UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Single-Mode Waveguide Coupling of Optical Antenna-LED for Optical Interconnects

Permalink

https://escholarship.org/uc/item/2244w3xq

Author

Andrade, Nicolas Miguel

Publication Date

2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Single-Mode Waveguide Coupling of Optical Antenna-LED for Optical Interconnects

by

Nicolas Miguel Andrade

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ming C. Wu, Chair
Professor Eli Yablonovitch
Professor Feng Wang

Spring 2021

Single-Mode Waveguide Coupling of Optical Antenna-LED for Optical Interconnects

Copyright 2021
by
Nicolas Miguel Andrade

Abstract

Single-Mode Waveguide Coupling of Optical Antenna-LED for Optical Interconnects

by

Nicolas Miguel Andrade

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ming C. Wu, Chair

Optical interconnects have displaced electrical wires in data centers and high performance computers due to higher efficiency and bandwidth. The development of on-chip optical interconnects is the next frontier for development, with the promise of reducing on-chip energy consumption. Light emitting diodes have high efficiencies and can achieve low footprint; however, spontaneous emission is typically far too slow to be useful in an interconnect. This can be understood by modeling spontaneous emission as dipole emission with a dipole length on the order of the atomic spacing ($\approx 1.3nm$) which is much smaller than the radiation wavelength at $1550nm$. This becomes a familiar radio frequency (RF) engineering problem: to make the light-emitting diode faster we just need to make it a better antenna.

By placing the radiating dipoles in the feed-gap of an optical antenna we can dramatically enhance the spontaneous emission rate, potentially allowing for >100GHz direction modulation. However, to create a useful on-chip interconnect, we need to couple the output light to a single mode waveguide so we can guide the light across the chip to a photodetector. Additionally, both the device and waveguide coupling scheme need to be compatible with top-down fabrication and electrical injection.

In this dissertation we focus on the cavity backed slot antenna geometry, showing that this device can be efficiently coupled to a single mode waveguide. The first part of this dissertation reviews the theory behind spontaneous emission enhancement, the cavity backed slot antenna, and waveguide coupling.

The waveguide coupled cavity backed slot antenna is then fabricated and measured. Clear evidence of waveguide coupling is demonstrated. The collected data is then compared to theory showing excellent agreement between experiment and finite difference time domain (FDTD) simulations. The estimated experimental waveguide coupling efficiency to a single mode waveguide is $\approx 85.9\%$.

The efficiency of the device is improved through a novel surface passivation process, leading

to a record low surface recombination velocity. Detailed time decay models are presented, including a method to model an arbitrary rate equation. The internal quantum efficiency of a 60nm wide LED ridge is estimated to be $10-20\%$ - with these devices showing $>180\times$ increase in photoluminescence.

Finally, the dissertation concludes with a full-link system model and a discussion of how to increase the power from the cavity backed slot antenna-LED and the importance of close integration with CMOS. The modeled end-to-end energy per bit can be $<1fJ/bit$, showing the great potential for on-chip optical interconnects.

To Rachel, Bingley, and Bea

Contents

Contents **ii**

List of Figures **iv**

List of Tables **x**

1 Introduction **1**
1.1 The metal interconnect problem 2
1.2 Fundamental limit of optical interconnects 3
1.3 On-chip optical interconnect transmitter requirements 4
1.4 Outline of this dissertation 7

2 Spontaneous emission enhancement **8**
2.1 Spontaneous emission 8
2.2 Spontaneous emission rate enhancement 9
2.3 Figures of merit 13

3 Cavity backed slot antenna-LED **19**
3.1 Cavity backed slot antenna 19
3.2 Aluminum oxide 24
3.3 Averaging 26

4 Waveguide coupled cavity backed slot antenna design **30**
4.1 Starting structure 30
4.2 Destructive interference 31
4.3 Total internal reflection 31
4.4 Flat facet reflector 33
4.5 Tapered waveguide coupler 34
4.6 Inverse design 38
4.7 Coupled cavity 41
4.8 Secondary design considerations 42
4.9 Towards electrical injection 46

5 Optically pumped waveguide coupled cavity backed slot antenna LED 50

5.1 Fabrication 50
5.2 Silver sputtering or evaporation 54
5.3 Measurement setup 57
5.4 Results 60

6 Surface Passivisation 64

6.1 Surface recombination 65
6.2 Time correlated single photon counting 66
6.3 Fabrication 68
6.4 Results 69
6.5 Thinner oxide 74
6.6 Field effect passivation 76

7 Link Model 83

7.1 Transmitter macro model 83
7.2 Receiver macro model 85
7.3 End-to-end link performance 86
7.4 Increasing power 87
7.5 Integration and outlook 89

8 Conclusion 91

A Fabrication Notes 92

A.1 Alignment Marks 93
A.2 Nanoscale Vernier Marks 93
A.3 Adhesion 93
A.4 High Contrast EBL 94
A.5 Silver agglomeration 95
A.6 Substrate removal 97

B Time Correlated Single Photon Counting 98

B.1 Theory 98
B.2 Avalanche photodiode 99
B.3 Setup details 101
B.4 Auto-alignment 102

C Inverse Design 104

Bibliography 106

List of Figures

1.1 Evolution of optical interconnects from global length scales to on-chip optical interconnects [3]. . . . 1
1.2 Capacitance per unit length versus the design rule (wire pitch). Reproduced from [8] ©1983, IEEE 3
1.3 Cartoon depiction of (a) stimulated emission in two level system with mirrors on both ends to create an optical cavity - note stimulated light is coherent so frequency and phase of photons match (b) stimulated emission in semiconductor composed of many dipoles (c) light intensity (i.e. number of photons per second) as a function of current (i.e. number of electrons per second) (d) modulation bandwidth as a function of current 5
1.4 Cartoon depiction of (a) spontaneous emission in two level system (b) spontaneous emission in semiconductor composed of many dipoles - note random orientation and phase (c) light intensity (i.e. number of photons per second) as a function of current (i.e. number of electrons per second) (d) modulation bandwidth as a function of current 6

2.1 Modulation bandwidth as a function of mode volume and Q factor. Q_{opt} is plotted as a dashed line indicating the optimal Q factor for modulation bandwidth, the area above this line is in the strong coupling regime which would require different analysis. Reproduced from [17] ©2009 Optical Society of America 11
2.2 (a) Electric dipole in free space with dipole length x_0 and (b) its equivalent circuit model. (c) Electric dipole in feed gap of dipole antenna and (d) its equivalent circuit model. 12
2.3 Dashed black and solid red lines show the experimental non-enhanced material spectrum $[L(\omega)]$ and the simulated enhancement spectrum $[F(\omega)]$ of the cavity-backed slot antenna on a bulk InP substrate, respectively. 16

3.1 Comparison of the dipole antenna, slot antenna, and cavity backed slot antenna geometry (a) - (c), farfield radiation pattern ($|E|$) (d)-(f), and circuit model with radiation resistance (g)-(i). 20
3.2 Cartoon showing electrical injection for cavity backed slot antenna-LED. The metal serves as both the antenna and the n-contact to inject electrons. The holes can be injected through the p-InP substrate. Note this is not drawn to scale. . . . 21

3.3 Contour plots of the (a) resonance frequency ω_0, (b) overall emission enhancement, (c) antenna efficiency on resonance, and (d) quality factor of the III–V antenna-LED with 1nm Al_2O_3 for a 20nm wide device. The solid white contour line denotes the combinations of slot length and cavity depth that achieve antenna resonance at 200 THz. 22
3.4 (a) Spectral average $\mu_{spectral}$ as a function of Q and (b) average enhancement F_{avg} and radiated power $\propto F_{avg} \times \eta_{antenna} \times V_{active}$ for slots with 200THz resonance. . 23
3.5 Simulation of spontaneous emission enhancement dependence on Al_2O_3 thickness. Antenna length was not adjusted for increasing thickness leading to a blue-shift of antenna resonance. 24
3.6 (a) Experimental decay curves for 200nm ridge device lifetime dependence based on Al_2O_3 thicknesses between 0-5nm. (b) Integrated decay intensity for variable ridge widths. 25
3.7 Normalized $|E|^2$ field with respect to $(x = 0, y = 0, z = 10nm)$ excited from electric dipole 500nm below slot opening (a) XZ profile (b) XY profile and (c) YZ profile. 28

4.1 Cross section, power flow, and waveguide coupling efficiency to the fundamental mode (η_{WC}) for (a) antenna-LED on single-mode InP waveguide and SiO_2 ridge, (b) antenna-LED on single-mode InP waveguide with metal wrapped around waveguide facet, and (c) antenna-LED on single-mode InP tapered waveguide with metal wrapped around waveguide facet and sidewalls (see Fig. 4.6(a) for perspective view, Fig. 4.8(b) for top view cross section). See *Appendix: Field profiles* for the E_{x} and E_{y} field profiles of the mode in the InP waveguide. . . . 32
4.2 Waveguide coupling efficiency (accounting for both directions) at resonance as for a) lower cladding refractive index and b) ridge height. The waveguide coupling efficiency reported is for bi-directional coupling, half of the power travels in each direction. Graph insets show width-height cross section of waveguide. 33
4.3 Waveguide coupling efficiency at 1500nm as a function of the separation between the center of the antenna-LED and the metal reflector. 34
4.4 Available modes in the metal coupler region. a) Waveguide mode that couples well to InP waveguide and b) waveguide mode that does not couple well to InP waveguide 35
4.5 Mode in waveguide when metal is wrapped around the waveguide sidewalls. This mode is the only mode that exists at this wavelength. 36
4.6 (a) Perspective view of tapered waveguide coupler with a waveguide height of 180nm and width of 550nm on a 500nm tall SiO_2 ridge, and (b) enhancement, antenna efficiency, and waveguide coupling efficiency spectra. 36
4.7 E_x and E_y field profiles for (a) antenna-LED on single-mode InP waveguide and SiO_2 ridge, (b) antenna-LED on single-mode InP waveguide with metal wrapped around waveguide facet, and (c) antenna-LED on single-mode InP tapered waveguide with metal wrapped around waveguide facet and sidewalls. 37

4.8 (a) Cross section schematic (XZ) of tapered waveguide coupler showing dashed cutline, and (b) top view XY cross section of waveguide along dashed cutline. (c) XY cross section of coupler after optimization, showing perturbations to Ag-InP boundary. Note (b) and (c) also show the projection of the LED base. 38
4.9 Enhancement, antenna efficiency, waveguide coupling efficiency spectra and top view XY cross sections for (a) single frequency optimization and (b) multi frequency optimization. For reference, the LED material spectrum [$L(\omega)$] between its 50% power points is shown by the gray shaded region. 39
4.10 (a) Avoided crossing between the optical antenna resonance and the inverse design coupler resonance. For reference, dashed black and green lines show independent resonances of the antenna-LED on a bulk InP substrate and the coupler section as a function of LED length, respectively. Enhancement spectra for LED lengths of (b) 110nm and (c) 122nm. 41
4.11 Waveguide coupling efficiency, enhancement, and antenna efficiency as a function of HSQ thickness. 43
4.12 Comparison of waveguide coupling efficiency for (a) abrupt taper coupler and (b) continual taper coupler with top down cross section. 45
4.13 (a) Top down geometry of electrically injected waveguide, and optical transmission efficiency in taper section for different taper geometries for (b) 1000nm wide waveguide and (c) 1500nm wide waveguide. 47

5.1 Waveguide coupled antenna-LED epitaxial layers 51
5.2 Abbreviated fabrication flow for LED ridge etch, antenna deposition, and substrate removal. 52
5.3 Left: scanning electron micrograph of tapered coupler after metal deposition. Right: focused ion beam cross section micrograph showing conformal coverage of silver. 53
5.4 Fabrication flow 55
5.5 FDTD simulation using total-field-scattered-field source. Silver loss modeled by increasing imaginary part of the refractive index from [59] for InP ridge (length: 680nm width: 60nm). 56
5.6 Comparison of sputtered silver (solid lines) and evaporated silver (dashed lines) for various antenna lengths for an antenna width of 60nm. Both samples show similar resonances for same antenna lengths. 56
5.7 LOR and PMMA bilayer showing re-entrant profile after exposure, 60s PMMA develop using MIBK:IPA, and 30s LOR develop using OPD4262. 57
5.8 Schematic showing expected areas where photons are emitted from. “Not-coupled” is all the light that is immediately emitted from the coupler section; “scattered” is the light that scatters when going from the coupler section to single mode waveguide; and “grating” is the light that is directed up and down through the grating. 58

5.9 Schematic of waveguide coupling efficiency measurement. Device under test (DUT) is optically pumped with laser source at higher photon energy. Emitted light at lower photon energy is filtered with dichroic mirror and directed to 2D InGaAs camera or cooled 1D InGaAs array. 59
5.10 $A(\omega)$ factor plotted as a function of wavelength. $A(\omega)$ takes into account the grating efficiency and collection efficiencies as well as the waveguide coupling and not-coupled count efficiencies. 60
5.11 Top: Optical microscope image of fabricated device after substrate removal. Bottom: Dark field image overlaid with heatmap of emission and solid blue line showing counts along 1D pixel line. A crack near the sample was label, it does not propagate into the device or waveguide. 61
5.12 Air side measurement of light emission for waveguide coupled cavity backed slot antenna LED. Not-coupled (blue) counts refer to light coming from the device that is not coupled to the waveguide, and grating counts (orange) refer to the light that was coupled to the waveguide and is emitted from the grating. 62
5.13 Spectra for grating counts (orange) and not-coupled counts (blue), as well as the sum of the two curves (green). Note, there is an optical filter at 1300nm to remove laser light. Inset shows GDS file for the structure, the length explains the Fabry-Perot etalon. 63
5.14 Waveguide coupling ratio as a function of wavelength. Calculated from diving the grating spectrum with the total spectrum from Fig. 5.13. The simulated waveguide coupling ratio is also plotted in orange. 63

6.1 Electrically active bonds and impurities at the surface of crystal lattice can lead to high recombination. As we scale the device to the nanoscale the surface to volume ratio becomes larger the surface recombination becomes more significant. 64
6.2 (a) Schematic of passivated InGaAsP LED ridge (length:1000nm, height:210nm, width:variable). (b) SEM of 200nm wide LED ridge after surface passivation ($S + (NH_4)_2S$ process) 69
6.3 Detected photoluminescence vs laser pump power (L-L) curves for LED widths of (a) 40nm (b) 200nm (c) 400nm and (d) 700nm. Samples were measured immediately after etching (blue) and again after surface passivation (orange) - these results are for the $S + (NH_4)_2S$ passivation 71
6.4 (a) L-L curve for continuous-wave μ-photoluminescence measurements at room-temperature displaying increase in photoluminescence and (b) spectra for $4\mu W$ pump power. L-L curve and spectra from 200nm wide ridge with length 1000nm. 72
6.5 Decay curve for 200nm wide ridge measured after etch (blue), for the $S+(NH_4)_2S$ passivation (orange), and for the $Old(NH_4)_2S$ passivation (green). Dashed lines show decay curve fit from $A'B$ decay model. 73

6.6 Surface recombination velocity as a function of width. At narrow ridge widths the surface recombination velocity increases exponentially (plotted as a dashed line). At large widths the surface recombination is approximately constant (mean plotted as a solid black line with ± a standard deviation). 74
6.7 Comparison of decay for $Al_2O_3/TiO_2/Al_2O_3$ trilayer chip (3nm/4nm/3nm) and bilayer Al_2O_3/TiO_2 (3nm/4nm) after annealing. Bilayer has significantly shorter lifetime. 75
6.8 Decay for InGaAs active region test chip for different Al_2O_3 and TiO_2 thickness (a) after annealing, (b) after 30:1 BHF removal of outer Al_2O_3, and (c) after antenna deposition and substrate removal for 1.5nm/3nm Al_2O_3/TiO_2 bilayer - the antenna deposition sample was pumped at 10x less power. 77
6.9 (a) h density as a function of position for different widths assuming $Q_{ox} = -3E13$, (b) effective electrostatic doping level, P_{ES}, for $Q_{ox} = -3 \times 10^{13} cm^{-2}$ and $Q_{ox} = -2 \times 10^{12} cm^{-2}$, and (c) $1/\tau$ as a function of 1/width - dashed line is for A' fit only and solid black line includes average between electrostatic doping terms from (b) - gray bars represent $|Q_{ox}| = 2 \times 10^{12} - 3 \times 10^{13} cm^{-2}$ range 78
6.10 (a) Expected η_{IQE} assuming we have electrostatic doping (solid) and assuming decreased lifetime is from higher non-radiative emission (dashed line). Experimental η_{IQE} using adjusted counts and powers from Table 6.1 for after etch and after passivation in $S + (NH_4)_2S$ sample assuming perfectly aligned Gaussian beam (b) and with a 15° polarization offset (c). 80
6.11 Projected η_{IQE} as a function of enhancement for a 20nm width for (a) undoped, unpassivated device, and (b) passivated device with $P_{ES} = 2 \times 10^{18} cm^{-3}$. 81

7.1 Full Link Image 83
7.2 Optimal energy per bit breakdown, required LED power, and receiver topology vs data rate, using the parameters in Table 7.2 86
7.3 Energy per bit at 50Gbps with 260GHz f_T for additional integration capacitances. 1.5fF roughly corresponds to the self capacitance in a through oxide via. 90

A.1 GDS file layout of waveguide coupling design highlight global, local, and Vernier alignment marks. 92
A.2 SEM of vernier marks showing no y-offset and an x-offset of +20nm for the waveguide layer relative to the LED ridges. This vernier pair was on the right edge of the 600μm field. 94
A.3 Atomic force micrograph (AFM) taken after FOX-15 exposure in Crestec 130keV and development for 20s in OPD 4262 (dilute TMAH). Profile in top right clearly shows evidence of forward scattering near waveguide edges causing resist to be under-developed in that region. 95

A.4 Comparison of waveguide coupled LEDs (a) after liftoff, (b) after atomic layer deposition capping with Al_2O_3 at 200C showing agglomeration, and (c) sample capped with 70nm Nb and heated on hotplate to 200C for more than an hour showing no obvious signs of agglomeration. 96
A.5 Optical micrograph after substrate and etch stop removal for a device with no additional Al_2O_3 deposited showing InP waveguide was unintentionally etched. . 97

B.1 (a) Laser pulse at fixed repetition rate. (b) Three different photon arrival times, binned at t_1, t_4, and t_2. (c) Histogram built up from binned photon arrival times. (d) Basic implementation of TCSPC. Drawing inspired by [28]. 99
B.2 Block diagram of the TCSPC optical setup at near-IR wavelengths. Drawing inspired by [28]. 100
B.3 Biasing trade-off - left: Photon detection efficiency as a function of excess bias and right: dark count rate as a function of hold-off time with three excess biases plotted. Reproduced from [99]. 101
B.4 Background subtraction for MPD InGaAs SPAD with $V = 2.5V$ and $T_{HO} = 15\mu s$ with laser rep rate of 4MHz. The power dependent dark count rate was determined by measuring a short lifetime device which decayed to the noise floor quickly. The dark count average was taken between 24-28ns with a 4ps bin (1000 samples). On the right we plot the dark count average as a function of the total counts minus the thermal dark counts (no photons). 102
B.5 Top: three background-subtracted decay curves for three separate pump powers. Bottom: optimal time delay calculated from least squares analysis. 103

List of Tables

2.1 Quantum well matrix element with quantized dimension oriented along $\hat{z}$ [12]. . . 15

3.1 Summary of approximate dipole averages for single quantum well, multiple quantum wells, and double heterostructure active regions. The polarization average is based on the matrix element, the spatial average is calculated from Lorentz reciprocity and the analytical formula, and the spectral average is an approximation based on the Q of the device (this is assumed not to change). 29

4.1 Summary of best design results for antenna on bulk InP, taper coupler, and inverse design structures designed at a single frequency and multiple frequencies, respectively. These structures are for a width of 20nm and an Al_2O_3 thickness of 1nm. 40

4.2 Calculation of series resistance for p-doped InP waveguide sections assuming a doping of $N_a = 1 \times 10^{19} cm^{-3}$ and waveguide height of 200nm. 49

6.1 Summary of efficiencies used to calculate the internal quantum efficiency η_{IQE}. η_{in} and η_{in} ($\theta_{gausian} = 15°$) are the input efficiencies assuming a Gaussian beam profile perfectly aligned to the ridge length and with a $\theta_{gausian} = 15°$ polarization offset. η_{out} is the collection efficiency, and η_{optics} is the product of optical and detector efficiencies. 79

7.1 The projected average enhancement F_{avg}, active region volume V, f_{3dB}, and output power for different active region designs biased to a carrier concentration of $N = 3 \times 10^{19} cm^{-3}$ for a 20nm ridge width. 89

Acknowledgments

I am very grateful for all the people who were critical in the development of this research - either directly or indirectly - over the years.

First, I would like to thank my graduate advisor, Professor Wu, for all the support during my time here. I really appreciate the collaborative group atmosphere you created. That support network was a huge help in the development of the research.

I would also like to thank Professor Yablonovitch, your questions have helped me gain a deeper insight. Thank you to both Professor Yablonovitch and Professor Wang for serving on my dissertation committee. And, Professor Kante for serving on my qual committee. I would also like to acknowledge my undergraduate research mentor, Professor Ümit Özgür, who initially sparked my interest in this field and guided me throughout my undergraduate career.

I really appreciated the weekly meetings with the antenna-LED subgroup: Dr. Seth Fortuna, Dr. Kevin Han, and Sean Hooten. Seth's mentorship in guiding my initial research in both simulation and fabrication was truly invaluable. Sean, I really appreciated going over research ideas with you, that process helped refine a lot of the ideas presented in this dissertation.

I will miss the collaborative nature of the 253M office - the discussions with graduate students in Professor Chang-Hasnain's group, Professor Wu's group, Professor Yablonvitch's group, and Professor Kante's group were incredibly helpful for troubleshooting fabrication issues, learning more about other projects, or cathartic after a long week. Those discussions were sorely missed during COVID.

Thank you to the staff of the Marvell Nanolab for running such a great facility and your expertise for process development. Additionally, the administrative staff at the E3S center, BSAC, and department were a tremendous support.

Finally, I would like to thank my family: my parents for all their help along the way, and my brother, Luis, for generously offering to copy edit. Thank you to my finacée Rachel, for bringing me so much happiness even during some of the most stressful times. You are my light, and I look forward to the journey ahead.

Chapter 1

Introduction

Information and communication technology (ICT) energy usage is expected to reach nearly 20% of global energy consumption by 2030, in large part due to increased demand from data centers [1]. This is not surprising because it is estimated that in just a few years there will be 30 billion devices connected to the internet exchanging zettabytes (10^{21} bytes) of data per year - mostly driven by an increase in mobile traffic [2]. Unfortunately, while the number of devices and computations are growing rapidly, scaling of transistors and energy per bit reduction has shown signs of slowing, and energy per bit reduction for on-chip interconnects is nearly stagnate.

	MAN/WAN	Cables (long)	Cables (short)	Card-to-card	Intra-card	Intra-module	Intra-chip
Length	Multi-km	10-300m	1-10m	0.3-1m	0.1-0.3m	5-100mm	0-20mm
Number of lines per system	Tens	Tens to thousands	Tens to thousands	Tens to thousands	Thousands	Tens of thousands	Millions
Use of optics	1980s	1990s	2005	2005-2010	2010-2015	After 2015	To be determined

Figure 1.1: Evolution of optical interconnects from global length scales to on-chip optical interconnects [3].

Regardless of scale, most of the energy cost in computing is consumed in moving a bit of information from point A to B. Due to the high traffic in data centers, interconnects need to operate at high bandwidths and low energy consumption. However, metal interconnects are severely limited in both bandwidth and energy consumption, which is why metal has been replaced by optical interconnects when available for distances longer than a meter. Fig. 1.1 lays out the historical transition from metal interconnects to optical. The main factor stopping a complete optical takeover on the data center scale (short cable) is cost. As the

market continues to mature it is expected that this scaling will continue with intra-module (chip-to-chip) [4] and intra-chip (on-chip) being the next frontiers.

To quantify the problem for on-chip interconnects, in 2004 interconnects consumed over 50% of the power in high-performance microprocessors at the 130nm technology node [5], and likely over 80% of power in modern CPUs [6]. In addition to high energy consumption of the interconnects, it is estimated that 50-80% of the gates are used for repeaters to maintain high switching speed [7].

In order to achieve greater on-chip power savings we need to address the interconnect problem.

1.1 The metal interconnect problem

To send a bit of information on-chip we need to charge the capacitance of a metal line, and to flip the bit back we need to discharge this metal line back to ground. That power used to charge the metal line will be dissipated as thermal loss over the wire resistances. So, on average over this cycle we would dissipate half the energy stored in the wire, which is given by:

$$E = \frac{1}{2} \times c_l \times l \times V^2 \tag{1.1}$$

where E is the energy per bit, c_l is the capacitance per unit length of the wire, l is the length of the wire, and V is the voltage we charge the line to (the required transistor gate voltage). Typical numbers for c_l and V are 2pF/cm and 0.8V, leading to an energy per bit per length consumption of:

$$\frac{E}{l} \approx 640\ fJ/bit/cm \tag{1.2}$$

Surely we must be able to scale this down by decreasing capacitance or voltage! Unfortunately, this is where the problem arises. As shown in Fig. 1.2, the total wire capacitance is a combination of capacitance to ground and to adjacent wires. These capacitances trade-off as we scale down the wire leading to a minimum capacitance of ≈2pF/cm [8, 9].

Likewise, it is difficult to scale down the transistor V because we want to maintain low leakage current in the off-state and a high on-off ratio. This is due to the ≈60mV/decade subthreshold swing for conventional transistors. Therefore, to ensure circuit stability, fast operation, and low leakage the supply voltage has been stabilized to be around 0.5-2V [10]. Conventional transistors cannot be operated below 0.2V [11], so there is very limited scaling of the voltage using this platform.

There has been a lot of promising work with new transistor technology to improve the subthreshold swing, which could lead to reductions in V. It is expected that a tunnel-FET could be scaled to 0.4V, ferro-FET to 0.2V, and spintronic MESO to 0.1V, leading to a ≈10x-100x scaling of the interconnect energy. While this is a promising avenue for interconnect energy scaling it is outside the scope of this dissertation.

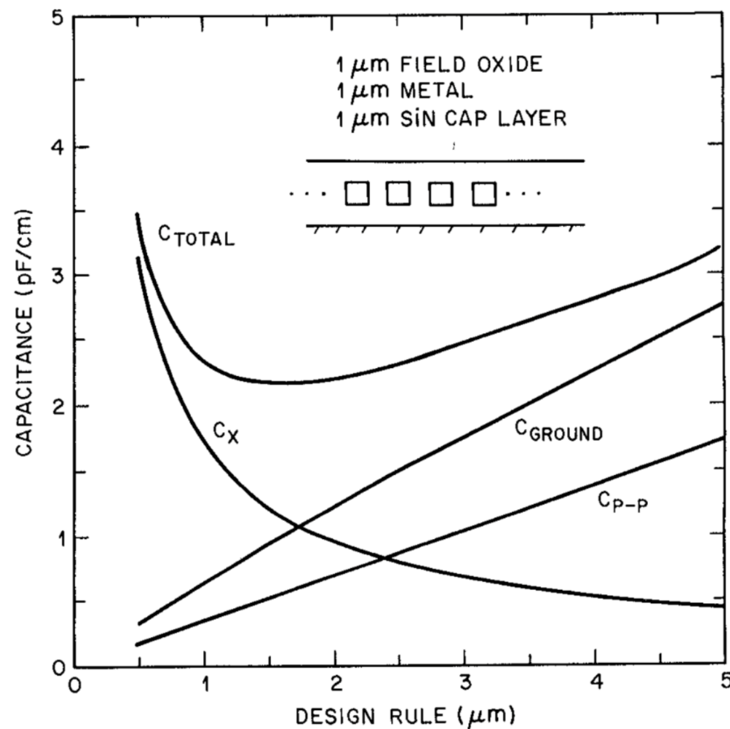

Figure 1.2: Capacitance per unit length versus the design rule (wire pitch). Reproduced from [8] ©1983, IEEE

Another important limit to consider is bandwidth. Due to wiring resistance and capacitance, electrical interconnects are limited in their bandwidth to [9]:

$$B \approx B_o \frac{A}{l^2} \tag{1.3}$$

where $B_o \approx 10^{15} - 10^{16} bits/s$, A and l are the wire cross-sectional area and length, respectively. Note that this limit applies to "on-off" signaling. For long metal lines, repeaters can be inserted to decrease l, but this comes at the expense of additional gates and increased power consumption.

