1) = vy0;
        obj.v(2,2) = vz0;
        load './y.mat'
        load './z.mat'
        z = z';
        obj.y = y;
        obj.z = z;
        obj.Emax = Emax;
        obj.tmin = tmin;
        obj.tmax = tmax;
        obj.t_init = obj.dt*obj.tmin;
        obj.g = obj.dt*obj.q/obj.m_e;
        obj.t = 1;

        % find weighting factor from fields
        [~,index] = min(abs(obj.y-obj.d0(1,1)));
        if obj.y(index) > obj.d0(1,1)
            y1 = index - 1;
            y2 = index;
        elseif obj.y(index) < obj.d0(1,1)
            y1 = index;
            y2 = index + 1;
        end

        [~,index] = min(abs(obj.z-obj.d0(1,2)));

        if obj.z(index) > obj.d0(1,2)
            z1 = index - 1;

```

```

        z2 = index;
    elseif obj.z(index) < obj.d0(1,2)
        z1 = index;
        z2 = index + 1;
    end

    % populate field vectors at current position
    % using bilinear interpolation

    m1 = (obj.y(y2)-obj.d0(1,1))*(obj.z(z2)...
    -obj.d0(1,2));
    m2 = (obj.d0(1,1)-obj.y(y1))*(obj.z(z2)...
    -obj.d0(1,2));
    m3 = (obj.y(y2)-obj.d0(1,1))*(obj.d0(1,2)...
    -obj.z(z1));
    m4 = (obj.d0(1,1)-obj.y(y1))*(obj.d0(1,2)...
    -obj.z(z1));

    Wy = obj.Emax*obj.c*(Ey(y1,z1)*m1...
    + Ey(y2,z1)*m2 + Ey(y1,z2)*m3 ...
    + Ey(y2,z2)*m4);

    Wz = obj.Emax*obj.c*(Ez(y1,z1)*m1...
    + Ez(y2,z1)*m2 + Ez(y1,z2)*m3 ...
    + Ez(y2,z2)*m4);

    obj.weight = (Wy^2 + Wz^2)^2;

end

end

function obj = step(obj,Ey,Ez)

    obj.tm1 = obj.t;
    obj.t = obj.t + 1;

    % check if electron has collided with gold surface
    if obj.d(obj.tm1,2) >= obj.z_metal
        obj.gold_flag = 1;
        obj.t_fin = obj.t_init+obj.dt*(obj.tm1);
        return
    % check if electron has collided with silica surface
    elseif obj.d(obj.tm1,2) <= obj.z_oxide
        oxide_flag = 1;
        obj.t_fin = obj.t_init+obj.dt*(obj.tm1);
        return
    elseif obj.d(obj.tm1,1) <= obj.y_input

```

```

        ifacet_flag = 1;
        obj.t_fin = obj.t_init+obj.dt*(obj.tm1);
        return
    elseif obj.d(obj.tm1,1) >= obj.y_output
        ofacet_flag = 1;
        obj.t_fin = obj.t_init+obj.dt*(obj.tm1);
        return
    end

% find closest mesh points
[~,index] = min(abs(obj.y-obj.d(obj.tm1,1)));
if obj.y(index) > obj.d(obj.tm1,1)
    y1 = index - 1;
    y2 = index;
elseif obj.y(index) < obj.d(obj.tm1,1)
    y1 = index;
    y2 = index + 1;
end

[~,index] = min(abs(obj.z-obj.d(obj.tm1,2)));

if obj.z(index) > obj.d(obj.tm1,2)
    z1 = index - 1;
    z2 = index;
elseif obj.z(index) < obj.d(obj.tm1,2)
    z1 = index;
    z2 = index + 1;
end

% populate field vectors at current position using
% bilinear interpolation
m1 = (obj.y(y2)-obj.d(obj.tm1,1))*(obj.z(z2)...
-obj.d(obj.tm1,2));
m2 = (obj.d(obj.tm1,1)-obj.y(y1))*(obj.z(z2)...
-obj.d(obj.tm1,2));
m3 = (obj.y(y2)-obj.d(obj.tm1,1))...
*(obj.d(obj.tm1,2)-obj.z(z1));
m4 = (obj.d(obj.tm1,1)-obj.y(y1))...
*(obj.d(obj.tm1,2)-obj.z(z1));

obj.E(obj.t,1) = obj.Emax*obj.c*(Ey(y1,z1)*m1...
+ Ey(y2,z1)*m2 + Ey(y1,z2)*m3 + Ey(y2,z2)*m4);

obj.E(obj.t,2) = obj.Emax*obj.c*(Ez(y1,z1)*m1...
+ Ez(y2,z1)*m2 + Ez(y1,z2)*m3 + Ez(y2,z2)*m4);

% update velocity
obj.v(obj.t,1) = obj.v(obj.tm1,1)...

```

```

+ obj.g*obj.E(obj.t,1);
obj.v(obj.t,2) = obj.v(obj.tm1,2)...
+ obj.g*obj.E(obj.t,2);

% update position
obj.d(obj.t,1) = obj.d(obj.tm1,1)...
+ obj.v(obj.t,1)*obj.dt;
obj.d(obj.t,2) = obj.d(obj.tm1,2)...
+ obj.v(obj.t,2)*obj.dt;

end

% check if electron has collided with any of the
% waveguide interfaces
function flag = check(obj)

    if obj.gold_flag == 1 || obj.oxide_flag == 1 || ...
        obj.ifacet_flag == 1 || obj.ofacet_flag == 1
        flag = 0;
    else
        flag = 1;
    end

end

end

end

end

```