1.2 Fundamental limit of optical interconnects

For optical interconnects we do not need to worry about charging the capacitance of a long metal line, because the information will be carried via an optical waveguide. For a very simplified calculation of the lower limit of optical interconnects we can assume we have a perfect transmitter, waveguide, and single photon detector. Additionally, in this simple

model we are doing intensity modulation, so we will get a "1" in the presence of one or more photons and a "0" if there are zero photons. So, the bit error rate (BER) can be found with Poisson statistics, with μ expected emitted photons per "1" and N actual received photons.

$$P(\mu; N) = \frac{\mu^N e^{-\mu}}{N!} \tag{1.4}$$

An error occurs when there are μ expected photons, but $N = 0$ received photons:

$$\begin{aligned} BER &= P(\mu; N = 0) \\ BER &= 10^{-22} \geq e^{-\mu} \end{aligned} \tag{1.5}$$

Therefore to get a BER of 10^{-22}, we need $\mu \geq 50$ photons. But if we want to send a "0" we do not need any photons, so on average we only need 25 photons per bit. Let's also assume we are operating around telecom wavelengths (1550nm), so the energy per photon is 0.8eV (0.13aJ).

$$E = 25\frac{photons}{bit} \times 0.13\frac{aJ}{photon} = 3.2\ aJ/bit \tag{1.6}$$

In this idealized link we would only need 3.2 aJ/bit! Even with introducing realistic waveguide losses <1 dB/cm the link, energy per bit would be nearly constant at on-chip length scales (<20mm). In chapter 7 we will introduce a more realistic optical link model, where we replace the zero-power transmitter and receivers with real models to account for the total system link energy consumption.

The maximum speed of "on-off" signalling in optical interconnects will be limited by the transmitter speed - both electrical and optical.

1.3 On-chip optical interconnect transmitter requirements

Optical interconnects would clearly have significant benefits over metal interconnects; however in order to create an efficient on-chip optical interconnect we should review the requirements:

1. **Integration with electronic IC.** The devices need to be electrically injected. This requires repeatable and deterministic placement of the device. Essentially, the fabrication needs to be top-down and excludes the use of randomly dispersed structures.

2. **Compatible with single mode waveguide.** The emission from the device needs to be coupled to a single mode waveguide. Single mode operation is important to achieve higher waveguide confinement for the photodetector, where we need ultra-small capacitance (aF - fF) for low-energy and high speed operation. Multi-mode structures are intrinsically larger, making that more challenging.

3. **Fast direct modulation.** The speed of the interconnect should be faster than on-chip rates. The clock speed is on the order of 1-6GHz. The speed has been kept at this rate due to the required voltage for higher speed operation and to reduce heating associated with higher power. Unconstrained, the transistor would be able to switch on the order of 3ps [9]. So the transmitter should be capable of a modulation speed $\geq$ 50GHz.

4. **High efficiency.** This requirement is fairly obvious: in order to create an efficient interconnect we cannot tolerate much loss in the system, including the transmitter. In practice, this can be difficult to achieve.

5. **Nanoscale size.** This requirement is related to the previous requirements. Due to the high density of interconnects we would need to have a small device footprint for integration. Generally, nanoscale devices can be operated at higher speeds and at lower powers. Finally, the nanoscale size means the mode is well defined - allowing for efficient coupling to a single mode waveguide.

Case for lasers

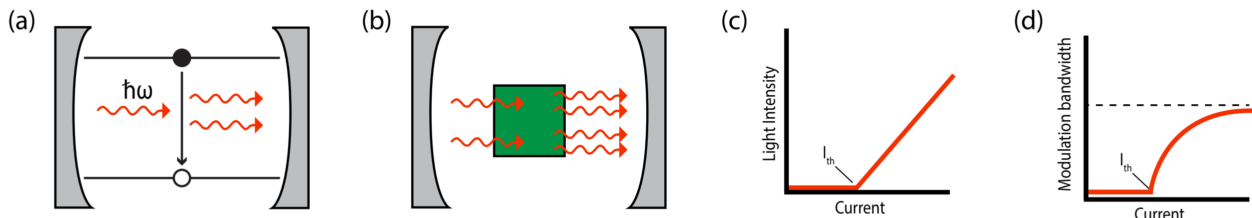

Figure 1.3: Cartoon depiction of (a) stimulated emission in two level system with mirrors on both ends to create an optical cavity - note stimulated light is coherent so frequency and phase of photons match (b) stimulated emission in semiconductor composed of many dipoles (c) light intensity (i.e. number of photons per second) as a function of current (i.e. number of electrons per second) (d) modulation bandwidth as a function of current

Lasers have been the preferred solution for creating optical interconnects due to their relatively high speed, power, single mode operation, and coherence - allowing for more complicated modulation formats. The general operating principle is shown in Fig. 1.3, and it requires a gain medium and an optical cavity. The device reaches threshold when the gain medium is injected with sufficient electron-hole pairs to achieve population inversion, and the cavity has enough photons to stimulate additional recombination. The speed of the laser can be increased by injecting more electron-hole pairs, which creates additional photons. We can keep ramping up the current until the gain saturates, which is typically around 50GHz for most lasers [12].

Looking back at our list of requirements, the laser can easily satisfy the integrated and coupled to a single mode waveguide condition. In terms of the speed, 50GHz would be sufficient for most applications, but does not leave much room for scaling. On the other hand, a recent demonstration from NTT used a photon-photon resonance to extend the f_{3dB} to 108GHz [13]. Generally, lasers have an issue with two requirements: high efficiency and nanoscale size.

As mentioned earlier, in order to achieve high speed operation the device needs to be biased at several times the threshold current. This comes with a relatively high static energy consumption, especially for large cavities. Scaling lasers to a nanoscale regime is challenging because dielectric cavities can only be scaled to $\approx (\lambda_0/2n)^3$, and requires mirrors to build up a sufficient photon density for high speed operation. Metal has been used to shrink down mode volumes and confine photons, but metal cavities generally have high loss which will further increases threshold current and decrease efficiency [14]. Other attractive approaches are photonic crystals or distributed bragg/feedback reflectors [13]. But these approaches clearly violate the nanoscale size requirement for high density optical interconnects.

Even though they violate the size requirement, lasers should not be discounted for on-chip optical interconnects. Because they can satisfy the other requirements, an efficient option could be to have a high-power centralized laser that is split through a series of 3dB couplers for global signal distribution (i.e. clock distribution). We would still need other light sources, however, for dense point-to-point communication.

Case for LEDs

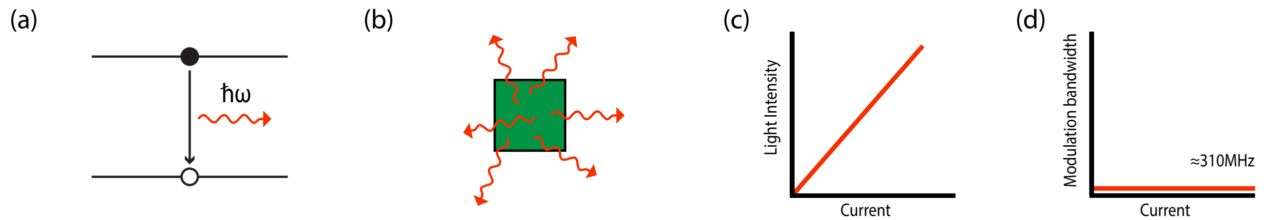

Figure 1.4: Cartoon depiction of (a) spontaneous emission in two level system (b) spontaneous emission in semiconductor composed of many dipoles - note random orientation and phase (c) light intensity (i.e. number of photons per second) as a function of current (i.e. number of electrons per second) (d) modulation bandwidth as a function of current

Unlike lasers, LEDs are capable of scaling down to the nanoscale and can operate without a threshold current. LEDs have been shown to have near unity efficiency but they are limited in speed by their spontaneous emission rate. The 3dB frequency is given below:

$$f_{3dB} = \frac{2B_0N}{2\pi} \tag{1.7}$$

where B_0 is the radiative recombination coefficient and N is the carrier concentration. For a typical InGaAs LED, $B_0 = 10^{-10} cm^3 s^{-1}$ [15]. If we pump the device to a very high carrier concentration of $2 \times 10^{19} cm^{-3}$ then we only get a 3dB frequency of 310MHz. This is why LEDs are generally not implemented for high speed communication applications. However, as I will show in Chapter 2, the spontaneous emission rate can be modified by the optical environment it is placed in, and, if properly designed, the f_{3dB} can exceed 100GHz.

1.4 Outline of this dissertation

In this dissertation, I report a method for coupling an electrically injected optical antenna-LED to a single mode waveguide. In Chapter 2, I review the model for spontaneous emission enhancement, as well as important details on figure of merit calculations for comparing different devices. In Chapter 3, I discuss the cavity backed slot antenna and its properties. Chapter 4 goes through the design methodology and finite difference time domain (FDTD) simulations for efficient waveguide coupling of the cavity backed slot antenna. In Chapter 5, I cover the fabrication and measurements of the optically pumped waveguide coupled antenna-LED. Chapter 6 lays out a novel surface passivation technique to reach high internal quantum efficiency for the device. And, in Chapter 7 the link model, avenues for increasing the power, and outlook are presented.

Chapter 2

Spontaneous emission enhancement

In this chapter the derivation of spontaneous emission and two models for spontaneous emission rate enhancement are reviewed. Additionally, how spontaneous emission rate enhancement relates to the device figures of merit and how to properly account for all the effects are discussed.

2.1 Spontaneous emission

Two popular ways for deriving spontaneous emission are using Einstein's phenomenological AB coefficient analysis or Fermi's Golden Rule [12]. We will go through Fermi's Golden Rule derivation for a two level system below, starting with the transition rate from an initial excited state state $|n_p, i\rangle$ with photon density n_p to final ground state $\langle n_p + 1, f|$ through the emission of a photon.

$$R_{i \to f} = \frac{2\pi}{\hbar} |H_{fi}|^2 \rho(\hbar\omega) \tag{2.1}$$

where H_{fi} is the matrix element describing the transition from the initial state to the final state and $\rho(\hbar\omega)$ is the density of optical states at energy $\hbar\omega$.

The matrix element for a dipole transition can be written as:

$$H_{fi} = \langle n_p + 1, f|\, qx \cdot \mathcal{E}\, |n_p, i\rangle \tag{2.2}$$

And using second quantization we can write the electric field operator as:

$$\mathcal{E} = -i\sqrt{\frac{\hbar\omega}{2V\epsilon}} \times (a_k^\dagger e^{-ik\cdot r} - a_k e^{ik\cdot r}) \times \hat{x} \tag{2.3}$$

where ϵ is the dielectric permittivity of the surroundings, V is the volume of space the dipole is in, $a_k^\dagger$ and a_k are the creation and annihilation operators respectively, and $\hat{x}$ denotes the polarization direction - in this case we assumed a polarization of x. We can drop the

annihilation operator because we are not concerned about absorption in this derivation, so we are left with:

$$H_{fi} = -i\sqrt{\frac{\hbar\omega}{2V\epsilon}} \langle n_p + 1, f|\, qxa_k^\dagger e^{-ik\cdot r}\, |n_p, i\rangle \tag{2.4}$$

From the dipole approximation we get:

$$H_{fi} = -i\sqrt{\frac{\hbar\omega}{2V\epsilon}} \langle n_p + 1|\, a_k^\dagger\, |n_p\rangle \langle f|\, qx\, |i\rangle$$

This reduces to:

$$H_{fi} = -i\sqrt{\frac{\hbar\omega}{2V\epsilon}} \sqrt{n_p + 1}\, \langle f|\, qx\, |i\rangle \tag{2.5}$$

$$|H_{fi}|^2 = (\frac{\hbar\omega q^2}{2V\epsilon})(n_p + 1)|x_{fi}|^2$$

The optical density of states in free space is given by:

$$\rho(\hbar\omega) = \frac{n^3\omega^2}{\hbar c^3 \pi^2} \tag{2.6}$$

We plug Eq. 2.5 and 2.6 into Eq. 2.1 and divide by 3 to account for the polarization averaging over all the dipole modes to get:

$$R_{i\to f} = \frac{1}{V}\frac{\omega^3 n q^2}{3\pi\epsilon_0\hbar c^3}(n_p + 1)|x_{fi}|^2 \tag{2.7}$$

This gives both the stimulated emission rate ($n_p >> 0$) and spontaneous emission rate ($n_p = 0$). For the full emitting volume of V we get the spontaneous emission lifetime ($n_p = 0$) of:

$$\frac{1}{\tau} = \frac{\omega^3 n q^2}{3\pi\epsilon_0\hbar c^3}|x_{fi}|^2 \tag{2.8}$$

2.2 Spontaneous emission rate enhancement

One of the most popular ways of looking at optical rate enhancement is through the Purcell effect; however, we can also derive rate enhancement from an antenna model.

The Purcell effect

In 1946, Purcell [16] noted that the recombination rate can be increased by placing the emitter in a cavity. He found that the optical density of states can be enhanced $\propto \frac{Q}{V}$ over free space, giving us the famous “Purcell factor.” This can be derived if we assume that the optical cavity has a Lorentzian lineshape:

$$L(\nu) = \frac{\frac{1}{\pi}(\frac{1}{2}\Gamma)}{(\nu - \nu_0)^2 + (\frac{1}{2}\Gamma)^2} \tag{2.9}$$

where Γ is the full width at half maximum. For a simple Lorentzian, we can get the Q factor $Q = \nu_0/\Gamma$, and the density of states at resonance is:

$$\rho_{cavity}(\nu_0) = \frac{1}{V}\frac{dN}{dE} = \frac{L(\nu_0)}{V} \tag{2.10}$$

$$\rho_{cavity}(\nu_0) = (\frac{2}{\pi\Gamma})(\frac{1}{V}) = (\frac{2}{\pi\nu_0})(\frac{Q}{V})$$

By dividing the recombination rate for a dipole in a medium over the recombination rate in the cavity we get the famous Purcell factor:

$$F = \frac{R_{i \rightarrow f}, dielectric}{R_{i \rightarrow f}, cavity} = \frac{\rho_{cavity}}{\rho_{freespace}} = \frac{3\lambda^3}{4\pi^2}\frac{Q}{V} \tag{2.11}$$

where F is the Purcell factor. Note, we added a factor of three to account for all the polarization directions.

So to increase the spontaneous emission rate we could either increase the Q or decrease the effective mode volume V_{eff}. Let's first consider the case where we increase the Q factor.

The Q factor is the ratio of energy stored to energy dissipated per cycle, another way of looking at this is through photon lifetime:

$$\tau_{cavity} = \frac{Q}{\omega} \tag{2.12}$$

The total lifetime will determine the direct modulation rate, it is approximately [17]:

$$\tau \approx \sqrt{\tau_{cavity}^2 + (\frac{\tau_0}{F(Q,V)})^2} \tag{2.13}$$

where τ_0 is the spontaneous emission lifetime of an unenhanced dipole and $F(Q,V)$ is the Purcell factor given in Eq. 2.11. From this expression there is an obvious trade-off in increasing the Q factor, the spontaneous emission rate scales with $1/Q$ while τ_{cavity} scales with Q. This is shown in Fig. 2.1.

The plot shows that scaling Q factor alone will be limited in the high mode volume case, but if the Q factor is low then we should be able to increase the modulation rate by scaling the effective mode volume. As discussed in Chapter 1, it is difficult to do this with lasers

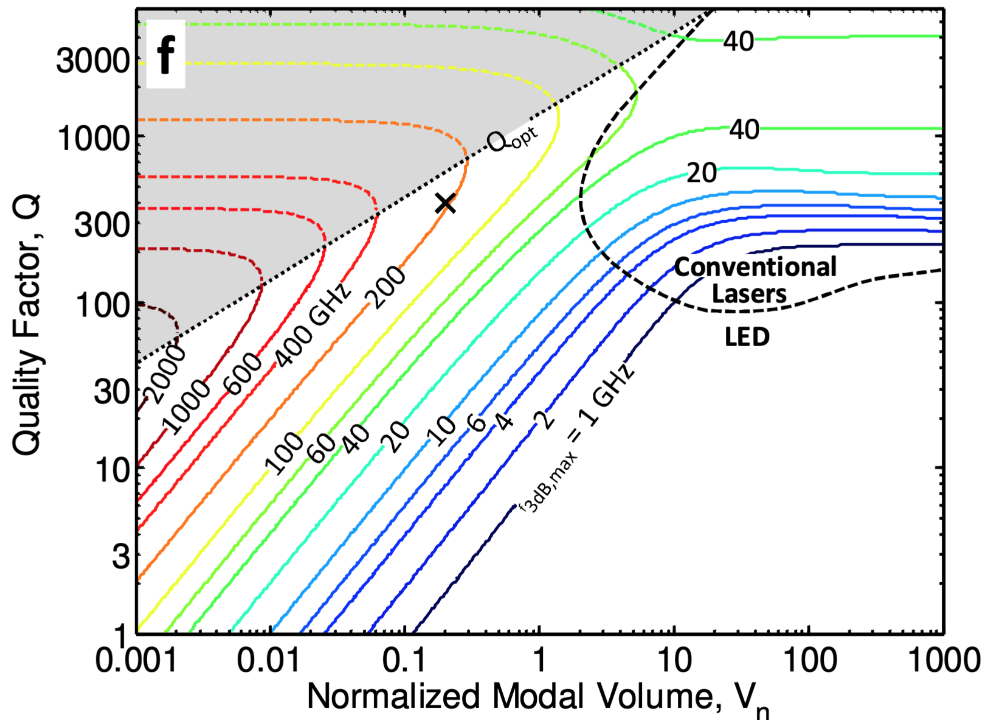

Figure 2.1: Modulation bandwidth as a function of mode volume and Q factor. Q_{opt} is plotted as a dashed line indicating the optimal Q factor for modulation bandwidth, the area above this line is in the strong coupling regime which would require different analysis. Reproduced from [17] ©2009 Optical Society of America

because dielectric cavities would violate the footprint limitation, while metal optics suffer from relatively high loss, making lasing difficult. Accordingly, it would be challenging to simultaneously achieve high Q (low threshold current) and low effective mode volume.

However, LEDs do not need a threshold emission, so we can trade-off the Q factor for a much lower effective mode volume. Because the Q factor is low, the metal-optic cavity can radiate faster and still achieve relatively low loss.

The Purcell factor analysis works well for dielectric cavities where the effective mode volume and Q factor are relatively simple to calculate; however, in low Q metal-optic structures it is harder to get design intuition from Purcell factor alone.

Antenna model

We can also derive the enhancement factor using antenna analysis. First, we can look at the picture for a dipole in free space. As shown in Fermi's Golden Rule analysis of spontaneous emission, the emission can be modeled by a dipole (on the atomic scale) with

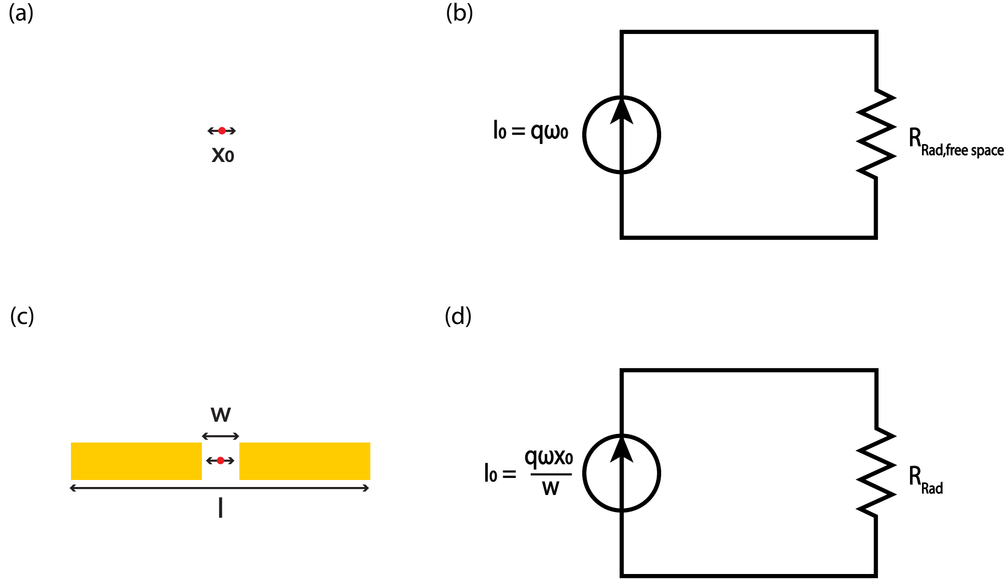

Figure 2.2: (a) Electric dipole in free space with dipole length x_0 and (b) its equivalent circuit model. (c) Electric dipole in feed gap of dipole antenna and (d) its equivalent circuit model.

dipole moment of $p = qx_0$ and oscillation frequency of ω_0. Because the dipole is on the order of $x_0 \approx 0.5nm$ and it is radiating to a free-space wavelength of $\lambda \approx 1550nm$ the power radiated can be modeled as coming from a Hertzian dipole with a current of $I = q\omega_0$ and radiation resistance $R_{Rad,free\ space} = \frac{2\pi}{3} Z_0 n (\frac{x_0}{\lambda_0})^2$, where Z_0 is the impedance of free space [18]. Therefore the power for an unenhanced dipole is:

$$P_0 = \frac{1}{2} I_0^2 R_{Rad,free\ space} \tag{2.14}$$

$$P_0 = \frac{\pi}{3} Z_0 n (\frac{x_0}{\lambda_0})^2 (q\omega_0)^2$$

The electric dipole in free space and equivalent circuit model is shown in Fig. 2.2a and 2.2b.

Now, consider that same dipole when placed in the feed gap of the an antenna as shown in Fig. 2.2c. Similar to free space, the dipole antenna in Fig. 2.2c can be modeled as an equivalent circuit as shown in Fig. 2.2d. If we look at resonance we can simplify the circuit because the reactance goes to zero. Additionally, for first order analysis, we can neglect metal loss. Therefore, the equivalent circuit on resonance just consists of a current source driving a radiation resistance where the current is the induced current from the dipole placed in the antenna feed gap. The induced current in the two metal arms is given by the Shockley–Ramo Theorem as $I_0 = q\omega_0 x_0 / w$ [19, 20]. For the dipole antenna with length l we no longer have a

Hertzian dipole, and instead we can use the radiation resistance for a dipole antenna, given by $R_{rad,dipole} = \frac{\pi}{6} Z_0 n (\frac{L}{\lambda_0})^2$ [18]. The power radiated by the dipole moment coupled to the dipole antenna is then:

$$P_{Rad} = \frac{1}{2} I_0^2 R_{Rad,dipole} \tag{2.15}$$

$$P_{Rad} = \frac{\pi}{12} Z_0 n (\frac{L}{\lambda_0})^2 (\frac{x_0}{w})^2 (q\omega_0)^2$$

Then the rate enhancement is given by the ratio of powers, in other words:

$$F = \frac{P_{Rad}}{P_0} = \frac{1}{4} (\frac{L}{w})^2 \tag{2.16}$$

L is not completely a free parameter. This analysis was done for resonance, so $L \approx \lambda_0/2$ for the fundamental length. For a feed gap of $w = 10nm$ and resonance wavelength of $\lambda_0 = 1550nm$ we could get an enhancement factor of:

$$F = \frac{1}{4} (\frac{775nm}{10nm})^2 \approx 1500\times!$$

From a quick calculation, for an LED with a 3dB frequency of 310MHz, if we put this in the optical antenna we could get a rate of nearly 465GHz - much faster than a conventional laser. However, that quick calculation is only looking at a single dipole; in a typical semiconductor we need to consider all the dipoles in the active region.

2.3 Figures of merit

As discussed in the previous subsection, the presence of the optical antenna causes the dipole to radiate more power than if it was in bulk semiconductor, the ratio of these powers provides the enhancement spectrum [21].

$$F(\omega) = \frac{P_{total}(\omega)}{P_0(\omega)} \tag{2.17}$$

where $P_{total}(\omega)$ is the total amount of power radiated into the simulation (including lost to metal) and $P_0(\omega)$ is the power a dipole in a bulk semiconductor would radiate. On resonance, $F(\omega = \omega_0)$ is the same as the Purcell enhancement calculation or the antenna-LED models.

The convenient part of this equation is that it provides instructions on how to simulate emission enhancement. We can place a dipole in the cavity of our choosing, then place a transmission box around the dipole to capture all the radiated power. This quantity is normalized by a simulation of the dipole in a homogeneous material, also using a transmission box to capture all radiated power.

While $F(\omega)$ is a useful quantity, it is not our figure of merit. To find the amount of rate enhancement we need to consider every dipole in the simulation, which means we need to

account for all the dipoles in the active region, and their polarization, position, and overlap with the material spectrum (essentially $F(\omega, x, y, z, p)$ where p is the polarization). After averaging we can get an average enhancement (F_{avg}), which is directly related to both the output power and the modulation rate. Essentially, we are looking for the multiplicative constants on our maximum enhancement rate:

$$F_{avg} = \mu_{polarization} \times \mu_{spatial} \times \mu_{spectral} \times F_{max} \tag{2.18}$$

where $\mu_{polarization}$, $\mu_{spatial}$, and $\mu_{spectral}$ are the polarization, spatial, and spectral averages, respectively, and, F_{max} is the peak enhancement value for a dipole on resonance with the optimal polarization and position in the cavity.

However, the average enhancement calculation could take hundreds to thousands of simulations for each structure, which is a huge computational demand and impractical without tactics to reduce complexity. In the following subsections we will discuss certain averaging techniques and how we can simplify these calculations:

Polarization averaging

Because we are considering all the dipoles in the active region, is helpful to identify any terms that will drop out to simplify the calculation. The full polarization average for an enhancement spectrum looks like:

$$F(\omega, x, y, z, p) = C_x \frac{P_{total,x}(\omega, x, y, z)}{P_{0,x}(\omega)} + C_y \frac{P_{total,y}(\omega, x, y, z)}{P_{0,y}(\omega)} + C_z \frac{P_{total,z}(\omega, x, y, z)}{P_{0,z}(\omega)} \tag{2.19}$$

where C_x, C_y, and C_z are the coefficients of the matrix element for the particular structure. Note $P_{0,x}(\omega) = P_{0,y}(\omega) = P_{0,z}(\omega)$ because it is a homogeneous medium, so we only need one normalization simulation.

In a simple double heterostructure, we would have $C_x = C_y = C_z = 1/3$. And in a quantum well, with the thickness oriented in $\hat{z}$, the matrix element has a TE/TM polarization dependence given in Table 2.1. For perfectly TE polarized light ($\theta = 0$) the matrix element reduces to $3/2M_b^2$ (i.e. $C_x = C_y = 1/2$).

Because optical antenna-LEDs have their enhancement polarized along one direction (we are choosing $\hat{x}$), we have $P_{total,x} >> P_{total,y}$ and $P_{total,x} >> P_{total,z}$. This means we can drop the $\hat{y}$ and $\hat{z}$ polarizations and are left with the polarization average:

$$\mu_{polarization} \approx C_x \tag{2.20}$$

So, our polarization average, $\mu_{polarization}$, is 1/3 for a double heterostructure, and 1/2 for a conduction band to heavy hole band (C-HH) transition in a quantum well. Note, care needs to be taken to ensure that the width of the LED ridge is not too narrow to create quantization along $\hat{x}$ because the structure would resemble a quantum wire-like and $\mu_{polarization}$ would become small.

Quantum Well Matrix Element			
	TE ($\hat{x}$ or $\hat{y}$)	TM ($\hat{z}$)	2xTE + TM
C-HH Transition	$\frac{3}{4}(1+cos^2\theta)M_b^2$	$\frac{3}{2}(sin^2\theta)M_b^2$	$3M_b^2$
C-LH Transition	$(\frac{5}{4}-\frac{3}{4}cos^2\theta)M_b^2$	$(\frac{1}{2}+\frac{3}{2}cos^2\theta)M_b^2$	$3M_b^2$
Sum Rule: HH+LH	$2M_b^2$	$2M_b^2$	$6M_b^2$

Table 2.1: Quantum well matrix element with quantized dimension oriented along $\hat{z}$ [12].

Position averaging

Now that we effectively eliminated the $\hat{y}$ and $\hat{z}$ polarizations, a position average is much easier to solve. We would still need to numerically integrate the following quantity - we use sum notation for practicality because this could be solved in FDTD.

$$\mu_{spatial} = \frac{\sum_{x,y,z} F(\omega, x, y, z, p = \hat{x})\Delta_x\Delta_y\Delta_z}{\sum_{x,y,z} Max(F(\omega, x, y, z, p = \hat{x}))\Delta_x\Delta_y\Delta_z} \tag{2.21}$$

For our devices, a very accurate approximation is that the spectral shape of the enhancement is unaffected by the position - only the magnitude changes. Therefore we just look at the value on resonance.

There are three methods for calculating this average:

1. **Brute Force.** This is where we calculate nested parameter sweeps over x, y, and z so we can look at every point in the active region. The brute force method is obviously the most computationally intensive solution, but should be the most accurate when performed on a fine mesh.

2. **Analytical Formula.** This method relies on an analytical solution of the shape of the mode $F(x, y, z)$. If this known or modeled ahead of time then we can simply change the sum notation to an integral and integrate over the normalized x, y, z enhancement.

3. **Lorentz Reciprocity.** This is a clever method for calculating the spatial distribution of the enhancement from a single simulation. Lorentz reciprocity states:

$$\int E_{antenna}J_{farfield} = \int E_{farfield}J_{antenna}$$

where $E_{antenna}$ is the electric field produced by current source $J_{antenna}$, and $E_{farfield}$ is the electric field produced by current source $J_{farfield}$. So to excite the antenna mode we can place a current source in the antenna farfield mode distribution and then measure

the electric field profile. Because $F \propto |E_{antenna}|^2$ to get the normalization we can just take:

$$\mu_{spatial} = \frac{\sum_{x,y,z} |E_{antenna}(x,y,z)|^2 \Delta_x \Delta_y \Delta_z}{\sum_{x,y,z} |E_{antenna,max}|^2 \Delta_x \Delta_y \Delta_z}$$

For a constant mesh spacing, which is common in FDTD, we get a simple average of the normalized electric field squared:

$$\mu_{spatial} = \frac{1}{n_{points}} \sum_{x,y,z} \frac{|E_{antenna}(x,y,z)|^2}{|E_{antenna,max}|^2}$$

In Chapter 7 we will compare these methods for our antenna geometry, but the spatial averaging for antenna-LEDs is generally on the order of $\approx 0.4 - 0.75$.

Spectrum overlap averaging

The final average concerns dipoles at different energies, this can be visualized in Fig. 2.3.

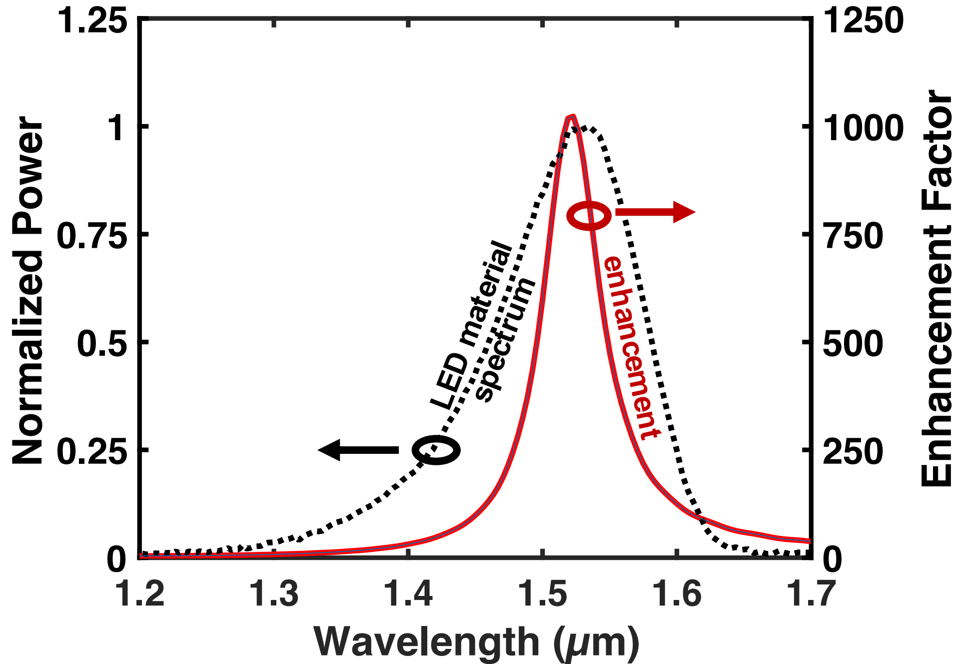

Figure 2.3: Dashed black and solid red lines show the experimental non-enhanced material spectrum $[L(\omega)]$ and the simulated enhancement spectrum $[F(\omega)]$ of the cavity-backed slot antenna on a bulk InP substrate, respectively.

The spectral average is then given by:

$$\mu_{spectral} = \frac{\int F(\omega)L(\omega)d\omega}{\int F_{max}L(\omega)d\omega} \tag{2.22}$$

where $F(\omega)$ is the overall enhancement spectrum seen by a dipole with the optimal polarization and position, and $L(\omega)$ is the experimental material spectrum without an antenna present. In principle the material spectrum $L(\omega)$ is dependent on the carrier concentration; however, in this report we fixed $L(\omega)$, and therefore the carrier concentration in order to simplify the analysis. For high injection this assumption should be revisited.

Antenna and waveguide-coupling efficiency

In addition to the average enhancement, we considered the antenna efficiency (η_{antenna}) and waveguide coupling efficiency to the fundamental mode (η_{WC}). The antenna efficiency only accounts for the metal loss, and the waveguide coupling efficiency only accounts for the scattering loss. The explicit definitions for the antenna efficiency and waveguide coupling efficiency spectra are shown below:

$$\eta_{\text{antenna}}(\omega) = \frac{P_{\text{total}}(\omega) - P_{\text{metal loss}}(\omega)}{P_{\text{total}}(\omega)} \tag{2.23}$$

$$\eta_{\text{WC}}(\omega) = \frac{1}{\eta_{\text{antenna}}(\omega)} \frac{P_{\text{fundamental mode}}(\omega)}{P_{\text{total}}(\omega)} \tag{2.24}$$

$$P_{\text{total}}(\omega) = P_{\text{fundamental mode}}(\omega) + P_{\text{scattering}}(\omega) + P_{\text{metal loss}}(\omega) \tag{2.25}$$

where $P_{\text{total}}(\omega)$ is the total optical power leaving the dipole source, $P_{\text{metal loss}}(\omega)$ is the power lost to metal, and $P_{\text{fundamental mode}}(\omega)$ is the power in the fundamental mode of the waveguide which was found by taking an overlap integral between the eigenmode solution and the simulated waveguide field profile. This is given by the overlap monitor in Lumerical FDTD which calculates the quantity:

$$\eta_{\text{WC}}(\omega) = Re\left(\frac{(\int \vec{E}_{measured}(\omega) \times \vec{H}^*_{TE00}(\omega) \cdot d\vec{S})(\int \vec{E}_{TE00}(\omega) \times \vec{H}^*_{measured}(\omega) \cdot d\vec{S})}{(\int \vec{E}_{measured}(\omega) \times \vec{H}^*_{measured}(\omega) \cdot d\vec{S})(\int \vec{E}_{TE00}(\omega) \times \vec{H}^*_{TE00}(\omega) \cdot d\vec{S})}\right) \tag{2.26}$$

where *measured* subscript denotes that it is the measured field in the simulation, and the $TE00$ subscript indicates simulated fundamental TE mode fields.

Note that the product of antenna efficiency and waveguide coupling efficiency gives the fraction of the total optical power coupled to the fundamental waveguide mode. Additionally, we calculated the average antenna efficiency and waveguide coupling efficiency. Below are the explicit definitions for η_{antenna} and η_{WC}:

$$\eta_{\text{antenna}} = 0.96 \times \frac{\int \eta_{\text{antenna}}(\omega)F(\omega)L(\omega)d\omega}{\int F(\omega)L(\omega)d\omega} \tag{2.27}$$

$$\eta_{\text{WC}} = \frac{\int \eta_{\text{WC}}(\omega)\eta_{\text{antenna}}(\omega)F(\omega)L(\omega)d\omega}{\int \eta_{\text{antenna}}(\omega)F(\omega)L(\omega)d\omega} \tag{2.28}$$

where 0.96 is the spatial average for the antenna efficiency determined by the brute force method. Note that the polarization dependence was negligible for both average efficiencies, because a dipole oriented along the width of the LED sees much greater enhancement than a dipole oriented along the length. Additionally, the spatial dependence was negligible for the waveguide coupling efficiency.

Device metrics

These average values could now be used to calculate relevant device metrics because they represent the average response of a carrier in the device. Two important metrics are the power in the fundamental mode of the waveguide and the 3dB frequency, given in Eqs. 2.29 and 2.30, respectively.

$$P_{\text{fundamental mode}} = F_{\text{avg}}\eta_{\text{antenna}}\eta_{\text{WC}}\hbar\omega B_0 N^2 V \tag{2.29}$$

$$f_{\text{3dB}} = \frac{2F_{\text{avg}}B_0 N}{2\pi} \tag{2.30}$$

where B_0 is the radiative recombination coefficient, N is the carrier concentration, V is the active region volume, and f_{3dB} is the 3dB modulation frequency assuming the radiative recombination rate is dominant. If we assume $F_{\text{avg}} = 164$, $B_0 = 10^{-10}\text{cm}^3\text{s}^{-1}$ [15], and $N = 2 \times 10^{19}\text{cm}^{-3}$ we could reach a 3dB frequency of 104GHz.

Chapter 3

Cavity backed slot antenna-LED

In the previous chapter we discussed spontaneous emission enhancement and how decreasing mode volume is an effective method for achieving high speed operation.

There are many geometries that have been considered to accomplish this in the literature: inelastic scattering of metal [22], using plasmonic nanopatch antenna to enhance quantum dots and 2D semiconductors [23, 24], a dipole antenna [21, 25], slot antenna [26], and cavity backed slot antenna [27, 28, 29]. And although different, J-aggregates can be thought of as creating a longer molecular dipole to enhance the emission [30].

Out of these reports, only a few have demonstrated electrical injection [27, 31, 32, 22], with the electrically injected cavity-backed slot antenna demonstrating ~200x peak enhancement [27]. We will accordingly focus on the cavity backed slot antenna-LED for the remainder of this dissertation.

3.1 Cavity backed slot antenna

In the previous chapter we discussed the circuit model for the dipole antenna to determine the enhancement. However, the dipole antenna has some drawbacks, namely: the fabrication is difficult, especially when trying to form contacts that do not perturb the antenna mode [28]. The complimentary structure to the dipole antenna is a slot antenna, shown in Fig. 3.1b, but this structure shares similar problems to the dipole antenna. Electrical contacts will still be difficult without perturbing the slot antenna mode, and the radiation pattern is concentrated primarily in two lobes going in opposite directions (Fig. 3.1e), which makes it difficult to achieve efficient waveguide coupling.

A solution to this is problem is to form a cavity on one side, which will make it a cavity backed slot antenna, Fig. 3.1c. The cavity backed slot antenna has several advantages over the dipole and slot antennas:

1. **Self-Aligned.** Because the antenna wraps around the LED ridge alignment and deposition of the metal layer becomes trivial.

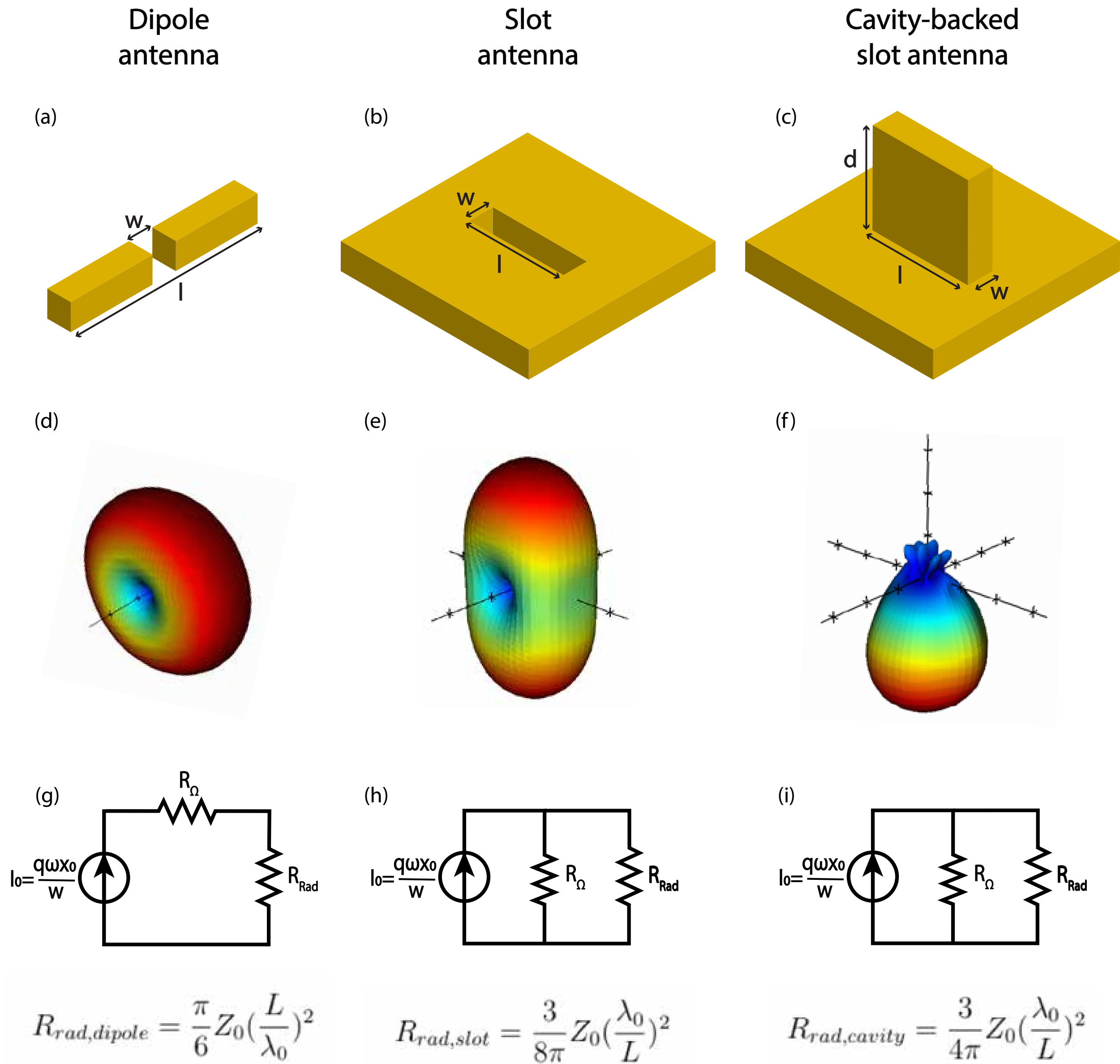

Figure 3.1: Comparison of the dipole antenna, slot antenna, and cavity backed slot antenna geometry (a) - (c), farfield radiation pattern ($|E|$) (d)-(f), and circuit model with radiation resistance (g)-(i).

2. **Electrical Injection.** The metal for the cavity backed slot antenna can also serve as a relatively large contact to the top of the LED ridge. This is important to reduce contact resistance, which can be very high for nanoscale devices. A cartoon of the

structure is shown in Fig. 3.2.

3. **Thermal Heat Sink.** In order to get sufficient power for the link we will need to pump the device to the $N \approx 10^{19} cm^{-3}$ range. Even though the total current can be small, the current density can be very high due to the nanoscale cross-section. Moderate losses can contribute to excessive heating without proper thermal management.

4. **Directional Emission.** As we will discuss in the next chapter, having directional emission is beneficial for coupling to a waveguide.

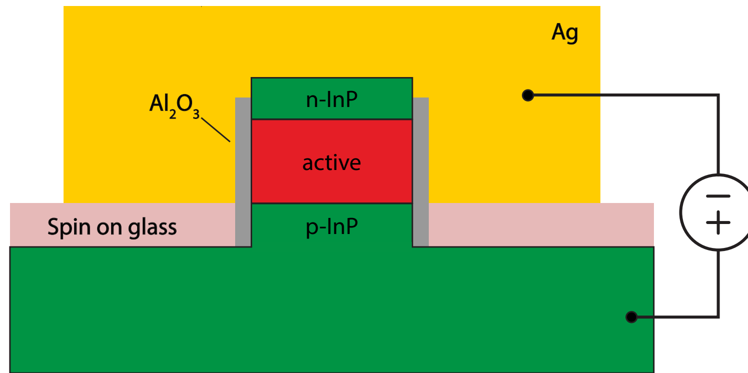

Figure 3.2: Cartoon showing electrical injection for cavity backed slot antenna-LED. The metal serves as both the antenna and the n-contact to inject electrons. The holes can be injected through the p-InP substrate. Note this is not drawn to scale.

Electrical injection can be achieved by doping the LED ridge. The electrons can be injected into the n-type layer by using the antenna as a contact, and the holes can be injected into the p-type layer. When the electron and hole recombine in the active region, it serves as an excitation of the cavity-backed slot antenna mode. The Al_2O_3 acts as isolation between the antenna and active region to prevent electrical shorting; likewise, the spin on glass (SOG) electrically isolates the antenna and the p-InP substrate.

Additionally, as shown in Fig. 3.1e, compared to the slot antenna, the cavity-backed slot antenna directs power in one direction, which helps with waveguide coupling, but the formation of the cavity gives rise to another condition on the resonance. The round-trip length of the light bouncing off the top of the cavity needs to satisfy [28]:

$$\Delta\phi = 2k_z d + \pi = 2\pi \tag{3.1}$$

where k_y is the propagation constant for the fundamental TE mode in the antenna. The propagation constant k_z for a fundamental TE mode is given by:

$$k_z = \sqrt{\omega^2 \mu\epsilon - (\frac{\pi}{L})^2}$$

which means we can now solve for ω to give us the resonance condition:

$$\omega_0 = \frac{\pi c}{n}\sqrt{(2d)^{-2} + L^{-2}}$$

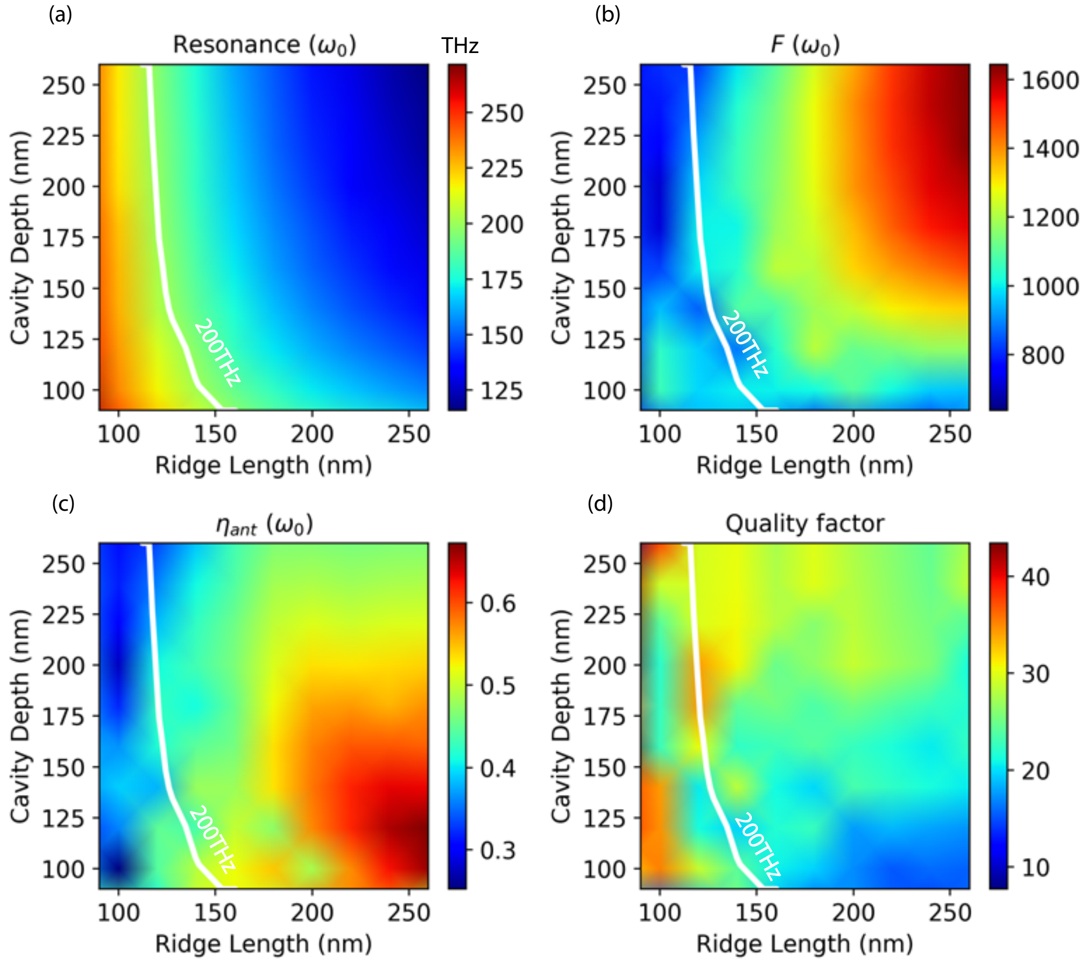

Figure 3.3: Contour plots of the (a) resonance frequency ω_0, (b) overall emission enhancement, (c) antenna efficiency on resonance, and (d) quality factor of the III–V antenna-LED with 1nm Al_2O_3 for a 20nm wide device. The solid white contour line denotes the combinations of slot length and cavity depth that achieve antenna resonance at 200 THz.

Because the cavity-backed slot antenna re-redirects the radiation into the lower half-space ($z < 0$), the magnitude of the electric field is doubled with respect to the slot antenna without a cavity. Doubling of the electric field yields a power density that is four times as high. However, the cavity-backed slot antenna radiates into half as much volume as the slot

antenna and therefore the overall radiated power is increased only by a factor of two. This can be modeled as twice the radiation resistance of the slot antenna:

$$R_{rad,cavity} = 2 \times R_{rad,slot} = \frac{3}{4\pi} Z_0 (\frac{\lambda_0}{L})^2 \tag{3.2}$$

Plugging this into the enhancement equation we get the following for the fundamental mode:

$$F = \frac{18}{\pi^2} (\frac{L}{w})^2 \tag{3.3}$$

A summary of how the cavity backed slot antenna's properties change for cavity depth (d) and ridge length (L) are shown in Fig. 3.3 for an LED ridge with 20nm width and 1nm Al_2O_3.

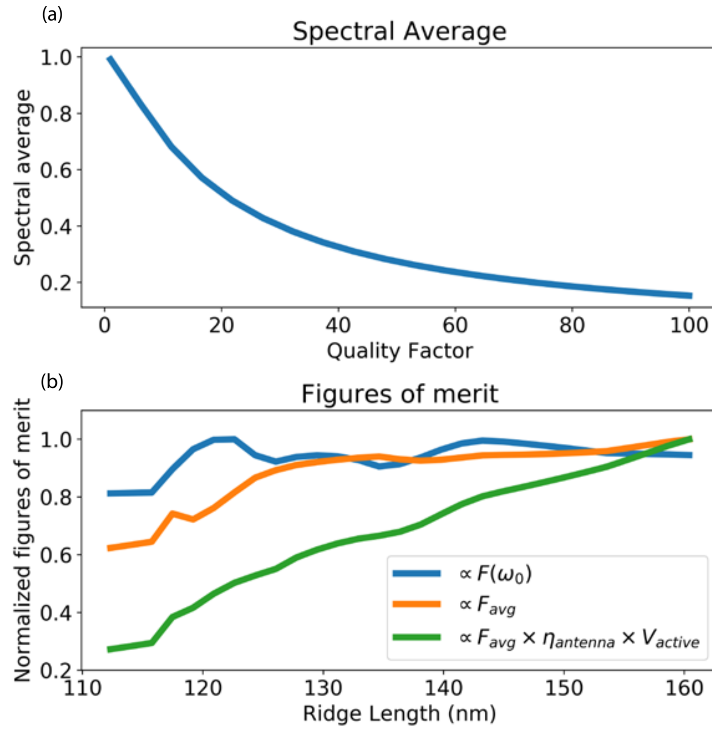

Figure 3.4: (a) Spectral average $\mu_{spectral}$ as a function of Q and (b) average enhancement F_{avg} and radiated power $\propto F_{avg} \times \eta_{antenna} \times V_{active}$ for slots with 200THz resonance.

As we increase the slot length or depth, we shift the resonance to longer wavelengths. Second, we see that for larger cavity depth the antenna loss is higher, because the photons are in the metal-optic cavity for longer. Finally, the resonator loses power more slowly so Q increases.

While the properties of the cavity backed slot antenna are useful for understanding the device trade-offs, we need to convert this into our system figures of merit - namely average enhancement. We can track this by modeling the spectral average $\mu_{spectral}$ as a function of Q, shown in Fig. 3.4a, discussed in Chapter 2.3. For simplicity, we used a Lorentzian for the enhancement $F(\omega)$ for various Q factors.

We then assume that the spatial average change is negligible for different slot lengths using a quantum well active region (i.e. height difference is neglected for quantum wells near the opening of the slot). We see in Fig. 3.4b that the average enhancement increases by around 1.7x as we increase the ridge length, and that the power increases by approximately a factor of 3x. But this comes with a trade-off, namely that as we increase the length of the device we need to decrease the height to compensate, which can make electrical injection difficult because we need contact and barrier layers. In order to leave enough room for electrical injection we fix this height to be 130nm, which gives us a target length of 130nm as well. We see in 3.4b that the average enhancement nearly levels out for this value, and the power would be approximately 60% of the peak value.

3.2 Aluminum oxide

Adding an oxide on the surface of the LED-ridge is important to electrically isolate the device from the metal. However, there is an important trade-off with the spontaneous emission enhancement factor $(F(\omega) \times \eta_{antenna})$ as shown in Fig. 3.5. This result can be explained by looking at Shockley-Ramo theorem with an inhomogeneous medium [28].

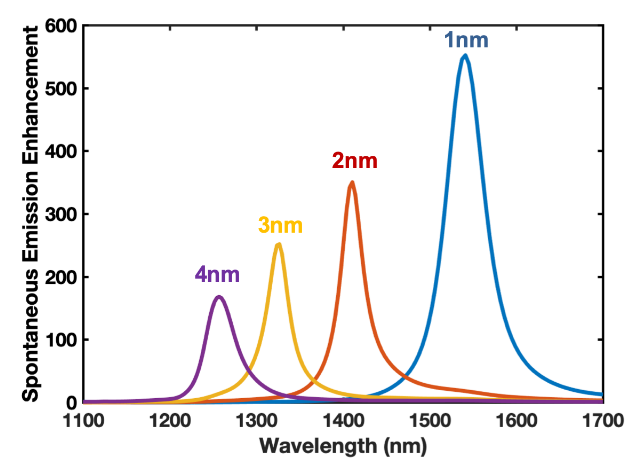

Figure 3.5: Simulation of spontaneous emission enhancement dependence on Al_2O_3 thickness. Antenna length was not adjusted for increasing thickness leading to a blue-shift of antenna resonance.

$$w_{eff} = w + 2\frac{\epsilon_s}{\epsilon_{ox}}t_{ox}$$

where w_{eff} is the effective width, ϵ_s and ϵ_{ox} are the permittivities of the semiconductor and oxide, respectively, and t_{ox} is the oxide thickness between the semiconductor and metal. The factor of 2 is from the structure symmetry. So, by increasing the effective width by adding an oxide, we blue-shift the resonance and decrease the peak enhancement.

From the device metrics in Chapter 2, we know that power and speed are proportional to enhancement. We therefore want to make the oxide as thin as possible without compromising electrical performance. To determine the minimum thickness we fabricated a simple cavity backed slot antenna clad with different Al_2O_3 thicknesses then measured the surface decay as a function of the ridge width, shown in Fig. 3.6a.

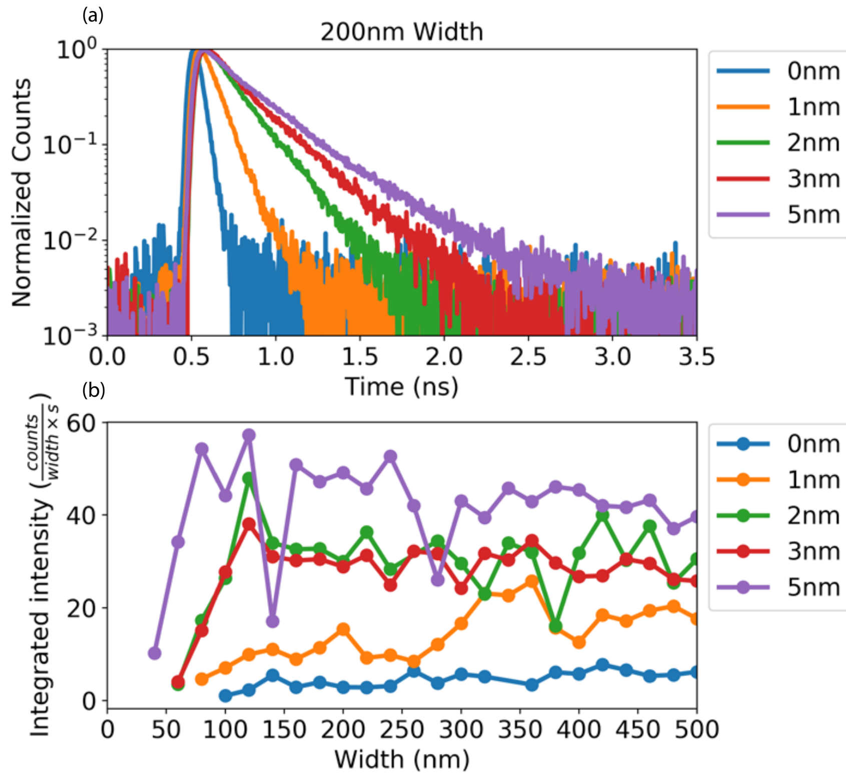

Figure 3.6: (a) Experimental decay curves for 200nm ridge device lifetime dependence based on Al_2O_3 thicknesses between 0-5nm. (b) Integrated decay intensity for variable ridge widths.

We determined that the lifetime of the device decreases with decreasing Al_2O_3 thicknesses. The 0nm and 1nm thicknesses are too thin - the decay became prohibitively fast when we

made narrower ridges $w \approx 20nm$, making the devices very inefficient at these dimensions. In order to have higher efficiency we need to use $\geq 2nm$ Al_2O_3 thicknesses. This trend can also be seen in the intensity data. In 3.6b, both the 0nm and 1nm thicknesses have low intensity at narrow ridge widths, and become prohibitively dim for devices with widths below 80nm; however, the thicker oxides still have sufficient intensity for measurements.

3.3 Averaging

As mentioned in Chapter 2.3, we need to calculate the average quantities to get the figures of merit. We will compare the three methods for averaging for a quantum well active region and double heterostructure active region.

Active region design

For a cavity-backed slot antenna oriented along the x-axis, in our simulations we observed that the distribution of voltage is nearly constant along the width ($\hat{y}$), a cosine distribution along the length ($\hat{x}$), and a square root distribution along ($\hat{z}$) for the fundamental mode. Finally, we assume that the x, y, and z components are approximately separable:

$$V(x,y,z) = V_0 Cos(\frac{\pi}{L+2\delta}x)\sqrt{1-\frac{z}{d+\delta}}$$

where δ is the skin depth of the metal, and w, L, and d are the width, length, and depth of the cavity-backed slot antenna, respectively. Because the enhancement is proportional to the electric field squared, the spatial average is:

$$\mu_{spatial} = \frac{\int_{x=-L/2}^{L/2}\int_{y=-w/2}^{w/2}\int_{z=0}^{h} Cos^2(\frac{\pi}{L+2\delta}x)(1-\frac{z}{d+\delta})dxdydz}{\int_{x=-L/2}^{L/2}\int_{y=-w/2}^{w/2}\int_{z=0}^{d} dxdydz}$$

where h is the height of the active region. In our design the total depth is 130nm, and for a single quantum well the active region height is approximately 5nm, where the rest of the height can be used for barriers and cladding. In a double heterostructure we assume that the active region height is approximately 110nm ($\approx d - 20nm$). If we assume a perfect metal ($\delta = 0$) then we get:

$$\mu_{spatial,QW} \approx \frac{1}{2}$$

$$\mu_{spatial,DH} \approx \frac{1}{4}$$

We can approximate the optical skin depth in silver to be $\approx 20-30nm$ - this adjustment gives us:

$$\mu_{spatial,QW} \approx 0.638$$

$$\mu_{spatial,DH} \approx 0.399$$

However, when we tried to verify the quantum well numbers using the brute force method, we found that the assumption that the profile along the width is constant does not hold - the electric field actually increases near the interface. When we account for that, the spatial average ends up being:

$$\mu_{spatial,QW} \approx 0.79$$

where the reference point is the center of the slot located at the opening (opposed to the maximum enhancement, which actually occurs near the interface). Due to computational time and memory limitations we did not do brute force for multiple quantum well or double heterostructure geometries.

Lorentz reciprocity

As discussed in Chapter 2.3, we can calculate the spatial average through Lorentz reciprocity by placing a dipole in the farfield of the antenna and measuring the excited mode. This is shown in Fig. 3.7, where we plot the normalized $|E|^2$ field, normalized with respect to the field at $(X = 0, Y = 0, Z = 10)$ for antenna with length 132nm, width 20m, height 140nm measured on resonance ($\lambda_0 = 1537nm$).

We can now now take the average field for different active region configurations:

$$\mu_{spatial,Single\ QW} \approx 0.771$$

$$\mu_{spatial,3x\ QWs} \approx 0.624$$

$$\mu_{spatial,DH} \approx 0.387$$

Note the single quantum well has a higher spatial average because it was calculated at the opening of the slot, similar to the brute force method.

We can now create an expected enhancement map by multiplying the enhancement at $(X = 0, Y = 0, Z = 10)$ by the normalized electric field - this is shown in Eq. 3.4.

$$F_{predicted}(x, y, z) = \frac{|E(x, y, z)|^2}{|E(0, 0, 10)|^2} F(0, 0, 10) \tag{3.4}$$

In order to test this method we placed dipoles at random locations and then compared to the predicted value. The error was typically $< 5\%$, which shows that this method provides a good approximation of the enhancement. This is illustrated by comparing the single quantum well case using brute force with Lorentz reciprocity - we get averages of 0.79 and 0.77, respectively ($\approx 2.5\%$ error). A summary of the calculations is shown in Table 3.1.

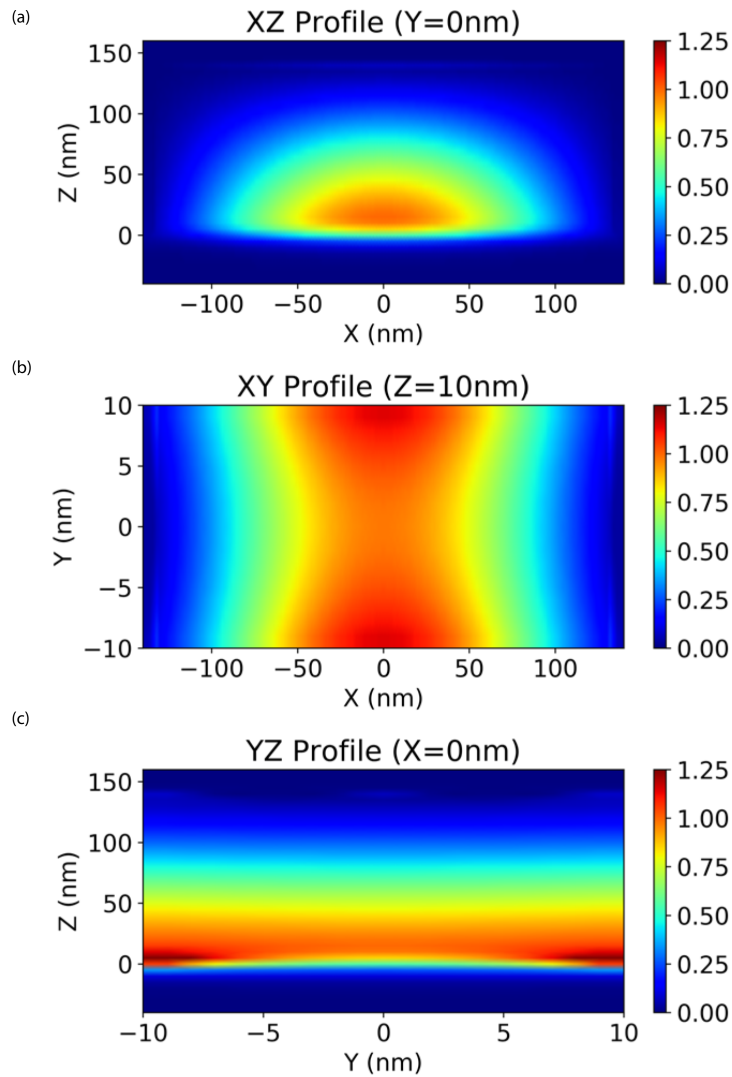

Figure 3.7: Normalized $|E|^2$ field with respect to $(x = 0, y = 0, z = 10nm)$ excited from electric dipole 500nm below slot opening (a) XZ profile (b) XY profile and (c) YZ profile.

Active Region Design				
Design	$\mu_{polarization}$	$\mu_{spatial}$	$\mu_{spectral}$	Height
Single quantum well	$\frac{1}{2}$	$\approx 0.77 - 0.79$	0.394	$\approx 6nm$
Multiple quantum well (3x)	$\frac{1}{2}$	$\approx 0.588 - 0.624$	0.394	$\approx 18nm$
Double heterostructure	$\frac{1}{3}$	$\approx 0.399 - 0.456$	0.394	$\approx 110nm$

Table 3.1: Summary of approximate dipole averages for single quantum well, multiple quantum wells, and double heterostructure active regions. The polarization average is based on the matrix element, the spatial average is calculated from Lorentz reciprocity and the analytical formula, and the spectral average is an approximation based on the Q of the device (this is assumed not to change).

We see that the single quantum well design has the highest product of averages, so will have the highest average enhancement, and, therefore, the highest modulation speed. But this design has the lowest active region volume and accordingly the lowest output power ($Power \propto F_{avg} \times V_{active}$). The double heterostructure has the highest output power but the lowest modulation rate. Among these designs the multiple quantum well structure is a good compromise for high speed with moderate power.

Chapter 4

Waveguide coupled cavity backed slot antenna design

In the previous chapter, we showed that the cavity-backed slot antenna has been demonstrated to satisfy the following requirements: nanoscale, efficient, direct modulation, and compatible with top-down fabrication. However, in order to create a full-link, it needs to be coupled to single mode waveguide. As shown in Fig. 3.1(f), the radiation of the cavity-backed slot antenna is primarily directed down towards the substrate, making it a non-trivial problem to couple to a photonic waveguide.

Many methods have been used to couple nanoscale devices to waveguides, including coupling an optically pumped dipole antenna to a multi-mode waveguide using the waveguide height to cancel the electric field propagating toward the substrate [33], an electrically injected metal cavity LED and laser on a single-mode waveguide using the mode shape in the metal cavity [34, 35], and using anti-symmetric second-order resonance for a double nanogap plasmonic antenna [36]. Overall, efficient devices that are compatible with electrical injection and have high enhancement are still needed.

In this chapter, the impact of each design step on efficient waveguide coupling to the fundamental mode is described.

4.1 Starting structure

As discussed in Chapter 3, the cavity-backed slot antenna is a promising candidate as an optical source due to its high spontaneous emission enhancement and compatibility for top down fabrication and electrical injection [27]. As shown in Fig. 3.1(c), the cavity-backed slot antenna is self-aligned to an InP/InGaAs/InP ridge (length: 130nm, width: 20nm, height: 140nm), where the height and length were chosen to tune the resonance frequency to best match the LED material spectrum while maximizing the radiated power for the fundamental antenna mode. The antenna is electrically connected to the top of the ridge, where it is used as a contact to inject electrons into the n-InGaAs contact layer. The holes are injected

into the p-InP layer, which is insulated from the antenna using a 40nm thick spin on glass (SOG). Finally, the InGaAs quantum well active region is electrically insulated from the antenna using a 1nm thick Al_2O_3 surrounding the ridge sidewalls. When an electron and hole recombine in the active region, the recombination acts as a dipole excitation of the antenna mode. In our 3D finite-difference time-domain (FDTD) simulations, we excited the antenna by placing an electric dipole source in the active region. Unless otherwise stated, we will be simulating our different designs using this base structure. The average enhancement for this structure is $F_{avg} = 160$ and $\eta_{antenna} = 54\%$

4.2 Destructive interference

The first two parameters we optimized were the waveguide height and width in order to cancel the fields propagating towards the substrate (similar to [33]). However, as shown in Fig. 4.1, this case is slightly different because there is metal over the waveguide to define the cavity-backed slot antenna. From a simple ray optics picture, the metal causes an additional $+\pi$ phase shift when reflecting from this surface. We were able to maximize the waveguide coupling efficiency with a waveguide height of 180nm and width of 550nm. With this structure we can only achieve a maximum waveguide coupling efficiency of 17% in each direction on resonance (34% if we account for both directions), which leaves room for improvement.

4.3 Total internal reflection

One way to understand waveguide coupling is by considering total internal reflection (TIR). The total internal reflection condition for light going from the waveguide to the cladding is shown below:

$$\theta_{TIR} = sin^{-1}(\frac{n_{cladding}}{n_{waveguide}}) \tag{4.1}$$

where $n_{waveguide}$ and $n_{cladding}$ are the refractive indices of the waveguide and cladding, respectively, and θ_{TIR} is the angle where total internal reflection begins. If the light experiences total internal reflection then it will be trapped in the waveguide, so we want this angle to be as small as possible. In our case, $n_{waveguide}$ is fixed by our active material, so we will be use the index of the InP/InGaAs material system. We have a little more control over $n_{cladding}$, and from the TIR equation we want to minimize the cladding refractive index to minimize θ_{TIR}. In Fig. 4.2 we plotted the waveguide coupling efficiency as a function of the lower cladding refractive index, and showed that expected relationship holds.

As we can see in Fig. 4.2, there is a considerable improvement in waveguide coupling efficiency by replacing the SiO_2 lower cladding with air (or vacuum); however, that would not be compatible with top-down fabrication. Nonetheless we found there is a method to achieve comparable improvement by etching the SiO_2 back to form an SiO_2 ridge for the waveguide

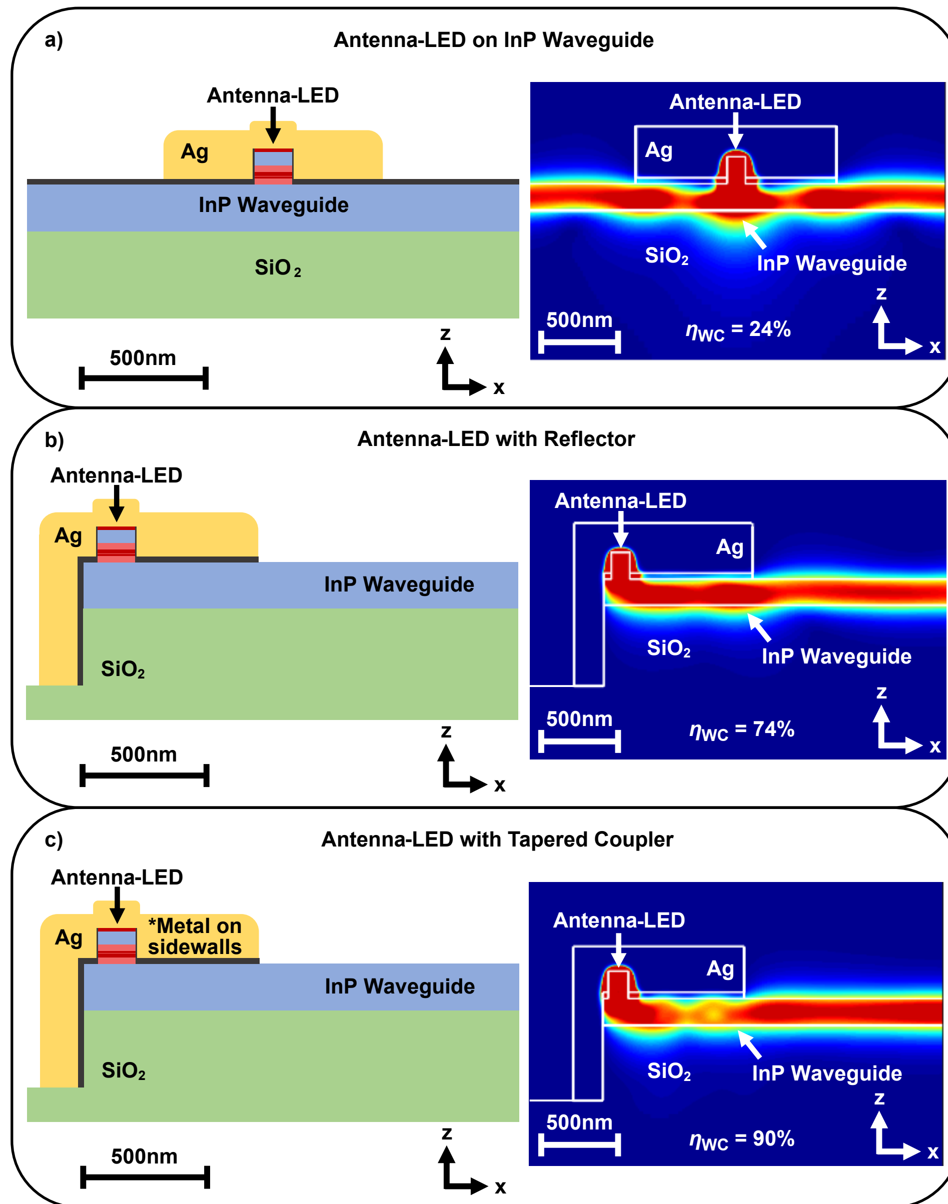

Figure 4.1: Cross section, power flow, and waveguide coupling efficiency to the fundamental mode (η_{WC}) for (a) antenna-LED on single-mode InP waveguide and SiO_2 ridge, (b) antenna-LED on single-mode InP waveguide with metal wrapped around waveguide facet, and (c) antenna-LED on single-mode InP tapered waveguide with metal wrapped around waveguide facet and sidewalls (see Fig. 4.6(a) for perspective view, Fig. 4.8(b) for top view cross section). See *Appendix: Field profiles* for the E_{x} and E_{y} field profiles of the mode in the InP waveguide.

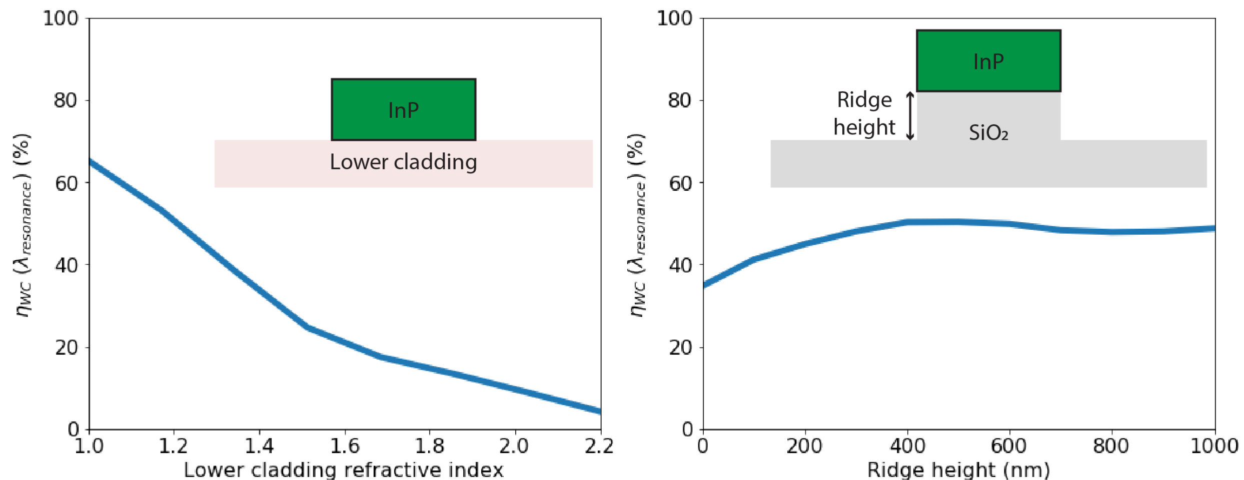

Figure 4.2: Waveguide coupling efficiency (accounting for both directions) at resonance as for a) lower cladding refractive index and b) ridge height. The waveguide coupling efficiency reported is for bi-directional coupling, half of the power travels in each direction. Graph insets show width-height cross section of waveguide.

to sit on top of. In Fig. 4.2, we plot the waveguide coupling efficiency as a function of SiO_2 etch depth. For larger etch depths we create a lower effective index, improving the waveguide coupling efficiency. Keeping fabrication and thermal effects in mind, we wanted the SiO_2 etch depth to be as small as possible while maximizing waveguide coupling efficiency, so we determined that a 500nm etch depth is optimal and will keep this parameter fixed for the subsequent simulations. The simulation results from this structure are shown in Fig. 4.1, we achieved a bi-directional 24% waveguide coupling on resonance (48% total).

4.4 Flat facet reflector

In Fig. 4.1(b), we truncated the waveguide and wrapped metal around the end of the facet to effectively act as a mirror. In addition to making the coupling unidirectional, the mirror created an image dipole 180° out of phase with the antenna-LED, which further suppressed fields propagating toward the substrate. By minimizing the separation between the antenna-LED and the back mirror, we achieved a waveguide coupling efficiency of 74% – note this was more than double the result from Fig. 4.1(a).

In Fig. 4.3, we see the expected constructive and destructive interference as we sweep the separation between the antenna-LED and the metal facet. We can consider the structure from an image dipole perspective or a reflection perspective to find the destructive interference condition.

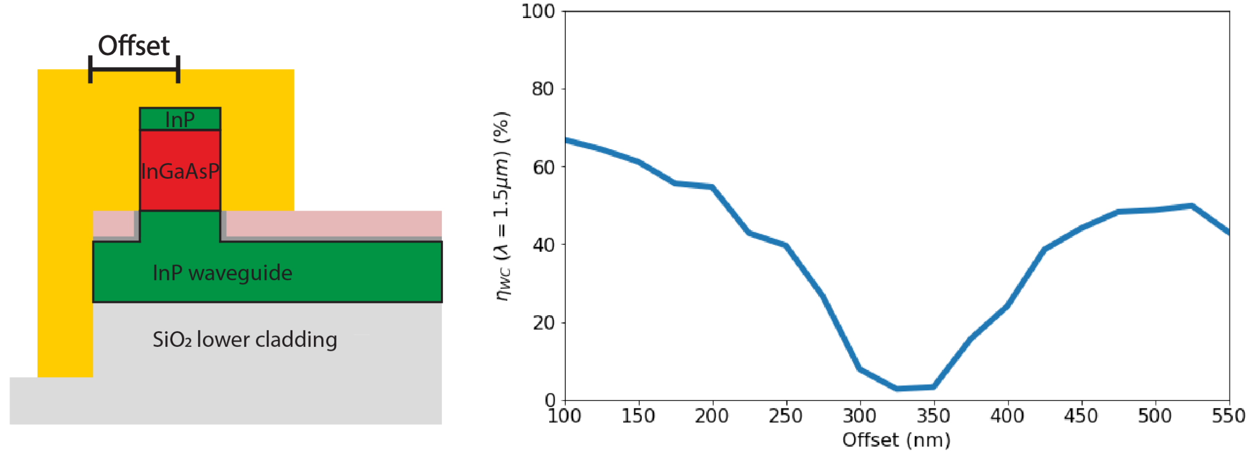

Figure 4.3: Waveguide coupling efficiency at 1500nm as a function of the separation between the center of the antenna-LED and the metal reflector.

Using a simple ray-bounce diagram, we know that we would acquire an extra $+\pi$ phase shift off the back metal plane, so there is destructive interference when the waveguide path length adds up to 2π. The image dipole standpoint is the same, where we have a dipole $+\pi$ out of phase at a distance of twice the offset, so we should see destructive interference when the following condition is met:

$$m\pi = \pi + 2k_x x \tag{4.2}$$

where m is an odd integer and x is the offset. Because the offset is defined from the center of the antenna-LED, the first minimum is at $k_x x = \pi$. We expect the minimum to occur at:

$$x = \frac{\lambda_0}{2n_{eff}} = \frac{1500nm}{2 * 2.08} \approx 360nm$$

where n_{eff} is the effective index of the fundamental mode under the metal. This agrees well with the minimum shown in Fig. 4.3. From this figure we can also see how the emission towards the substrate is suppressed at small offsets, and as we move to the second peak this suppression decreases.

4.5 Tapered waveguide coupler

In this subsection we will describe ways to prohibit higher order modes - in other words, how to shape the emission so it can couple to the desired waveguide mode. By using image theory, we can direct more light towards the waveguide; however, image theory alone cannot

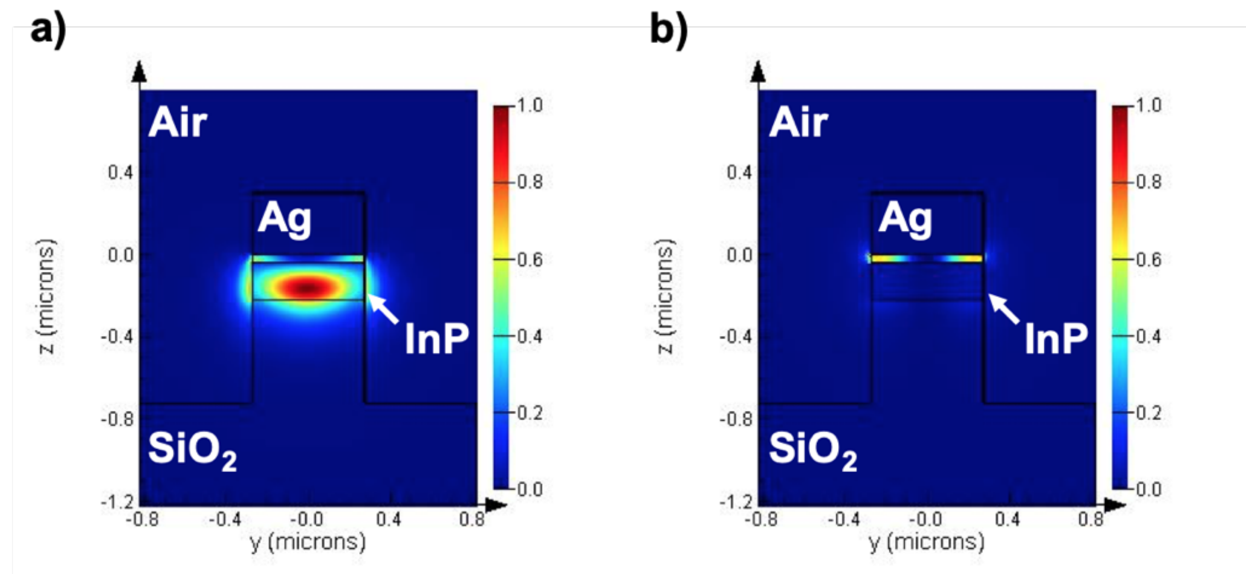

Figure 4.4: Available modes in the metal coupler region. a) Waveguide mode that couples well to InP waveguide and b) waveguide mode that does not couple well to InP waveguide

guarantee that we are coupling to the correct mode. In Fig. 4.4, we show two potential modes that exist under the metal in the waveguide coupler section; however, only Fig. 4.4a will efficiently couple to the InP waveguide after the metal is removed.

From Maxwell's equations we know that at the boundary between a perfect conductor and dielectric the only electric field component that can exist is the normal component - the tangential component must be zero. As shown in Fig. 4.4b, our undesired modes have an electric field normal to the top metal contact and near the edge of the waveguide. We found that by wrapping metal around the waveguide sidewalls and tapering the waveguide width down to 300nm we can cancel the undesired mode because the electric field becomes tangential to the waveguide sidewall. We can now calculate the modes for the new structure and see that only one mode can propagate under the metal, shown in Fig. 4.5.

While the overlap integral between this mode and the output waveguide is lower than the mode from Fig. 4.4, it is the only mode able to propagate under the metal, so the waveguide coupling efficiency is increased. Because the coupler section and output waveguide are both single mode, the power that is not coupled is reflected back to the coupler section, creating a fabry-perot resonator. The results are shown in Fig. 4.1(c), where we see the improvement in coupling to the fundamental mode by tapering the waveguide near the antenna-LED and wrapping metal around the sidewall of the tapered section. Figure 4.6(a) shows the perspective view and Fig. 4.6(b) shows the enhancement, antenna efficiency, and waveguide coupling efficiency spectra. With this structure we were able to achieve an average enhancement of 162, a waveguide coupling efficiency of 90%, and an antenna efficiency of

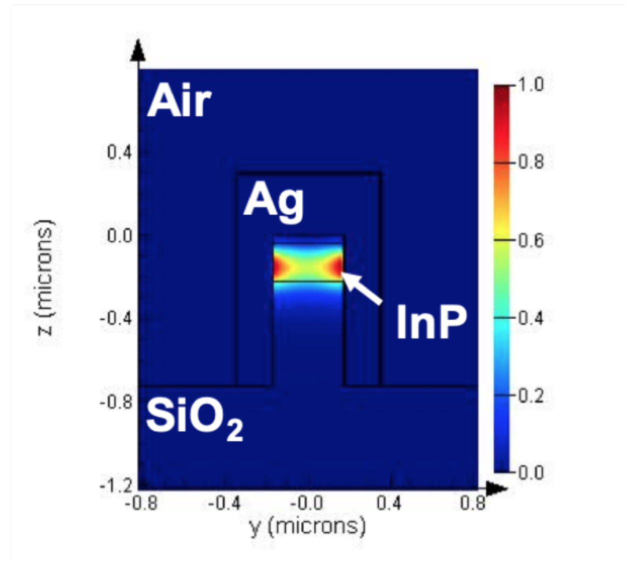

Figure 4.5: Mode in waveguide when metal is wrapped around the waveguide sidewalls. This mode is the only mode that exists at this wavelength.

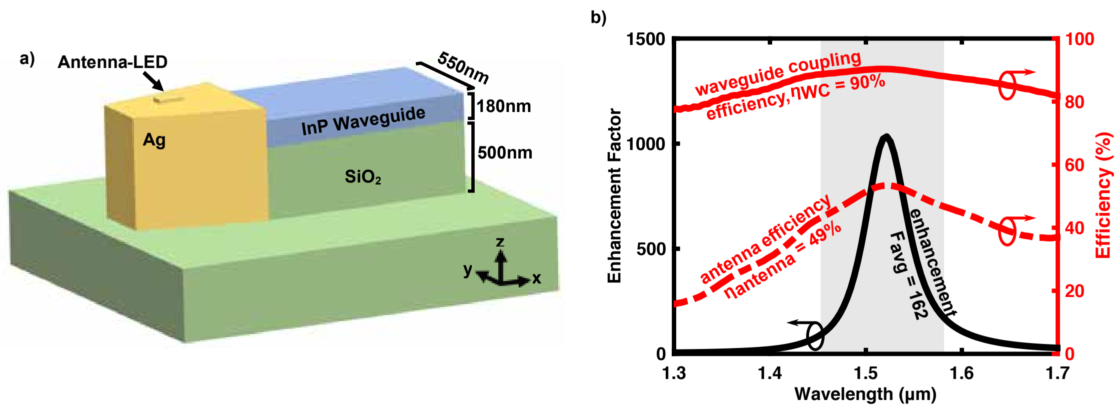

Figure 4.6: (a) Perspective view of tapered waveguide coupler with a waveguide height of 180nm and width of 550nm on a 500nm tall SiO_2 ridge, and (b) enhancement, antenna efficiency, and waveguide coupling efficiency spectra.

49%. The electric field plots for the hand design structures are summarized in Fig. 4.7.

Although our hand-optimized results are comparable to the best results in the literature, we were restricted to exploring only simple geometries of the waveguide coupler due to the immense computational resource requirements of simulating fine-meshed three-dimensional optical structures. In order to more completely explore the parameter space associated with this waveguide coupler, we applied computational inverse design techniques.

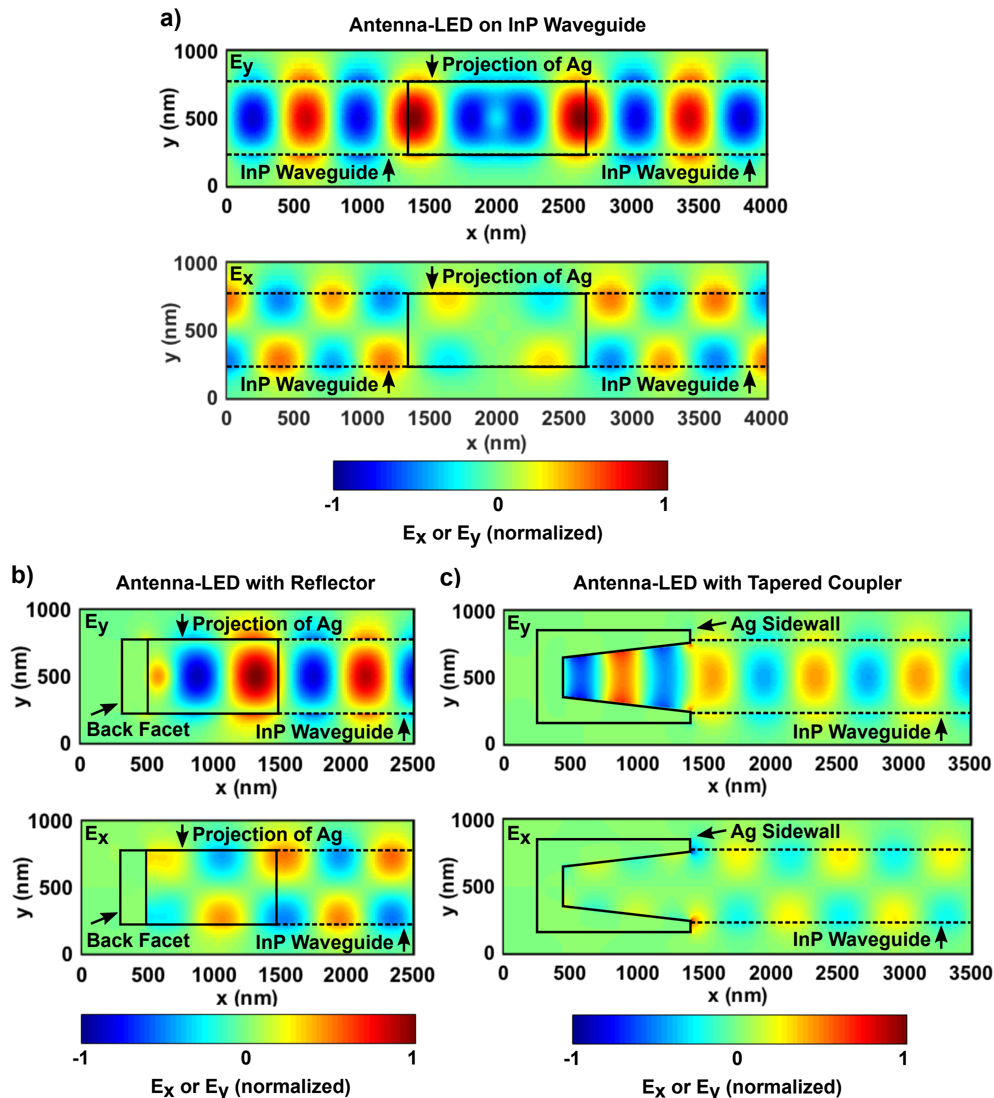

Figure 4.7: E_x and E_y field profiles for (a) antenna-LED on single-mode InP waveguide and SiO_2 ridge, (b) antenna-LED on single-mode InP waveguide with metal wrapped around waveguide facet, and (c) antenna-LED on single-mode InP tapered waveguide with metal wrapped around waveguide facet and sidewalls.

4.6 Inverse design

Electromagnetic inverse design has been used to improve characteristics of a multitude of photonic devices [37, 38, 39, 40, 41, 42, 43, 44, 45]. Inverse design methods allow one to efficiently find non-intuitive geometric structures that optimize electromagnetic figures of merit. For example, inverse design has been used to find high efficiency vertical grating couplers [37], to design a small footprint polarization beam splitter [38], to optimize a broadband two-mode de-multiplexer [39], to increase the near-field enhancement of an optical antenna while minimizing temperature rise [40], and to optimize fabrication-constrained silicon photonic devices [41].

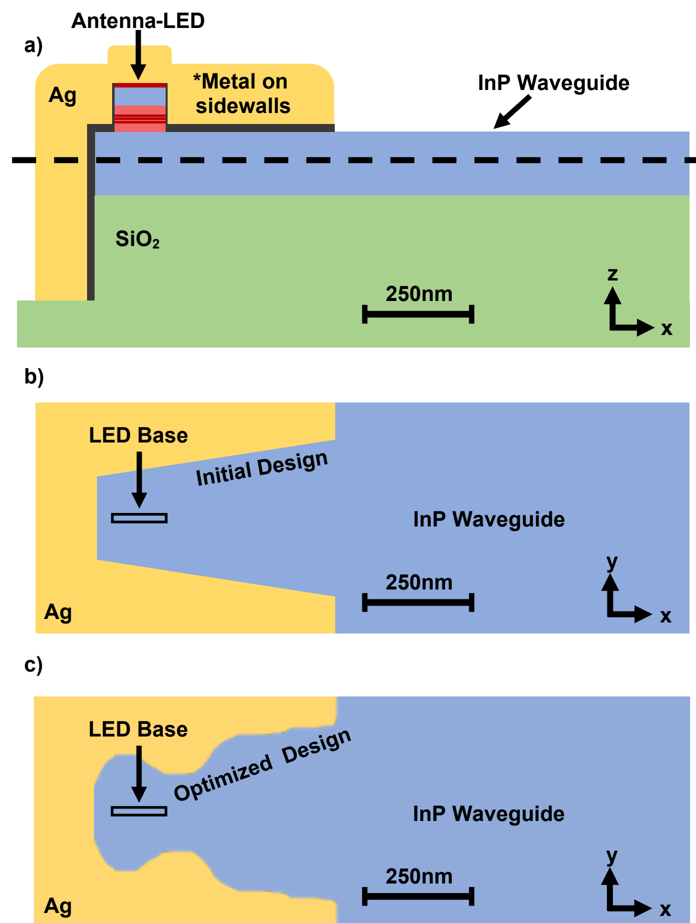

Figure 4.8: (a) Cross section schematic (XZ) of tapered waveguide coupler showing dashed cutline, and (b) top view XY cross section of waveguide along dashed cutline. (c) XY cross section of coupler after optimization, showing perturbations to Ag-InP boundary. Note (b) and (c) also show the projection of the LED base.

Gradient-descent based optimization using the adjoint method can be used to optimize almost any user-defined electromagnetic figure of merit over an arbitrarily large parameter space with minimal computational resource requirements [42, 43]. In the literature this optimization method and similar topology optimization methods are commonly referred to by the more general term inverse design, which we will adopt in order to help easily distinguish the various results in this dissertation. For brevity we will not delve into the details of the method, but we recommend the reader review the works in [42, 43, 41, 45, 44] for more information. See *Appendix C* for specifics regarding our implementation of inverse design.

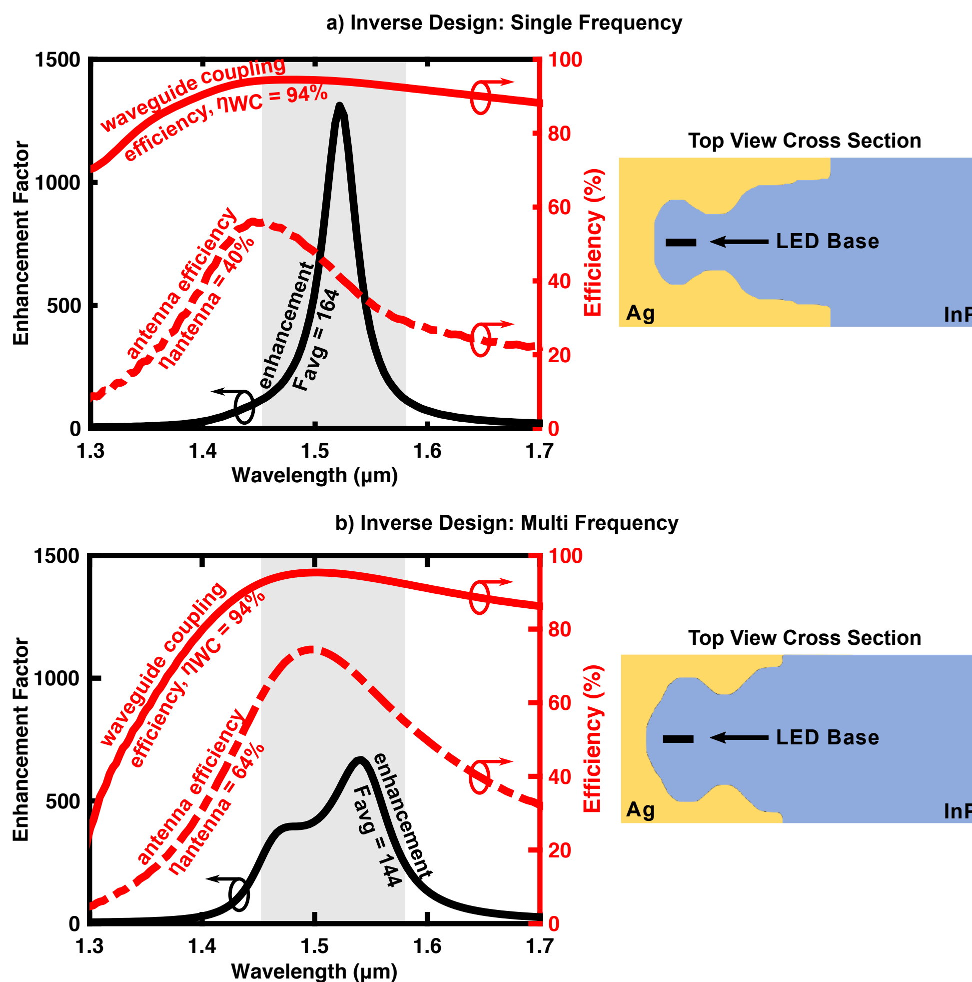

Figure 4.9: Enhancement, antenna efficiency, waveguide coupling efficiency spectra and top view XY cross sections for (a) single frequency optimization and (b) multi frequency optimization. For reference, the LED material spectrum $[L(\omega)]$ between its 50% power points is shown by the gray shaded region.

Inverse design was applied to the 2D cross section of the tapered coupler (Fig. 4.8(b)) to optimize enhancement, antenna efficiency, and waveguide coupling efficiency by perturbing the interface between InP and Ag (Fig. 4.8(c)) – the spectra before optimization are shown in Fig. 4.6(b). Our initial inverse design cost function was the power transmitted through the waveguide at a single frequency (spectral product at the resonant frequency of enhancement, antenna efficiency, and waveguide coupling efficiency). This led to a slight improvement in the power transmitted at resonance compared to the tapered coupler – shown in Fig. 4.9(a) and Fig. 4.6(b), respectively. However, when calculating the average values, we noticed there was a large trade-off between average enhancement and antenna efficiency. When compared to the tapered coupler, although the peak enhancement increased from 1034 to 1312, the average enhancement only increased from 162 to 164 and the antenna efficiency dropped from 49% to 40%. Waveguide coupling increased slightly from 90% to 94%. When we combine these numbers, we see that the average power of the single frequency optimization was lower than the tapered coupler. This is not surprising because the cost function did not represent an average value.

In order to increase the average power transmitted, we changed the inverse design cost function to be the weighted sum of the optical power at three frequencies. We weighted the power transmitted at resonance ten times less than the power transmitted at ±55THz (±40nm) from resonance to encourage a broader enhancement spectrum. As shown in Fig. 4.9(b), we were able to create a broader enhancement spectrum with a greater antenna efficiency, ultimately achieving $F_{\text{avg}} = 144$ and $\eta_{\text{antenna}} = 64\%$. A summary of the different designs is shown below in Table 4.1 - we included the antenna on bulk InP as a comparison for F_{avg} and $\eta_{antenna}$.

Design Summary				
Design	F_{avg}	$\eta_{antenna}$	η_{WC}	$\eta_{antenna} \times \eta_{WC}$
Antenna on bulk InP	160	54%	–	–
Taper	162	49%	90%	44%
Inverse Design: Single Frequency	164	40%	94%	38%
Inverse Design: Multi Frequency	144	64%	94%	60%

Table 4.1: Summary of best design results for antenna on bulk InP, taper coupler, and inverse design structures designed at a single frequency and multiple frequencies, respectively. These structures are for a width of 20nm and an Al_2O_3 thickness of 1nm.

As discussed in Chapter 2.3, the 3dB frequency is proportional to F_{avg} and the power in the fundamental mode is proportional to $F_{avg} \times \eta_{antenna} \times \eta_{WC}$. We accordingly expect the

Single Frequency design to be the fastest device and the Multiple Frequency design to be the most efficient with the highest output power.

4.7 Coupled cavity

Our design methodology is contingent on the LED material spectrum, shown in Chapter 2, Fig. 2.3. Given a narrower material spectrum, the single frequency design could be more desirable because the average enhancement would be much larger than the multi frequency design or tapered coupler. Even with our current material spectrum, the single frequency design will theoretically have the fastest direct modulation rate – though at a great expense to antenna efficiency. In contrast, the multi frequency design will have a slower direct modulation rate, but maintain high enhancement while achieving the highest efficiency, making it capable of delivering the most optical power to the waveguide. In fact, when we compare the product of F_{avg}, η_{antenna}, and η_{WC} from the multi frequency design with the cavity-backed slot antenna on a bulk InP substrate, we find that we could emit slightly more power in the fundamental mode of an InP waveguide than would be radiated in all directions for the bulk InP substrate case.

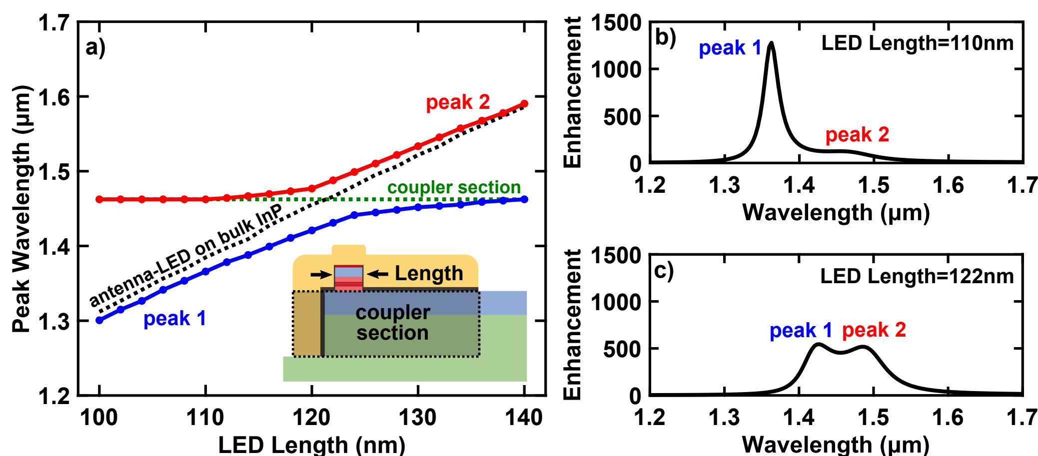

Figure 4.10: (a) Avoided crossing between the optical antenna resonance and the inverse design coupler resonance. For reference, dashed black and green lines show independent resonances of the antenna-LED on a bulk InP substrate and the coupler section as a function of LED length, respectively. Enhancement spectra for LED lengths of (b) 110nm and (c) 122nm.

Close observation of the multi frequency design enhancement spectrum in Fig. 4.9(b) reveals two distinct peaks. This can be explained by thinking of the antenna-LED and coupler section (see inset Fig. 4.10(a)) as coupled resonators. When they have the same resonance frequency, it will lead to a frequency split that can be observed in the enhancement

spectra. This was confirmed by sweeping the LED length in the multi frequency design, which resulted in an avoided crossing between the antenna-LED resonance and the coupler section resonance, as shown in Fig. 4.10. The dashed black line was generated by sweeping the length of the antenna-LED on a bulk InP substrate. The dashed green line was created by placing a dipole in the coupler section (see inset) and sweeping the length of an off-resonance antenna-LED. During the length sweep we found that the antenna efficiency always peaked at the coupler section resonance rather than at the antenna-LED resonance.

A similar observation was made in the single frequency design in Fig. 4.9(a): the antenna efficiency peak was associated with the coupler resonance. However, in contrast to the multi frequency design, the antenna-LED and coupler section resonances are detuned – evident by the offset between the peak enhancement and antenna efficiency wavelengths in Fig. 4.9(a).

To summarize, the spectra of the waveguide coupling designs can be explained by considering the antenna-LED and the coupler section as coupled resonators. When the resonances are tuned (multi frequency design), we have an impedance match and frequency splitting. Due to the impedance match, the optical power is able to quickly leave the lossy antenna-LED (lower Q factor) resulting in less metal loss (higher antenna efficiency). In contrast, when the resonances are detuned (single frequency design), we have an impedance mismatch which results in the optical power reflecting back to the lossy antenna region. This results in more metal loss (lower antenna efficiency) and higher enhancement. A similar conclusion was reached in [46], where detuned resonators were exploited to achieve higher peak enhancement. Note that regardless of how the coupler section resonance was tuned, both these designs yielded higher waveguide coupling efficiency than the tapered coupler.

4.8 Secondary design considerations

So far we identified the key parameters that affect the waveguide coupling efficiency, namely: waveguide height, lower cladding index/geometry, Al_2O_3 thickness, and metal geometry; however, there are a few secondary parameters that play a smaller role on waveguide coupling efficiency, but are nonetheless important. The three parameters we will focus on are the isolation oxides, taper geometry, and the waveguide width.

Isolation oxide

The oxide between the antenna and substrate is an important design parameter. It has been shown for a cavity backed slot antenna on bulk InP that adding a low index material around the LED ridge at the opening of the slot can increase the spontaneous emission factor (product of $F(\omega)$ and $\eta_{antenna}$) by $2\times$ due to a reduction of effective index and effective width [28]. We can see this trend in the middle and bottom graphs in 4.11. As we increase the thickness of the oxide, both the enhancement and antenna efficiency increase slightly. When combined, this only increases the peak spontaneous emission enhancement by 1.25x. In addition to the increase in the antenna efficiency we see a red-shift in the resonance. The

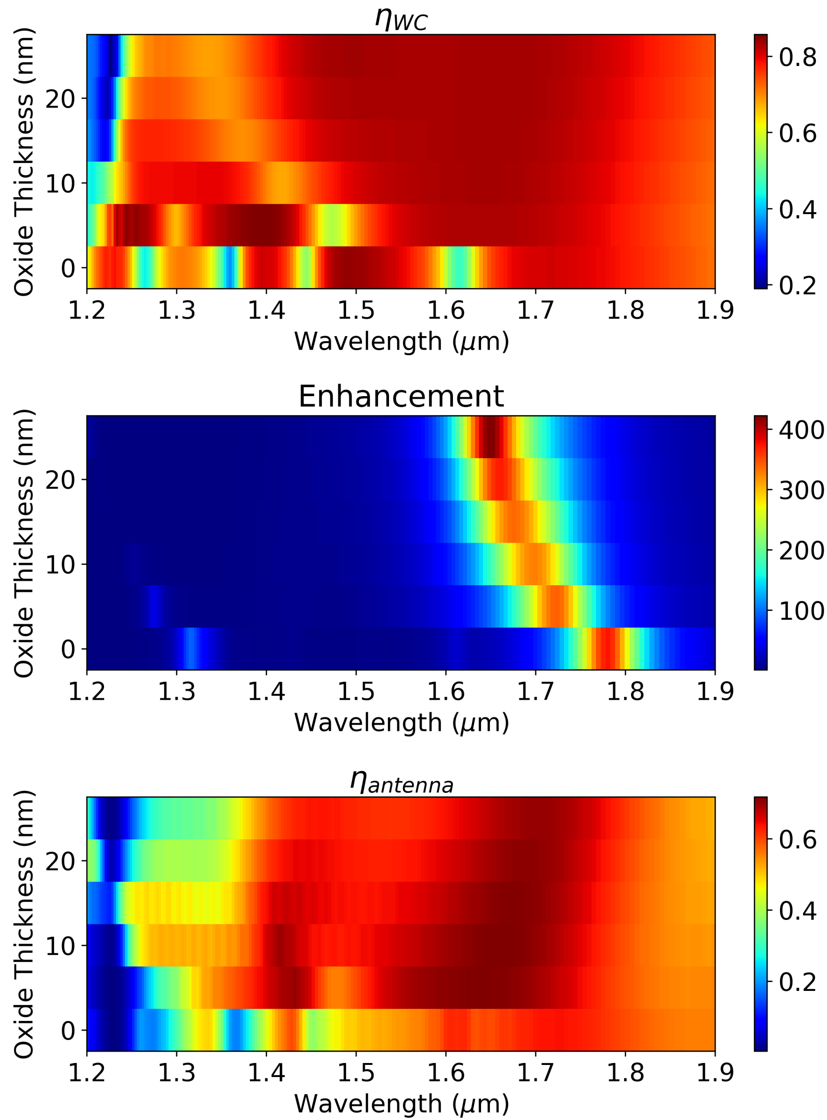

Figure 4.11: Waveguide coupling efficiency, enhancement, and antenna efficiency as a function of HSQ thickness.

red-shift can be understood by increasing the effective index at the opening - so the resonance red-shifts to compensate.

At the extreme of no spin-on-glass oxide (oxide thickness of 0) we see multiple resonances in the waveguide coupling efficiency and antenna efficiency. This can be understood as exciting higher order plasmonic modes in the coupler section. As we increase the oxide thickness above $\approx 5 - 10nm$, these are suppressed for our wavelengths of interest. This means we should not completely remove the spin on glass layer, and for thicknesses $\geq 15nm$ the separation oxide does not play a large role on the waveguide coupling efficiency.

In addition to higher spontaneous emission enhancement, the oxide is important to insure electrical isolation between the p-doped substrate and the antenna, which serves as the n-contact. Likewise, this oxide will be important to reduce the parasitic capacitance. So far the simulations have been for the oxide on top of the waveguide, but in a real device we would need an oxide on the sidewall of the waveguide to make sure it is electrically isolated. In our simulations this had a very small impact on the waveguide coupling efficiency and can be compensated with a smaller taper profile.

Waveguide width

Another parameter worth mentioning is the waveguide width after the coupler section. In all the designs so far we show a discontinuity in the waveguide which is designed to maximize the power transmission. However, the waveguide coupling efficiency is fairly insensitive to this parameter as long as the output waveguide is around single mode waveguide dimensions. If the output waveguide remains single mode then we should not have a problem coupling to higher order modes, and the waveguide width only affects the overlap integral between the coupler section and output waveguide, which determines the amount of reflection at this interface. A change in width ±100nm does not significantly change the design, allowing for some flexibility in fabrication and design. The waveguide can then be adiabatically tapered out to larger dimensions for lower waveguide loss.

Taper profile

An advantage of this coupler design is that it is largely insensitive to changes in the geometry. And while we showed a benefit in matching the resonance of the coupler section to the antenna-LED resonance, the device can operate in a non-resonant mode and achieve high waveguide coupling efficiency. This allows us to make changes like adding an oxide on the sidewall or perturbing the shape in inverse design without affecting the waveguide coupling efficiency dramatically.

We explored the fabrication tolerance of the structure and found that one important parameter is the metal offset. As shown in Fig. 4.12a, our initial design had an abrupt change in the width from 500nm to 700nm, but during metal liftoff it is difficult to control this offset within a few hundred nanometers. With offsets away from the abrupt change in width (negative offsets) the waveguide coupling efficiency only changes slightly due to

changing the taper resonance, but is largely unaffected. However, as we increase the taper offset past the abrupt change in width (positive offsets), there is a dramatic affect on the waveguide coupling efficiency. This can be mitigated by having a constant taper, as shown in Fig. 4.12b. Now, for small changes in the metal offset there is a minimal impact in the waveguide coupling efficiency, making the device more fabrication tolerant.

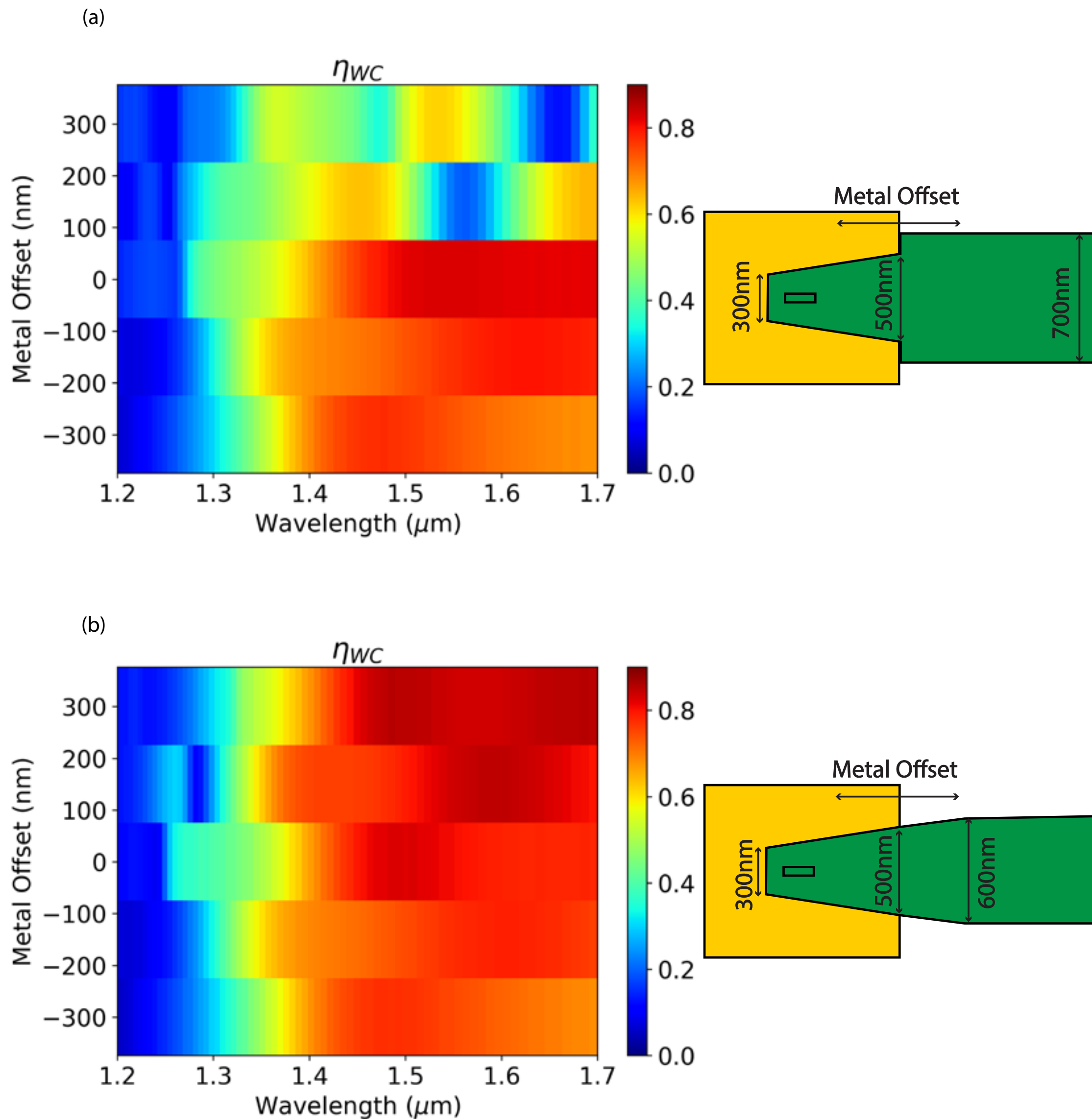

Figure 4.12: Comparison of waveguide coupling efficiency for (a) abrupt taper coupler and (b) continual taper coupler with top down cross section.

4.9 Towards electrical injection

As mentioned in the previous chapters, one constraint on these devices is that they need to be capable of electrical injection. In this section we will discuss a potential method for achieving electrical injection with the tapered geometry.

Optical considerations

The cavity backed slot antenna would still serve as the n-contact for the device; however, in order to inject holes we need a p-contact. A potential way of achieving this would be to use the metal wrapped around the p-InP waveguide as a contact, but this was dismissed as a possibility due to the difficultly in fabrication and because the waveguide sidewall would likely have high contact resistance. Another possibility considered was to directly contact the p-InP waveguide with metal, but this poses problems with disturbing the waveguide mode. Accordingly, taking some influence from ridge waveguide modulators, we want to place the metal sufficiently far away from the main waveguide mode. This can be accomplished in two ways:

1. **Rib-to-ridge conversion.** As discussed in previous sections, etching into the lower cladding can significantly increase the waveguide coupling efficiency, so we would still want to use an InP rib waveguide on top of a ridge-like lower cladding for the coupler section to maintain high waveguide coupling efficiency. In order to contact the waveguide outside the mode, we could convert the mode in the InP rib waveguide to a ridge waveguide mode, then contact the waveguide in the slab layer. In this geometry we quickly see a trade-off in slab thickness – thicker slabs will have a lower mode overlap between rib and ridge waveguide leading to a higher conversion loss, but the thicker slabs will have a lower series resistance.

2. **Tapered waveguide.** This solution involves an adiabatic taper to wider dimensions which will lead to greater confinement of the fundamental mode. The wider waveguide can then be contacted with perpendicular tapered waveguide that will not disturb the fundamental mode. This structure has the advantage of being fabricated in the same mask as the main waveguide, but gives rise to a similar trade-off to the rib-to-ridge conversion. Larger tapered waveguide tips will affect the fundamental mode more, but will have lower series resistance.

As shown in Fig. 4.13, we decided to design around the latter structure. As expected, there is a trade-off between the taper dimensions and the transmission. In Fig. 4.13b and 4.13c we compare the transmission efficiency for different taper geometries when the bus waveguide is 1000nm and 1500nm wide, respectively. As expected, for the 1500nm waveguide the fundamental mode is more confined to the center of the waveguide which allows for higher transmission for all taper geometries ($\geq 98\%$ transmission through this section).

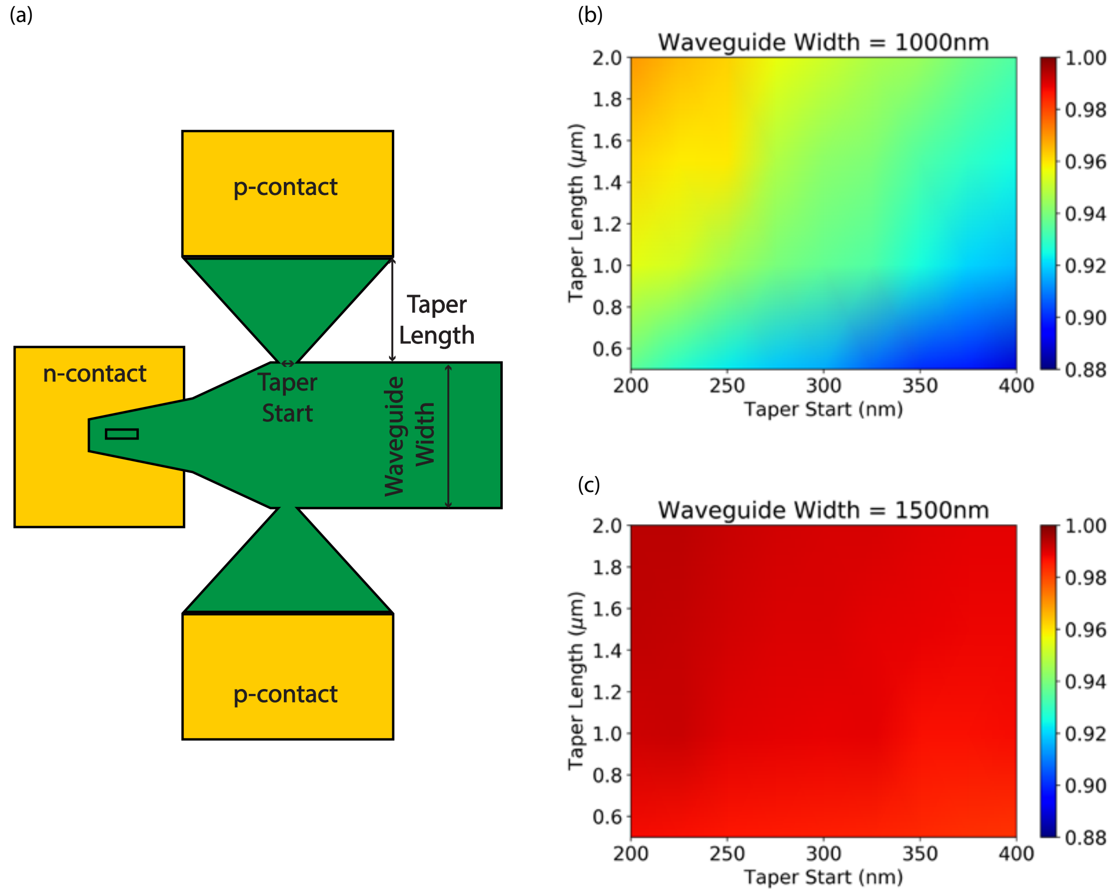

Figure 4.13: (a) Top down geometry of electrically injected waveguide, and optical transmission efficiency in taper section for different taper geometries for (b) 1000nm wide waveguide and (c) 1500nm wide waveguide.

One final consideration is that doping the InP waveguide will lead to higher waveguide loss (≈2dB/cm vs ≈0.3dB/cm in a Zn-doped waveguide compared to intrinsic) [47]. If the platform is compatible with Si photonics it would help to reduce link loss by coupling the light to a lower loss waveguide (Si or SiN) or only doping near the taper antenna sections. Assuming these losses, for links on the order of $1cm$ the waveguide will be practically lossless, so we can ignore the length dependence on the link energy.

Electrical considerations

The final consideration with electrical injection is the loss associated with the contacts and the series resistance. The important contact resistance is for the n+InGaAs - using

a nanoscale Mo contact this can be as low as $\rho_c = 1.3\Omega\mu m^2$ [48]. For a ridge with a 20nm width and 130nm length the contact resistance is on the order of:

$$R_{contact} = \frac{1.3}{0.13 * 0.02} = 500\Omega \tag{4.3}$$

The series resistance of the p-type waveguide is slightly more complicated. We need to calculate the series resistance of the various tapers:

$$R_{taper} = \int_0^L \frac{\rho}{h * (w_i + \frac{x}{L}(w_f - w_i))} dx \tag{4.4}$$

$$R_{taper} = \rho \frac{L}{h(w_f - w_i)}) \log \frac{w_f}{w_i} \tag{4.5}$$

We can double check this is correct by doing a change of variable $\delta w = w_f - w_i$ and setting to $\delta w = 0$ (i.e. not having a taper):

$$R_{taper} = \rho \frac{L}{h\delta w} \log 1 + \frac{\delta w}{w_i} \tag{4.6}$$

Using L'Hospital's Rule we get:

$$R_{taper} = \lim_{\delta w \to 0} \rho \frac{L}{h} \frac{\frac{1}{w_i}}{\frac{\delta w}{w_i} + 1} = \rho \frac{L}{h \times w} \tag{4.7}$$

This is the normal series resistance of a rectangular wire. Finally, to calculate the resistance for this structure we need the resistivity for p-doped InP:

$$\rho \approx \frac{1}{q\mu_p N_a} \tag{4.8}$$

where μ_p is the hole mobility and N_a is the p-doping. The hole mobility as a function of p-doping is approximately [49]:

$$\mu_p = 10 + \frac{160}{1 + (\frac{N_a}{4.87*10^{17}})^{0.62}} \frac{cm^2}{Vs} \tag{4.9}$$

At a doping of $N_a = 1 \times 10^{19} cm^{-3}$ the resistivity is $\rho = 2 \times 10^{-2}\Omega cm$. We can use Eq. 4.5 to estimate the series resistance for different sections, a summary of these are given below in Table 4.2.

The final resistance we need to consider is the LED ridge itself. We can again assume a p-doping of the ridge of $N_a = 1 \times 10^{19} cm^{-3}$. Accordingly, the series resistance of the ridge is:

$$R_{ridge} = \rho \frac{200nm}{130nm \times 20nm} \approx 15339\Omega \tag{4.10}$$

Series Resistance				
Section	Length (nm)	$W_{initial}$ (nm)	W_{final} (nm)	Resistance (Ω)
Coupler Section	1000	300	700	2112
Waveguide Taper	750	700	1000	889
Contact Taper	1500	3000	200	1446

Table 4.2: Calculation of series resistance for p-doped InP waveguide sections assuming a doping of $N_a = 1 \times 10^{19} cm^{-3}$ and waveguide height of 200nm.

So the total series resistance, assuming a negligible p-contact resistance and series resistance in the n-doped InP ridge, is:

$$R_{total} = R_{ridge} + R_{contact} + R_{coupler\ section} + R_{waveguide\ taper} + 0.5 \times R_{contact\ taper} \tag{4.11}$$

The total series resistance is $\approx 19k\Omega$, mostly dominated by the LED ridge at $15.3k\Omega$. The waveguide taper sections only contributes $3.7k\Omega$, and the remaining resistance comes from the n-contact.

In this chapter, we have demonstrated that the cavity-backed slot antenna-LED can be efficiently coupled to a single-mode waveguide, which was validated using relevant figures of merit in an optical interconnect. Then, using inverse design we further optimized the cavity-backed slot antenna coupling, ultimately achieving a waveguide coupling efficiency of 94%, antenna efficiency of 64%, while maintaining a high average enhancement of 144. We found that inverse design was able to achieve these results by tuning the optical resonance of the coupler section relative to the antenna-LED based on our cost function.

Due to its high efficiency, nanoscale size, compatibility with electrical injection and top-down fabrication, and speed, the cavity-backed slot antenna-LED is a very promising transmitter for an on-chip optical interconnect.

Chapter 5

Optically pumped waveguide coupled cavity backed slot antenna LED

In the previous chapter we showed the theory and simulations for efficient single mode waveguide coupling of the cavity-backed slot antenna. In this chapter we will discuss the fabrication and measurements of the waveguide coupled device.

5.1 Fabrication

Epitaxial layer design

We start with a metal organic chemical vapor deposition (MOCVD) grown epitaxial wafer shown in Fig. 5.1. The structure is an undoped double heterostructure (DH) with an InGaAsP active region and InP confinement layers. Normally, it is preferable to avoid quarternary alloys if possible due to a more complicated growth that can introduce more defects and yield a lower quality device; however, because we use a cooled InGaAs detector, it will blue-shift the absorption spectra leading to a poor responsivity at the peak emission wavelengths for a simpler InP lattice matched $In_{0.53}Ga_{0.47}As$ active region.

To have a sufficient overlap with the detector responsivity, we designed the active region to have a bandgap of 0.826eV (1500nm). The lattice matching condition for $In_{1-x}Ga_xAs_{1-y}P_y$ on InP and its bandgap are given below [50].

$$x = \frac{0.1893(1-y)}{0.4050 + 0.0132y} \tag{5.1}$$

$$E_g = 0.75 + 0.46y + 0.14y^2 (eV) \tag{5.2}$$

So, in order to get our designed bandgap, we target an $In_{1-x}Ga_xAs_{1-y}P_y$ alloy with $x = 0.391$ and $y = 0.159$.

The top 20nm InP and 30nm InGaAs layers are cap layers to protect the active region during transportation and are removed prior to alignment mark processing. The

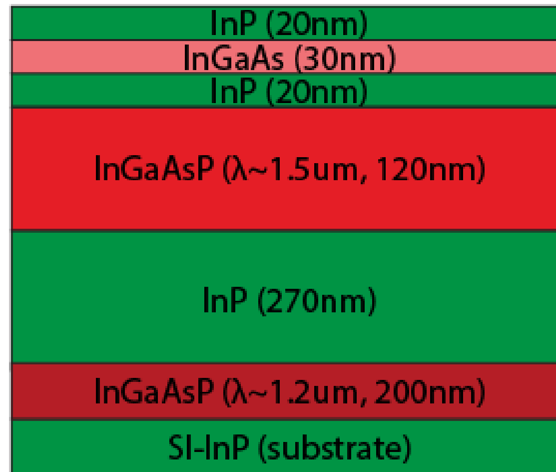

Figure 5.1: Waveguide coupled antenna-LED epitaxial layers

InP/InGaAsP/InP layers compose the active region, where the 270nm InP layer acts as both carrier confinement for the active region and the waveguide layer. Finally, the InGaAsP etch stop and InP substrate are removed to provide the index contrast for the InP waveguide.

Fabrication flow

The fabrication flow is shown in Fig. 5.2, with additional fabrication notes provided in the appendix. In order to pattern and align nanoscale features, all patterning was done using electron beam lithography in crestec CABL-UH 130kV system.

We start with the epitaxial wafer shown in Fig. 5.1, and top InP and InGaAs cap layers are removed. Next, a positive resist (PMMA 495 C4) is patterned with electron beam lithography (EBL) for electron-beam evaporated Cr/Au (8nm/92nm) alignment mark liftoff (Fig. 5.2a). Then, also in EBL, a negative resist hydrogen silsequioxane (HSQ) is patterned into a dry-etch hard mask for the nanoscale LED-ridge etch (Fig. 5.2b). The InP/InGaAsP ridges are then etched using a reactive ion etch inductively-coupled plasma (RIE-ICP) dry etch system with Ar/H2/CH4/Cl2 gas at room temperature (Figure 5.2c) - note this is a critical step because it will determine both the antenna-LED and waveguide heights, additionally it is a time-etch rather than a selective etch. After the hard mask removal in buffered oxide etch (5:1), the etch height is measured using atomic force microscopy (AFM) to confirm it is within specification before proceeding with the sample. The epi was designed with a target etch height of 210nm, which would create a waveguide height of 200nm after processing. If we underetch then the waveguide will be thicker and potentially support higher order modes; likewise, if we overetch then the waveguide could be too thin and no longer support a mode, so we target a thickness within $\pm$20nm.

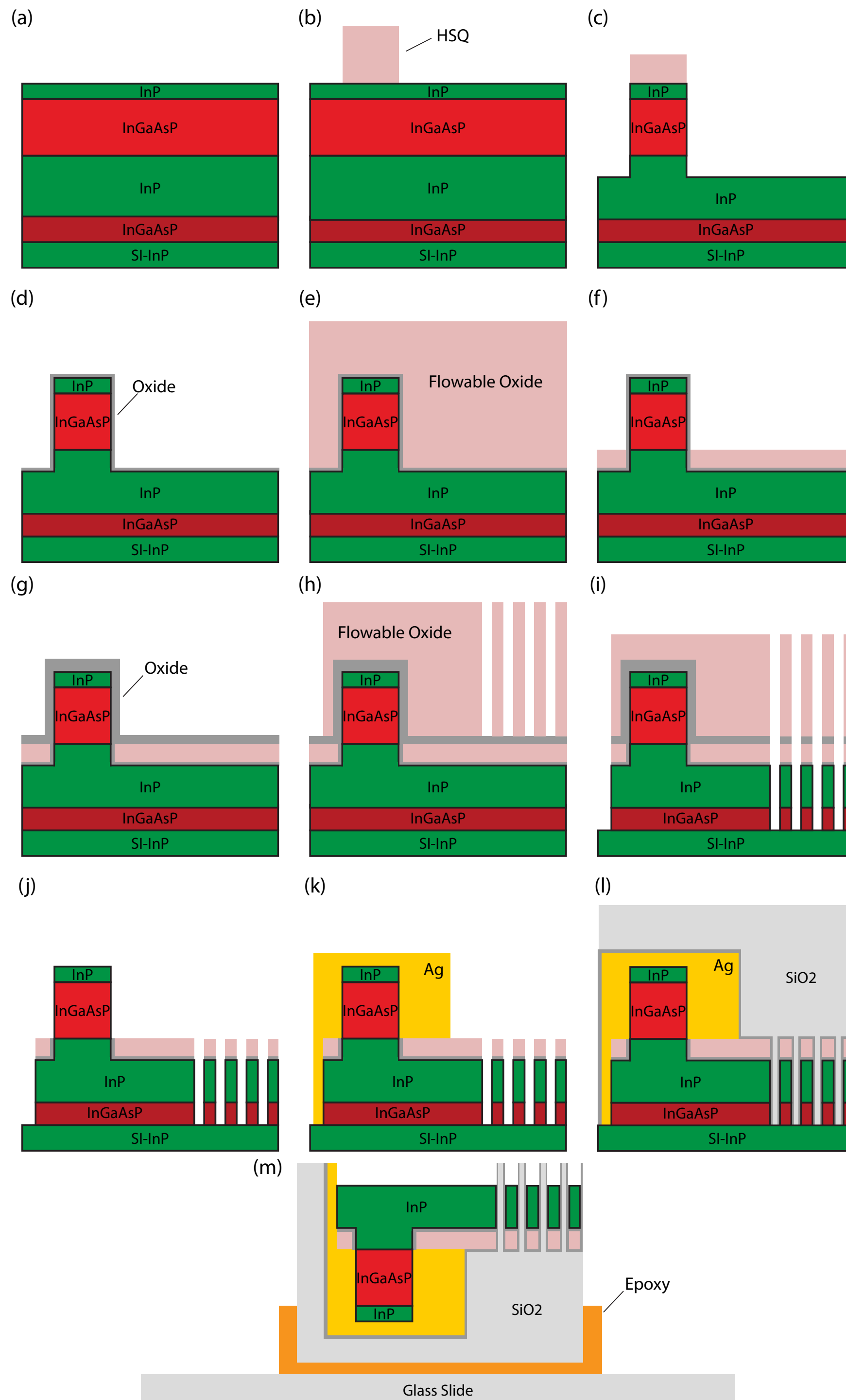

Figure 5.2: Abbreviated fabrication flow for LED ridge etch, antenna deposition, and substrate removal.

After confirming the etch is within our specifications the surface is cleaned with dilute tetramethylammonium hydroxide (TMAH) to remove native oxides [51] and buffered oxide etch (10:1) before an ammonium sulfide soak and atomic layer deposition of $Al_2O_3/TiO_2/Al_2O_3/TiO_2$ (3nm/4nm/10nm/4nm) (Fig. 5.2d). The inner Al_2O_3 will serve as electrical isolation from the metal antenna, while TiO_2 will act as a wet etch stop for the thicker outer Al_2O_3. This outer Al_2O_3 layer is used to protect the surface from degradation during additional plasma etches, and finally the outer TiO_2 acts as an adhesion layer for further processing.

The atomic layer deposition is followed by spinning a thick planarization flowable oxide (FOX), baking at 325C to crosslink the resist into spin-on-glass (Fig. 5.2e), and an isotropic reactive ion etch back using an SF6/O2 chemistry (Fig. 5.2f). Then the spin-on-glass is protected by a 10nm/4nm Al_2O_3/TiO_2 layer, where the Al_2O_3 will serve as a mask and the TiO_2 as an adhesion layer (Fig. 5.2g). Then we use a high contrast EBL process to pattern FOX resist into a hard mask for the waveguide etch step (Fig. 5.2h).

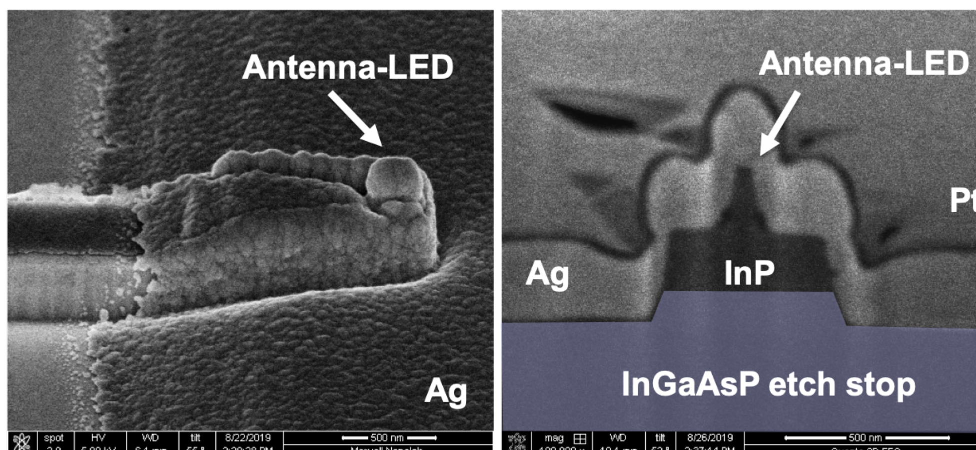

Figure 5.3: Left: scanning electron micrograph of tapered coupler after metal deposition. Right: focused ion beam cross section micrograph showing conformal coverage of silver.

The waveguide is etched through a series of selective wet and dry etches (Fig. 5.2i). First the TiO_2 is etched using an anisotropic RIE-ICP SF6/O2 plasma, the Al_2O_3 is etched using a dilute TMAH, the underlying spin on glass and TiO_2 is also etched using an anisotropic RIE-ICP SF6/O2 plasma, the remaining $Al_2O_3/TiO_2/Al_2O_3$ are selectively etched using dilute TMAH, 80C H_2O_2, and dilute TMAH, respectively. Finally, the InP waveguide is etched using the RIE-ICP Ar/H2/CH4/Cl2 gas chemistry. The remaining FOX mask is removed using a low-power anisotric SF6/O2 plasma RIE-ICP etch and the atomic layer deposition layers are selectively removed via wet etch to re-expose the surface of the InGaAsP LED (Fig. 5.2j). Then the surface is cleaned again with dilute TMAH and BOE 10:1, followed by another ammonium sulfide soak and a fresh Al_2O_3/TiO_2 (3nm/4nm) and a rapid thermal anneal and isotropic SF6/O2 to improve the surface. Then an LOR10A/PMMA 495 C4

bilayer is EBL patterned, the TiO_2 ALD is removed, and approximately 400nm of silver and 30nm niobium is sputtered. For more information on this process see the following subsection as well as the appendix (Fig. 5.2k). A scanning electron micrograph (SEM) and focused ion beam (FIB) SEM after this process is shown in 5.3. We can see from the FIB SEM that the sputtering silver step coverage on the sidewalls of the LED and the InP waveguide is sufficiently thicker than the optical skin depth to prevent light leakage.

Then we deposit mechanical oxides to keep the waveguides and grating in place during substrate removal, consisting of 10nm Al_2O_3 and 1μm room temperature electron cyclotron resonance chemical vapor deposition (ECR-CVD) (Fig. 5.2l). The sample is epoxy bonded to a glass slide using Norland Optical Adhesive 81 (NOA-81), ultraviolet (UV) and hotplate cured, and the substrate is removed by mechanical grinding down to approximately 100μm thickness then wet etched. Finally, the etch stop is selectively removed using a dilute piranha to complete the fabrication (Fig. 5.2m).

5.2 Silver sputtering or evaporation

One critical step of process development we would like to expand on is the silver deposition because it is important to have a low-loss metal to achieve high efficiency antenna efficiency. Likewise, the metal needs to conformal coverage of the LED-ridge to achieve a desirable antenna-LED resonance [28].

Electron-beam evaporation is regularly used to deposit high-quality and reasonably low-loss metal films for use in metal optics. However, it suffers from line of sight travel, making it a poor choice to cover sidewalls. This can be mitigated by an in-situ tilt stage, where the angle can be adjusted from 0-90°, which can effectively get conformal deposition along one axis of rotation [28]. But, for our waveguide coupled devices, it is desirable to have conformal metal deposition along the waveguide facet as well, which is perpendicular to the preferred deposition axis.

Sputtering is an alternative technique to deposit metal films, where a plasma is used to bombard a target. The operating pressure is several orders of magnitude higher than evaporation (≈1mTorr), but is still low enough to maintain line-of-sight deposition. Using this technique, more conformal coverage can be achieved by having a relatively large target with stage rotation. Additionally, high quality films have been demonstrated by optimizing the pressure and power [52].

Without an empirical test, it is difficult to predict which film would have lower loss. Because a plasma is required for sputtering, it operates at several orders of magnitude higher pressure than evaporation, which has the potential for more impurity or void incorporation in the film. On the other hand, sputtering can operate at much higher speeds than the tilt evaporation (especially in-situ tilt evaporation), which is known to decrease impurity concentration.

To determine the metal quality of evaporated and sputtered films we fabricated cavity-backed slot antenna LED test structures (height: 160nm, width: 60nm) with varying length

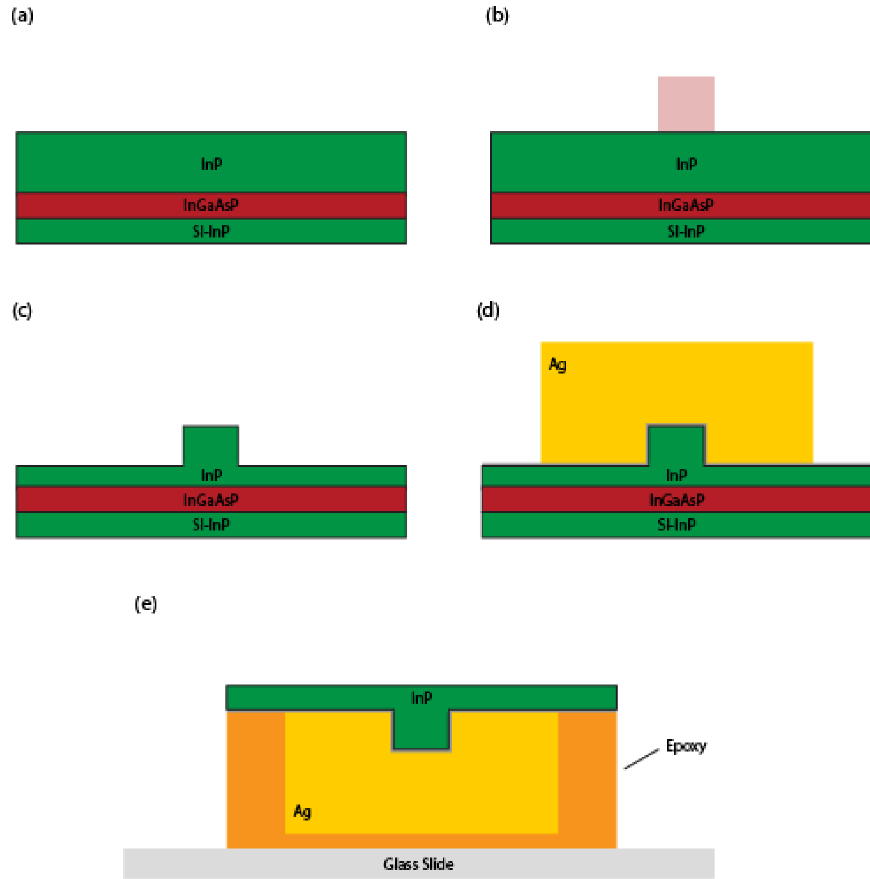

Figure 5.4: Fabrication flow

and measured the Q factor of the dark field scattering resonance, which corresponds to the metal quality [53, 54, 55, 56, 57, 58]. We confirmed that our design would be limited by the material quality factor with a total field-scattered field source in Lumerical FDTD, with results shown in Fig. 5.5. In this simulation we increase the imaginary part of the permittivity to simulate a higher loss film, and see a decrease in both scattering amplitude and Q.

The process flow for the metal test structures requires only a single mask and etch step. As shown in Fig. 5.4, the process flow is: (a) Wet etch past InGaAs active region; (b) EBL pattern HSQ ridges; (c) ICP etch InP ridge and atomic layer deposition of 3nm Al_2O_3; (d) Cleave samples and deposit Ag on sample halves (one half undergoes evaporation, the other sputtering); (e) Substrate + etch stop removal. The samples were processed at the same time to minimize fabrication variation that could affect the resonance.

The dark field scattering spectra is calculated using Eq. 5.3, it has three separate spectral measurements: 1) scattering from sample ($counts_{sample}(\omega)$), 2) scattering from flat metal

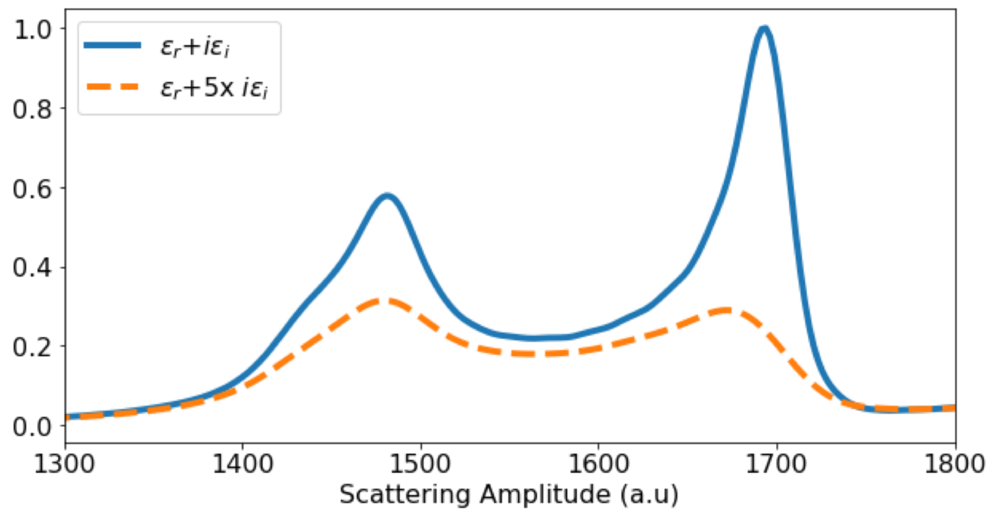

Figure 5.5: FDTD simulation using total-field-scattered-field source. Silver loss modeled by increasing imaginary part of the refractive index from [59] for InP ridge (length: 680nm width: 60nm).

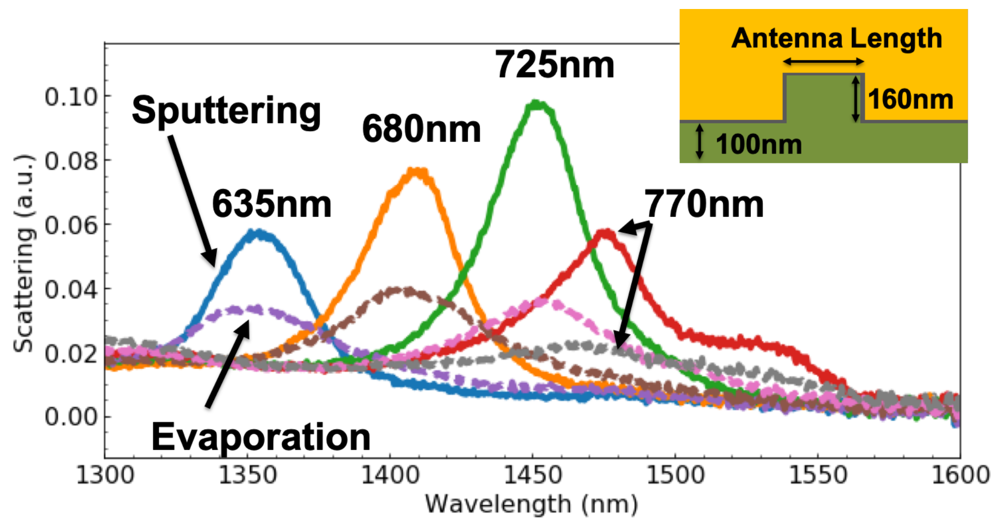

Figure 5.6: Comparison of sputtered silver (solid lines) and evaporated silver (dashed lines) for various antenna lengths for an antenna width of 60nm. Both samples show similar resonances for same antenna lengths.

region ($counts_{metal}(\omega)$) - i.e. this is the background measurement to remove the dependence on the planar metal, and 3) scattering from an efficient uniform scattering source to correct

for the illumination - we used simple white paper ($counts_{paper}(\omega)$).

$$Spectra(\omega) = \frac{counts_{sample}(\omega) - counts_{metal}(\omega)}{counts_{paper}(\omega)} \tag{5.3}$$

In Fig. 5.6, we show the dark field spectra for the silver evaporated sample (dashed lines) and sputtered sample (solid lines) for four different antenna lengths. As expected from antenna theory, longer antennas shift the resonance to longer wavelengths - likewise for the same antenna length we get the same resonance wavelength for both metals, a good confirmation that we minimized unintentional fabrication variations between the samples. For the 680nm long antenna the sputtered sample has a Q factor of 30.3, while the evaporated sample has a Q factor of 16.3. Additionally, the scattering intensity of the sputtered sample is higher, both these observations are consistent with a less lossy metal.

After it was determined sputtering provided a silver film of superior quality and coverage, we introduced it into our fabrication flow. As shown in Fig. 5.7, we achieved this with a PMMA and LOR bilayer to provide a sufficient re-entrant profile for liftoff [60].

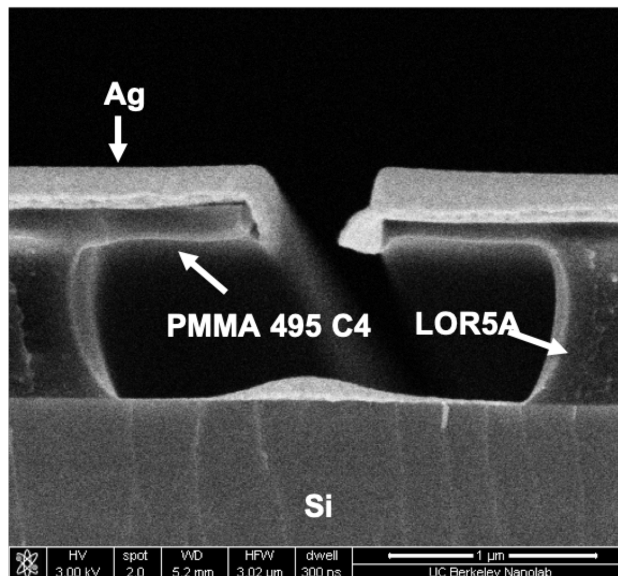

Figure 5.7: LOR and PMMA bilayer showing re-entrant profile after exposure, 60s PMMA develop using MIBK:IPA, and 30s LOR develop using OPD4262.

5.3 Measurement setup

After fabricating the device we can measure the spatial distribution of the counts, and assign the light to one of the emission areas in Fig. 5.8.

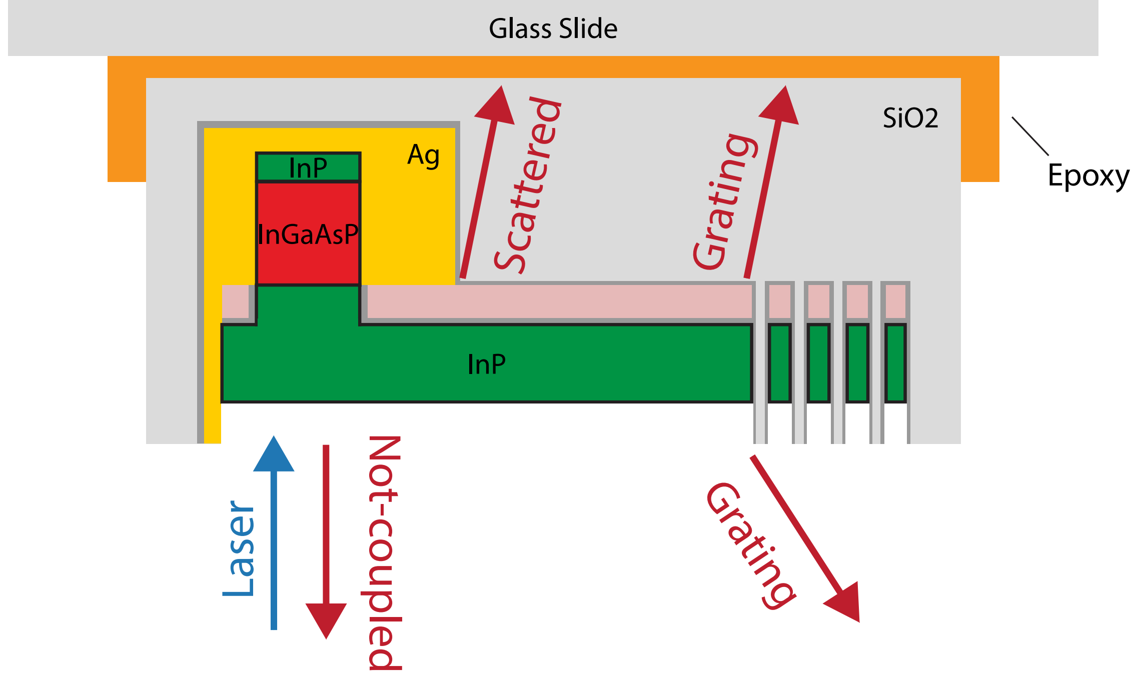

Figure 5.8: Schematic showing expected areas where photons are emitted from. "Not-coupled" is all the light that is immediately emitted from the coupler section; "scattered" is the light that scatters when going from the coupler section to single mode waveguide; and "grating" is the light that is directed up and down through the grating.

We will be approximating the waveguide coupling efficiency by taking the ratio of light coming from the grating to the total light collected. To distinguish this from the waveguide coupling efficiency we will be calling this experimental quantity $\eta_{WC,ratio}$. As shown in Fig. 5.9, only one side of the counts can be collected at a time (air-side or epoxy-side) - so the single side ratio is:

$$\eta_{WC,ratio}(\omega) = \frac{GratingCounts(\omega)}{GratingCounts(\omega) + NotCoupledCounts(\omega)} \tag{5.4}$$

where

$$GratingCounts(\omega) = \eta_{WC}(\omega)\eta_{grating,z+}(\omega)\eta_{grating,collection}(\omega)$$

$\eta_{WC}(\omega)$ is the actual waveguide coupling efficiency, $\eta_{grating,z+}(\omega)$ is the grating coupler efficiency radiating into the halfsphere, and $\eta_{grating,collection}$ is the microscope collection efficiency (NA = 0.5 for 50x objective).

$$NotCoupledCounts(\omega) = \eta_{NC}(\omega)\eta_{NC,collection}(\omega)$$

$\eta_{NC}(\omega)$ is ratio of light that is radiated from the coupler section in the lower half-sphere, and $\eta_{NC,collection}(\omega)$ is the microscope collection efficiency (NA = 0.5 for 50x objective).

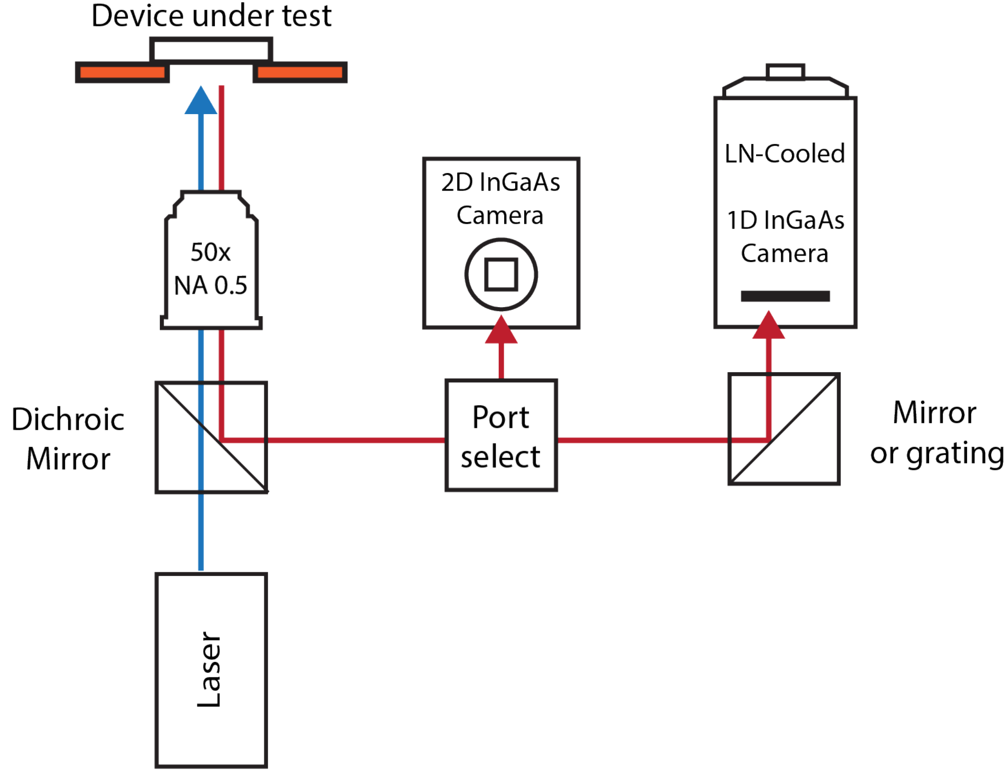

Figure 5.9: Schematic of waveguide coupling efficiency measurement. Device under test (DUT) is optically pumped with laser source at higher photon energy. Emitted light at lower photon energy is filtered with dichroic mirror and directed to 2D InGaAs camera or cooled 1D InGaAs array.

In order to understand how the experimental single-sided waveguide coupling ratio approximates the waveguide coupling efficiency we rewrite Eq. 5.4 as:

$$\eta_{WC,ratio}(\omega) = \frac{\eta_{WC}(\omega)\eta_{grating,z+}(\omega)\eta_{grating,collection}(\omega)}{\eta_{WC}(\omega)\eta_{grating,z+}(\omega)\eta_{grating,collection}(\omega) + \eta_{NC}(\omega)\eta_{NC,collection}(\omega)} \tag{5.5}$$

We can lump everything except for the η_{WC} in the numerator into a wavelength dependent coefficient:

$$\eta_{WC,ratio}(\omega) = A(\omega)\eta_{WC}(\omega) \tag{5.6}$$

where

$$A(\omega) = \frac{\eta_{grating,z+}(\omega)\eta_{grating,collection}}{\eta_{WC}(\omega)\eta_{grating,z+}(\omega)\eta_{grating,collection}(\omega) + \eta_{NC}(\omega)\eta_{NC,collection}(\omega)}$$

From Eq. 5.6, we clearly see that the waveguide coupling ratio would perfectly approximate the waveguide coupling efficiency when $A(\omega) = 1$, when $A(\omega) > 1$ the waveguide coupling ratio overestimates the waveguide coupling efficiency, and when $A(\omega) < 1$ it underestimates the waveguide coupling efficiency. This quantity is plotted in Fig. 5.10, and is slightly above 1 for most of our wavelengths of interest ($\lambda \approx 1400 - 1600nm$). Note that the simplification of $A(\omega)$ depends on the both the optical efficiencies of the setup (i.e. grating and collection efficiencies) as well as the waveguide coupling efficiency.

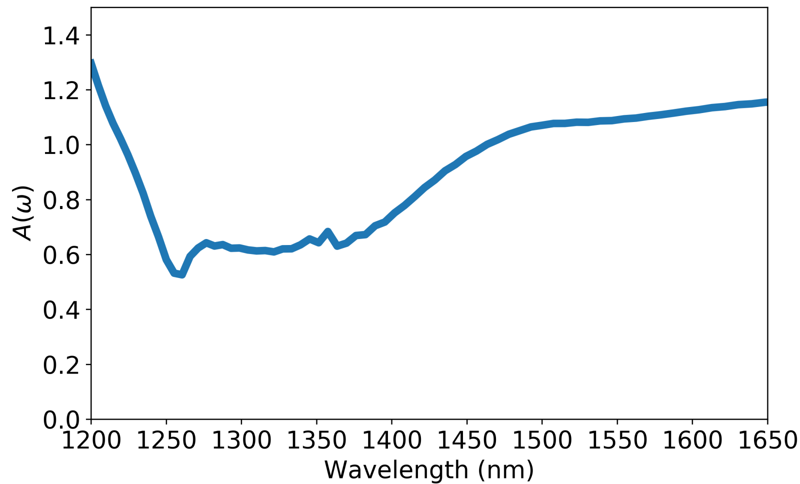

Figure 5.10: $A(\omega)$ factor plotted as a function of wavelength. $A(\omega)$ takes into account the grating efficiency and collection efficiencies as well as the waveguide coupling and not-coupled count efficiencies.

So far we have been considering the spectral dependence of the emission, this requires the light from the grating to be focused on the spectrometer aperture, then a separate measurement where the "not-coupled" counts are focused on the aperture. But, we can also do a single-shot measurement of the waveguide coupling efficiency by changing the spectrometer grating to a mirror. This essentially performs a weighted average over the emission spectra, and helps us distinguish between the grating emission and the "not-coupled" emission.

5.4 Results

In Fig. 5.11, we show an optical micrograph taken after substrate removal in the top image, and the dark field image with a heatmap of the counts overlaid from the 2D camera. In

order to get a more accurate measurement we use the LN-cooled 1D InGaAs detector - this measurement is overlaid as a solid blue line at the bottom of the figure. The rest of the data comes from the cooled 1D InGaAs detector.

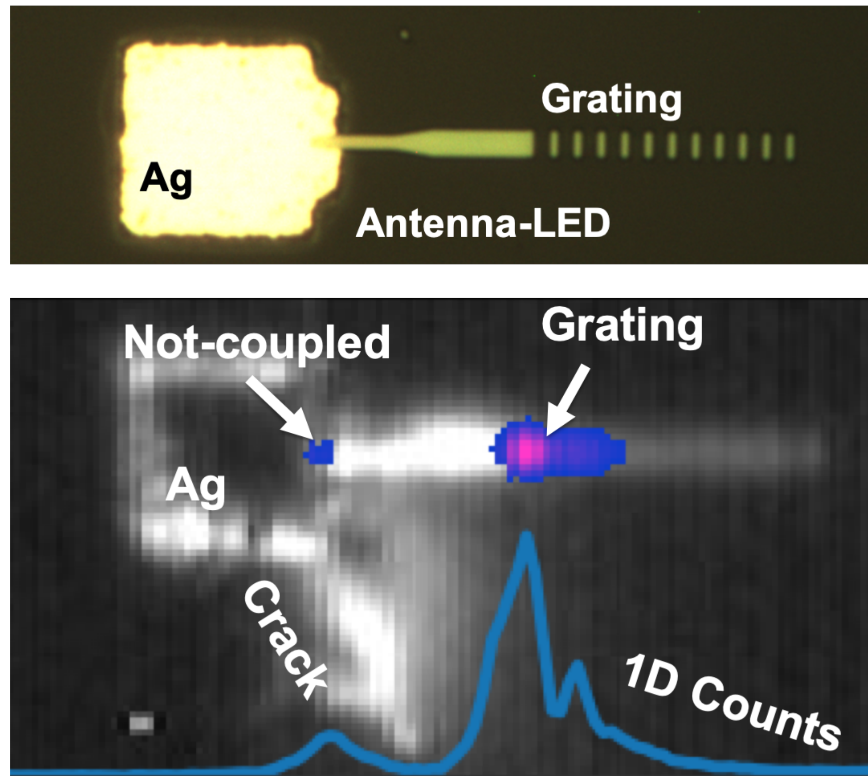

Figure 5.11: Top: Optical microscope image of fabricated device after substrate removal. Bottom: Dark field image overlaid with heatmap of emission and solid blue line showing counts along 1D pixel line. A crack near the sample was label, it does not propagate into the device or waveguide.

In Fig. 5.12 we calculate the waveguide coupling ratio using the air-side measurements of not-coupled counts and grating counts using a spatially integrated version of Eq. 5.4. From the air-side measure we calculate a waveguide coupling ratio of 85.9%. For these results we colored the counts that we included in not-coupled in blue and grating in orange.

$$\eta_{WC,ratio} = \frac{\sum_{pixel} GratingCounts(pixel)}{\sum_{pixel} GratingCounts(pixel) + NotCoupledCounts(pixel)}$$

From Fig. 5.12 we can now identify the not-coupled counts and grating counts easily. This allows us to position these counts in the center of the spectrometer aperture and close it down to only collect the grating counts or not-coupled counts. The light is passed through

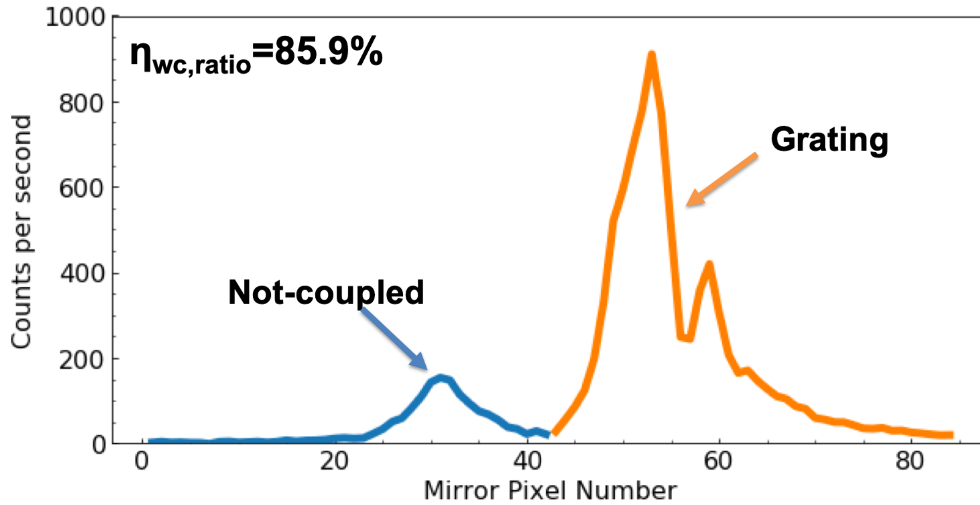

Figure 5.12: Air side measurement of light emission for waveguide coupled cavity backed slot antenna LED. Not-coupled (blue) counts refer to light coming from the device that is not coupled to the waveguide, and grating counts (orange) refer to the light that was coupled to the waveguide and is emitted from the grating.

the spectrometer and the spectrum is recorded on the 1D InGaAs detector, the results are plotted in Fig. 5.13. Note that the data has a Fabry-Perot etalon with a 28nm free spectral range, this comes from reflections between the grating coupler and coupler section. The inset in Fig. 5.13 shows the GDS file for the design, we can calculate the free spectral range in Eq. 5.7.

$$\Delta\lambda_{FSR} = \frac{\lambda^2}{n_g L} \tag{5.7}$$

where λ_{FSR} is the free spectral range, λ is the free space wavelength, n_g is the group index, and L is the round-trip length of the cavity. For an order of magnitude approximation we will assume a constant group index of $n_g = 2.45$, for a weakly reflecting grating the effective length is halfway through the grating [12], and the wavelength is $\lambda = 1550nm$. The the total round-trip length is $L \approx 35\mu m$. Plugging these numbers we get $\Delta\lambda_{FSR} \approx 28nm$.

The spectrally resolved waveguide coupling ratio is plotted in Fig. 5.14, which was calculated by directly applying Eq. 5.4. In addition to the experimental waveguide coupling ratio, we plotted the simulated waveguide coupling ratio in orange, which takes into account the grating and collection efficiencies of the setup. We see an excellent agreement between the experimental and simulated waveguide coupling ratio for the tapered coupler.

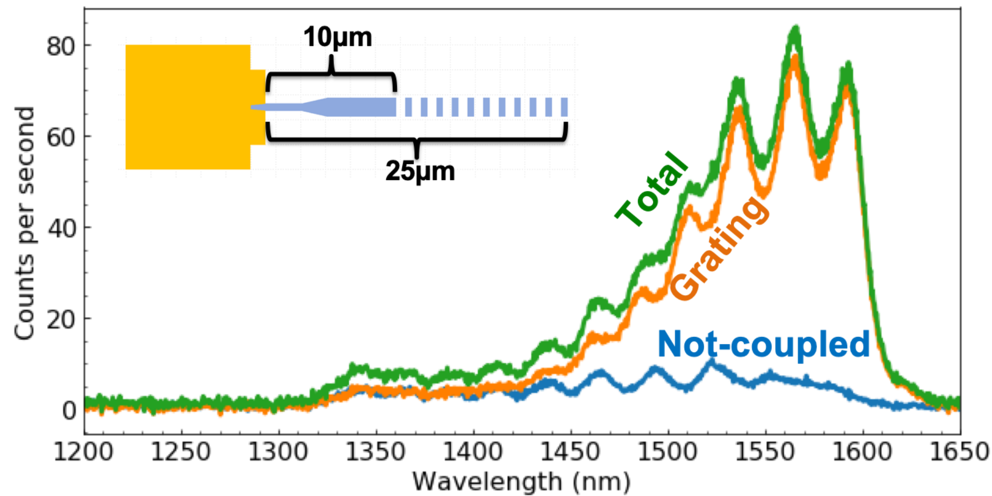

Figure 5.13: Spectra for grating counts (orange) and not-coupled counts (blue), as well as the sum of the two curves (green). Note, there is an optical filter at 1300nm to remove laser light. Inset shows GDS file for the structure, the length explains the Fabry-Perot etalon.

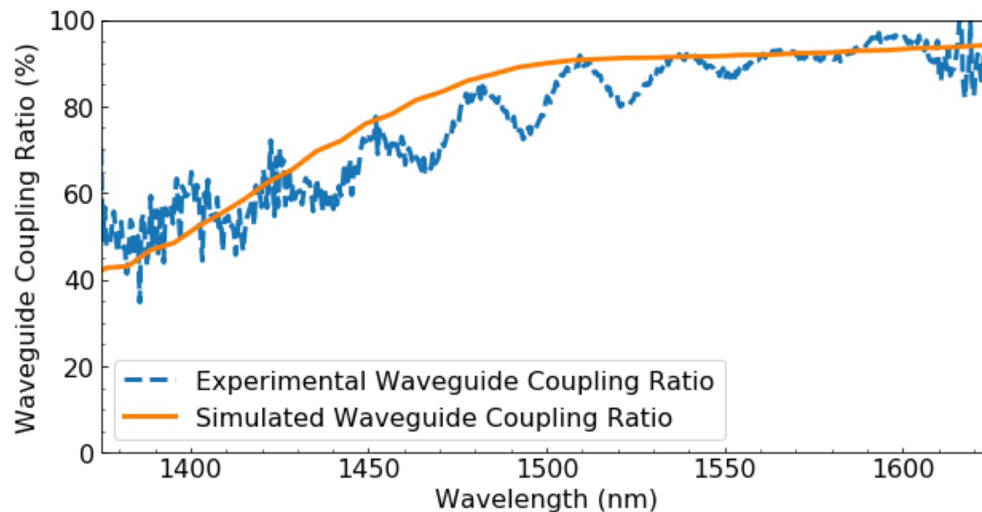

Figure 5.14: Waveguide coupling ratio as a function of wavelength. Calculated from diving the grating spectrum with the total spectrum from Fig. 5.13. The simulated waveguide coupling ratio is also plotted in orange.

Chapter 6

Surface Passivisation

Development of high speed nanoscale photonic components, in particular emitters and detectors, for integrated optical interconnects will be critical for achieving the optical interconnect budget of <10fJ/bit [6]. III-V semiconductors are attractive due to their high speed and bandgaps in the 1.55μm and 1.3μm telecommunication windows; in the past few years there has been interest in developing III-V photonic crystal lasers[61, 62], metal-dielectric nanocavity laser [63] and LEDs [34], nanowire lasers [64], antenna-LEDs[65, 33], and nanoscale detectors [66]. And, in addition to opto-electronic devices, III-V semiconductors have been explored for nanoscale FETs due to their high mobility [67, 68].

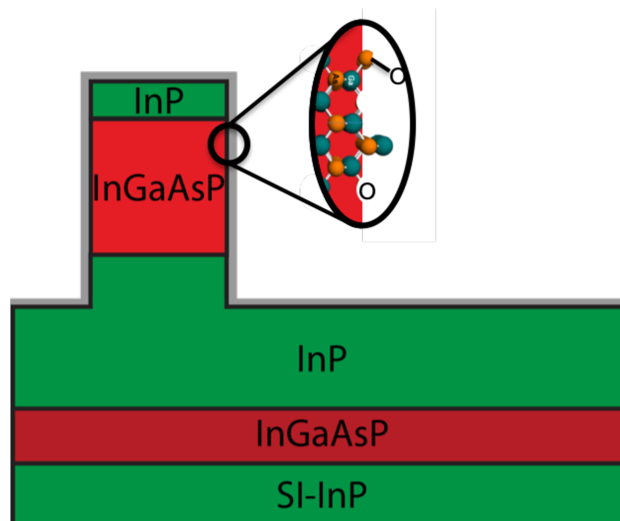

Figure 6.1: Electrically active bonds and impurities at the surface of crystal lattice can lead to high recombination. As we scale the device to the nanoscale the surface to volume ratio becomes larger the surface recombination becomes more significant.

However, when scaling devices down, the surface area to volume ratio increases, making

the device more sensitive to surface effects. If the surface of the semiconductor is not properly terminated the dangling bonds and impurities form an electrically active non-radiative recombination pathway, leading to lower efficiency. Surface passivation looks to mitigate these problems, and while a variety of different fabrication processes have been tested [69], the main mechanisms are regrowth [70], field effect passivation [71, 72, 68, 73, 74, 75] - where a large fixed charge is deposited on the surface, or chemically passivating the dangling bonds through solution treatment [76]. However, chemical passivation has shown stability problems without proper encapsulation [77, 78].

6.1 Surface recombination

From [79], the surface recombination can be written as:

$$R_{surf} = (\frac{A_s}{V_{act}})\frac{np - n_i^2}{\frac{n}{v_h} + \frac{p}{v_e}} \tag{6.1}$$

where A_s is the active region surface area, V_{act} is the active region volume, v_h and v_e are the capture velocities of holes and electrons at the surface, respectively. Under high injection $np >> n_i^2$, this simplifies to:

$$R_{surf} = (\frac{A_s}{V_{act}})v_s n \tag{6.2}$$

where v_s is the surface recombination velocity and is equal to $v_s^{-1} = v_e^{-1} + v_h^{-1}$. Additionally, we can take another approximation for long ridges with very narrow widths ($l >> w$), such that the surface recombination rate reduces to:

$$R_{surf} = (\frac{2h(w+l)}{wlh})v_s n \approx (\frac{2}{w})v_s n \tag{6.3}$$

While this is a useful approximation to understand the scaling dependencies on the width, the full ratio is easy to calculate. Finally, the non-radiative lifetime from surface recombination is then given by:

$$\tau_{surf}^{-1} = \frac{R_{surf}}{n} = (\frac{A_s}{V_{act}})v_s \tag{6.4}$$

Let us now consider our case, for an InGaAs(P) semiconductor, typical values for the surface recombination velocity are in the order of $\approx 10^4 cm/s$. For an antenna-LED with a length of 160nm and width of 20nm we get a non-radiative lifetime of 100ps. This could be acceptable for very high enhancements and pump powers where the radiative lifetime is $< 100ps$ but after processing the surface recombination velocity can be $> 10^5 cm/s$. So, to create an efficient device, we need to consider ways to protect or improve the surface during processing.

6.2 Time correlated single photon counting

Time correlated single photon counting (TCSPC) is by definition measuring the number of photons out of the device. This means the number of photons as a function of time is given by:

$$Photons(t) = \eta_c B_0 N(t)^2 \tag{6.5}$$

where η_c is the collection efficiency. Further, after an excitation from a laser pulse the rate equation is:

$$\frac{dN}{dt} = -AN - B_0 N^2 - CN^3 \tag{6.6}$$

In the following subsections we examine how to model this equation for various regimes. The most popular assumption is that the device is SRH dominated, which allows for a simple exponential fit.

SRH dominated

$$\frac{dN}{dt} \approx -AN \tag{6.7}$$

The solution to this differential equation is:

$$N(t) = N_0 e^{-\frac{t}{\tau}} \tag{6.8}$$

Which means the number of photons collected as a function of time is:

$$Photons(t) = \eta_c B_0 N(t)^2 \tag{6.9}$$

Plugging in our solution above we get:

$$Photons(t) = \eta_c B_0 N_0^2 e^{-\frac{2t}{\tau}} \tag{6.10}$$

Note when most papers fit the decay, they neglect the extra factor of 2 in the exponential lifetime. Generally the constants in front of the exponential could be hard to measure, so it is easier to normalize the equation and just fit the exponential:

$$NormalizedPhotons(t) = e^{-\frac{2t}{\tau}} \tag{6.11}$$

SRH + radiative dominated

A more complete model is to include the radiative recombination term in the rate equation.

$$\frac{dN}{dt} \approx -AN - B_0 N^2 \tag{6.12}$$

The solution to this equation is:

$$N(t) = \frac{A}{e^{A(t-c)} - B} \tag{6.13}$$

where

$$N_0 = \frac{A}{e^{-Ac} - B} \tag{6.14}$$

This means the photons decay curve would be:

$$Photons(t) = \eta_c B_0 \frac{A^2}{(e^{A(t-c)} - B)^2} \tag{6.15}$$

Again, generally it benefits us to normalize to remove the collection efficiency:

$$NormalizedPhotons(t) = \frac{A^2}{N_0^2(e^{A(t-c)} - B)^2} \tag{6.16}$$

Rate equation fit

As shown in the previous subsection, including the radiative rate can greatly improve the fit - we wanted to evaluate if it would be possible to include the Auger term. Unfortunately, the full rate equation does not have a closed form solution, so we have to solve it numerically; however, it follows the same general procedure of integration, normalization, and fitting.

In this case, to keep track of variables, it helps to normalize the rate equation first. To find the appropriate change of variable let us look at the counts again:

$$Photons(t) = \eta_c B_0 N(t)^2 \tag{6.17}$$

$$NormalizedPhotons(t) = N_{Norm}(t)^2 = \frac{N(t)^2}{N_0^2} \tag{6.18}$$

So if we look at the rate equation

$$\frac{dN}{dt} = -AN - B_0 N^2 + CN^3 \tag{6.19}$$

Substituting $N(t) \to N_0 \times N_{Norm}(t)$:

$$N_0 \frac{dN_{Norm}}{dt} = -AN_0 N_{Norm} - B_0 (N_0 N_{Norm})^2 + C(N_0 N_{Norm})^3 \tag{6.20}$$

$$\frac{dN_{Norm}}{dt} = -AN_{Norm} - B_0 N_0 N_{Norm}^2 - CN_0^2 N_{Norm}^3 \tag{6.21}$$

We can numerically integrate this equation:

$$N_{Norm}(t+1) = N_{Norm}(t) + \frac{dN_{Norm}}{dt}dt \tag{6.22}$$

where $N_{Norm}(0) = 1$. From this equation we see that the fit parameters are A, B_0N_0, and CN_0^2.

6.3 Fabrication

One of the most promising techniques reported to date is encapsulating the active region after sulfur passivation in 50nm PECVD SiOx, achieving a low surface recombination of 260 cm/s [80].

In this chapter we report an extremely low surface recombination velocity of 45cm/s for InGaAsP LED ridges with an increase of >180x in the photoluminescence for a device with 200nm width using an aged ammonium sulfide solution with a well controlled atomic layer deposited oxide ($< 17nm$). To our knowledge, this is the lowest surface recombination velocity reported for an InGaAs or InGaAsP active region.

In this section we will discuss two samples: the first was fabricated with an old bottle of ammonium sulfide 20% in water - for brevity we will call this $Old\ (NH_4)_2S$, and the second was fabricated using a new bottle of ammonium sulfide 20% in water that was artificially aged by adding 0.22g elemental sulfur to 7ml of ammonium sulfide 20% in water 20 minutes prior to passivation - likewise for brevity we will refer to this process as $S + (NH_4)_2S$. As will be shown in the section, the $Old\ (NH_4)_2S$ has the best surface recombination velocity, but when we tried to replicate this result with a new bottle of ammonium sulfide 20% in water, the results were significantly worse. However, we were able to demonstrate comparable results with the artificially aged $S + (NH_4)_2S$ solution, which suggests that aging the ammonium sulfide is important to achieve low surface recombination velocity.

The epitaxial layers were grown using metal organic chemical vapor deposition (MOCVD) on InP substrate. They are patterned with hydrogen silsequioxane (HSQ) electron beam lithography resist and etched into ridges with 210nm height and 1μm length with varying widths using an inductively coupled plasma reactive ion etcher with an Ar/H2/CH4/Cl2 chemistry. The HSQ hard mask is stripped with 5:1 buffered hydroflouric acid (BHF) for 2 minutes.

Both samples are pre-cleaned in OPD4262 developer (dilute tetramethylammonium hydroxide) for 1 minute followed by a 15s dip in 10:1 BHF. Between each step the sample is rinsed in DI water and dried with nitrogen. After pre-clean, the sample is then submerged for 20 minutes in the ammonium sulfide solution (either $Old\ (NH_4)_2S$ or the artificially aged $S + (NH_4)_2S$), then it is rinsed with isopropyl alcohol for 5-10s and immediately loaded into an atomic layer deposition (ALD) tool.

The sample that was soaked in $Old\ (NH_4)_2S$ was loaded into a Picosun ALD tool. The deposited stack was 30 cycles Al_2O_3 (≈3nm), 200 cycles TiO_2 (≈4nm), and 100 cycles Al_2O_3 (≈10nm) deposited sequentially without breaking vacuum at 250C. The Al_2O_3 precursors

are Trimethylaluminum (TMA) and DI water, and the TiO_2 precurors are Titanium Tetrakis Isopropoxide (TTIP) heated to 80C and DI water.

The sample soaked in $S + (NH_4)_2S$ was loaded into a Cambridge Fiji F200 ALD tool. The atomic layer deposition stack consisted of 30 cycles Al_2O_3 (≈3nm), 100 cycles TiO_2 (≈4nm), and 30 cycles Al_2O_3 (≈3nm) deposited sequentially without breaking vacuum at 200C. The Al_2O_3 precursors are Trimethylaluminum (TMA) and DI water, and the TiO_2 precurors are Tetrakis(dimethylamino)titanium (TDMAT) heated to 75C and DI water.

Finally, to complete the surface passivation, both samples underwent rapid thermal annealing at 350C in N_2 for 5 minutes. A schematic of the device and scanning electron micrograph (SEM) of the $S + (NH_4)_2S$ sample is shown in Fig. 6.2.

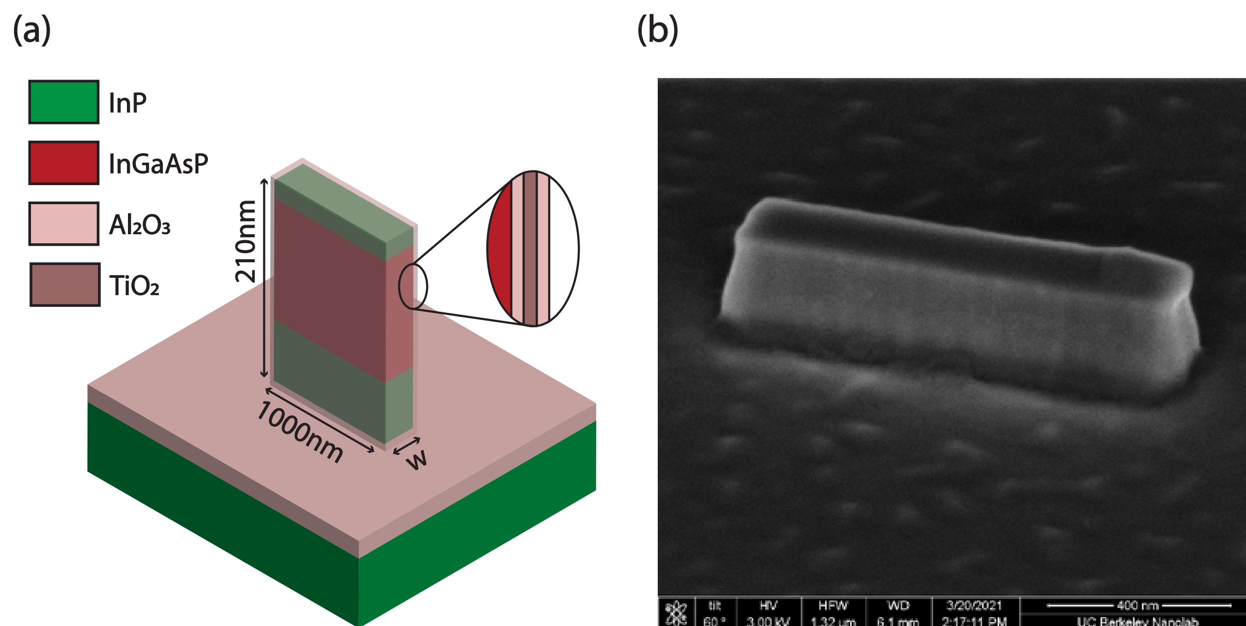

Figure 6.2: (a) Schematic of passivated InGaAsP LED ridge (length:1000nm, height:210nm, width:variable). (b) SEM of 200nm wide LED ridge after surface passivation ($S + (NH_4)_2S$ process)

6.4 Results

To characterize the efficiency of the surface passivation, we collected room temperature detected photoluminescence vs laser pump power (L-L) measurements and time correlated single photon counting (TCSPC) measurements of our LED ridges immediately after etching and after surface passivation. We used a 1170nm continuous wave (CW) laser for the L-L measurements to avoid a time-dependent quantum efficiency, and a femtosecond Ti:Sapphire

laser at 1000nm for the TCSPC measurements. Both setups use a linear polarizer for the input to align the electric field to the length of the ridge, a 50x microscope objective to focus the laser, and a dichroic mirror paired with a 1300nm longpass filter to remove laser light and etch stop emission.

The L-L curves from CW excitation are plotted in Fig. 6.3 for LED ridge widths of 40nm, 200nm, 400nm, and 700nm. We saw an increase in the photoluminesence (PL) after passivation compared to the measurement immediately after etching the sample for all ridge widths.

The rate equation for CW excitation is provided below in Eq. 6.23.

$$G = A'N + BN^2 + CN^3 \tag{6.23}$$

where G is the optical generation rate, A' is the combination of Shockley-Reed-Hall (SRH) recombination and surface recombination, B is the radiative recombination coefficient, and C is the Auger recombination coefficient. Because we are measuring photons ($\propto BN^2$), if the carrier recombination is dominated by SRH and surface recombination (A') we expect a slope of 2 dec/dec in the loglog L-L plot. Likewise, if the recombination is dominated by radiative recombination, we expect a slope of 1 dec/dec. As shown in Fig. 6.4(a) for a device with a width of 200nm, after etching, the L-L slope was 1.62 dec/dec, and, after passivation, the slope decreased to 1.02 dec/dec, which indicates that for these powers radiative recombination is dominating. Additionally, at the lowest pump power measured we saw an increase in the PL of 180× for the 200nm width. Extrapolating from these slopes we would expect the ratio to be even larger at lower pump powers.

Note that the <1.00 dec/dec slope roll-off at pump powers over $\approx 40\mu W$ in the after-passivation measurement can be explained by band-filling rather than onset of Auger recombination. In Fig. 6.4(b), the portion of the spectra below the 1300nm longpass filter is fairly small at $4\mu W$, but as we increased the pump power by an order of magnitude the peak power shifted to shorter wavelengths and the portion of the spectra below 1300nm was no longer insignificant.

To measure the time correlated single photon counting decay we used an MPD InGaAs/InP single photon avalanche photodiode with a 4ps timing resolution. Due to the relatively long lifetimes the device does not reach A' dominated recombination until low carrier concentrations. This corresponds to a low photon count, making it difficult to measure a purely exponential decay without falling below the detector noise floor. So, instead of a pure exponential, we fit the temporal decay curve with both A' and B terms: $G = A'N + BN^2$, referred to here as A'B decay curve. The initial part of the A'B decay curve is dominated by radiative recombination, and decays to an A' dominated rate. This decay curve, where A' and B both contribute, has been previously modeled [81], and, to confirm the quality of the fit, we introduced a new model that can account for Auger recombination by directly fitting to the normalized rate equation.

An example of the decay curves for after etching compared to after surface passivation is shown in Fig. 6.5. The A' lifetimes for the 200nm ridge width after etch, $S + (NH_4)_2S$,

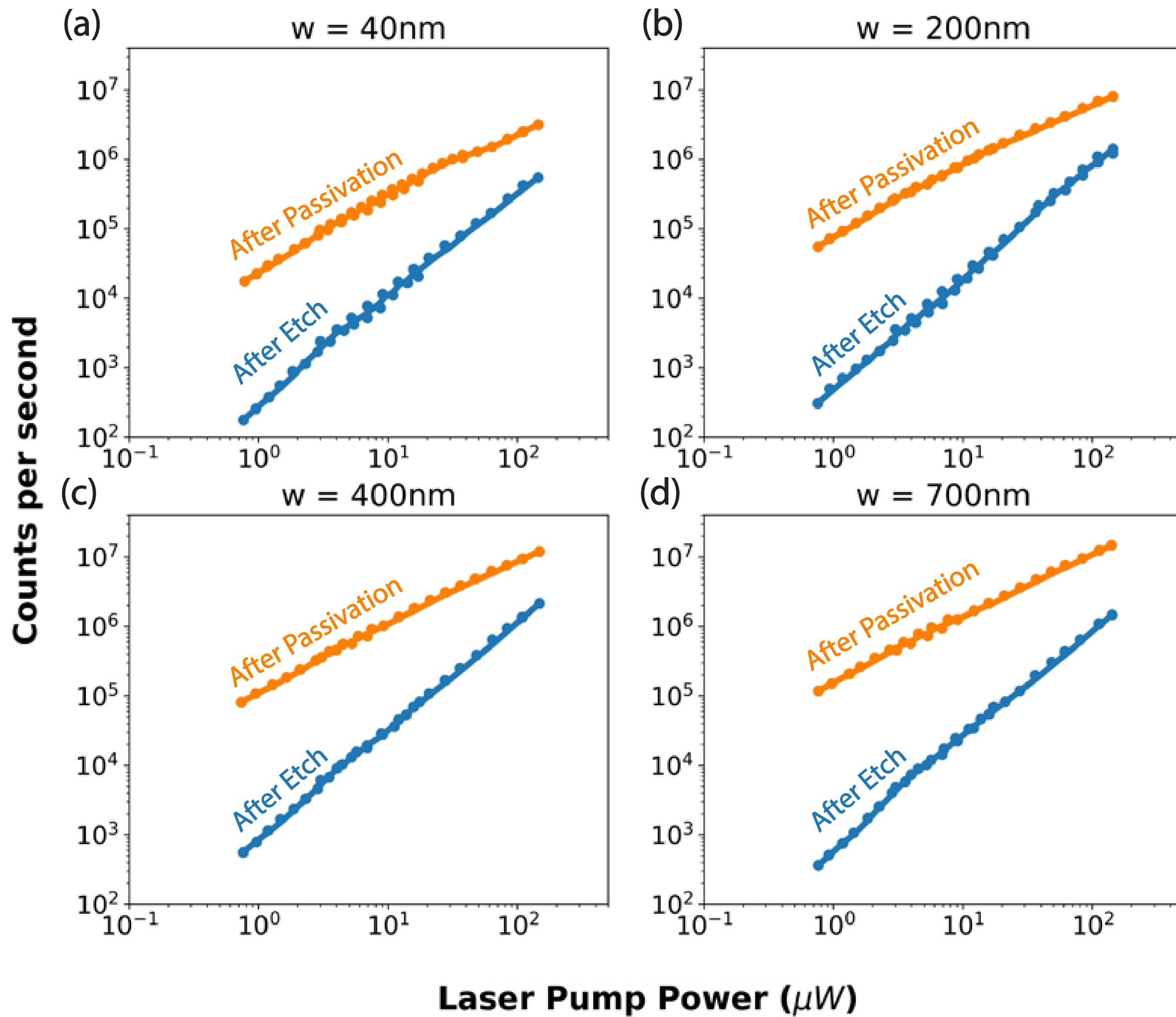

Figure 6.3: Detected photoluminescence vs laser pump power (L-L) curves for LED widths of (a) 40nm (b) 200nm (c) 400nm and (d) 700nm. Samples were measured immediately after etching (blue) and again after surface passivation (orange) - these results are for the $S + (NH_4)_2S$ passivation

and $Old(NH_4)_2S$ are 0.61ns, 29.95ns, and 207.47ns, respectively.

The connection between the surface recombination velocity (SRV) and the A' coefficient is given in Eq. 6.24.

$$A' = \frac{1}{\tau_{SR}} + \frac{1}{\tau_{SRH,bulk}} = \frac{2(w+l)}{wl}v_s + \frac{1}{\tau_{SRH,bulk}} \tag{6.24}$$

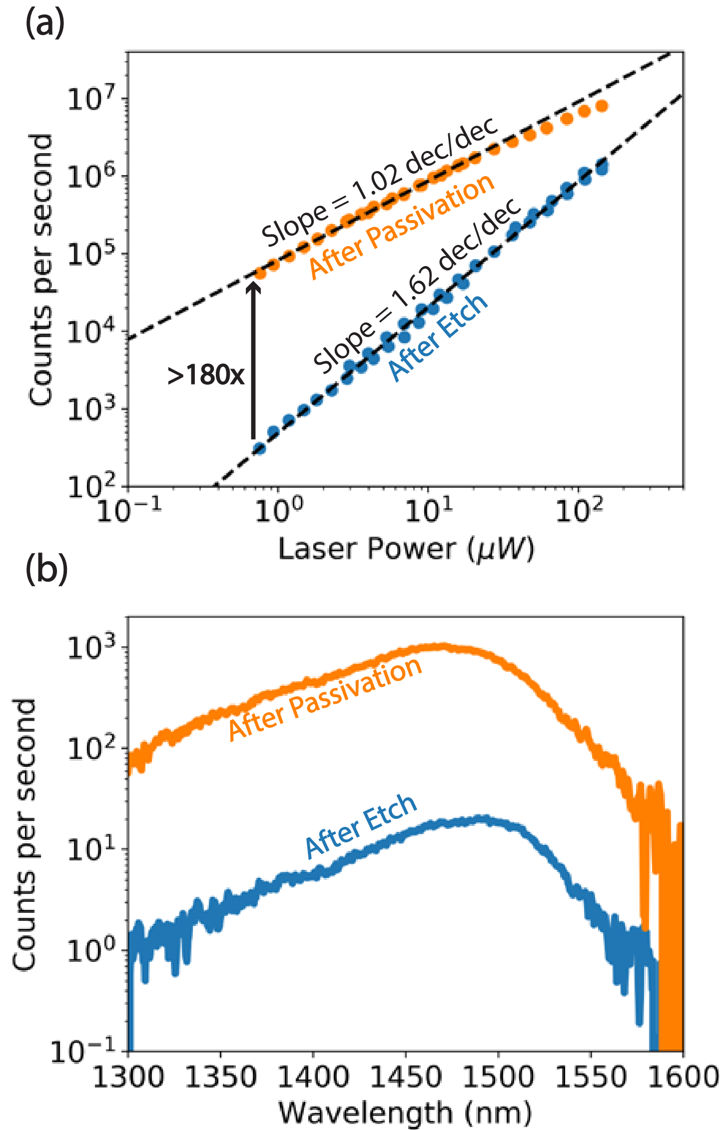

Figure 6.4: (a) L-L curve for continuous-wave μ-photoluminescence measurements at room-temperature displaying increase in photoluminescence and (b) spectra for $4\mu W$ pump power. L-L curve and spectra from 200nm wide ridge with length 1000nm.

where $\tau_{SRH,bulk}$ is the bulk SRH recombination, τ_{SR} is the lifetime from surface recombination, v_s is the surface recombination velocity, w is the width of the ridge, and l is the length of the ridge. For nanoscale devices generally $\tau_{SRH,bulk} >> \tau_{SR}$, so we can neglect the bulk contribution - this lets us directly convert between the measured lifetime and the surface recombination velocity, shown in Fig. 6.6. There are two regions in the graph: at narrow

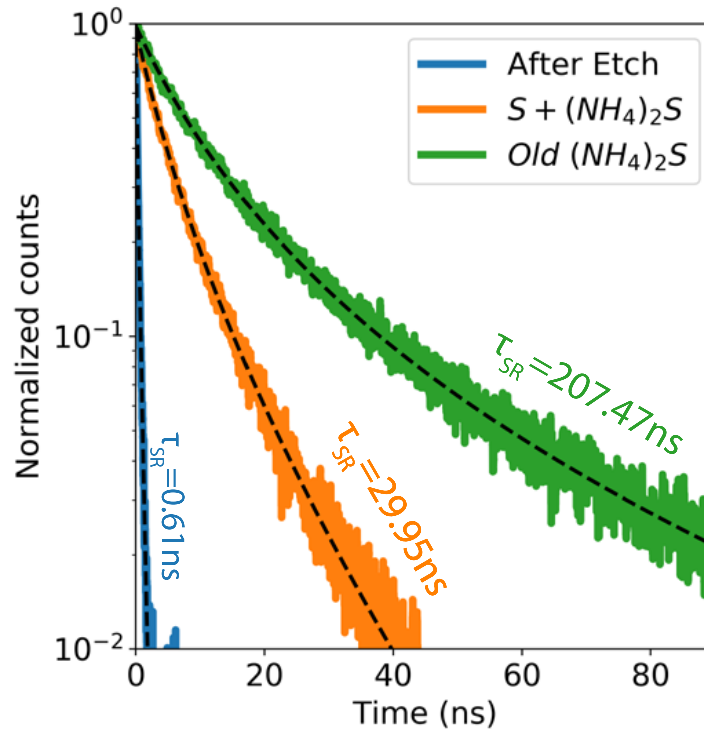

Figure 6.5: Decay curve for 200nm wide ridge measured after etch (blue), for the $S{+}(NH_4)_2S$ passivation (orange), and for the $Old(NH_4)_2S$ passivation (green). Dashed lines show decay curve fit from $A'B$ decay model.

ridge widths the surface recombination increases exponentially (plotted as a dashed line), then at wider ridges the surface recombination velocity is approximately constant (plotted as a solid black line, the gray region represents $\pm$ a standard deviation).

The average fit in the constant region of the surface recombination velocity gives us $1.3 \times 10^4 \pm 1075$cm/s after etching (blue dots), and after passivation this decreases to 190 ± 42cm/s for the $S + (NH_4)_2S$ passivation (orange dots). Likewise for the *Old* $(NH_4)_2S$ passivation (green dots), we get an average surface recombination velocity of 45 ± 15cm/s.

We also tested several small process variations. Without artificial aging of the $(NH_4)_2S$, the surface recombination velocity is lower than after etching, but is orders of magnitude higher than the $S + (NH_4)_2S$ process. In addition to aging the $(NH_4)_2S$ solution with elemental sulfur, we achieved similar results by aging the solution on a hotplate at 30C for 30 minutes prior to soaking the sample. We have not performed a full design of experiment on the process space, so it is possible that there exists a more optimal aging process (aging method, $(NH_4)_2S$ concentration, soak time), ALD conditions (temperature, thickness, precursor), and annealing condition (gas, time, temperature) that would allow for even lower

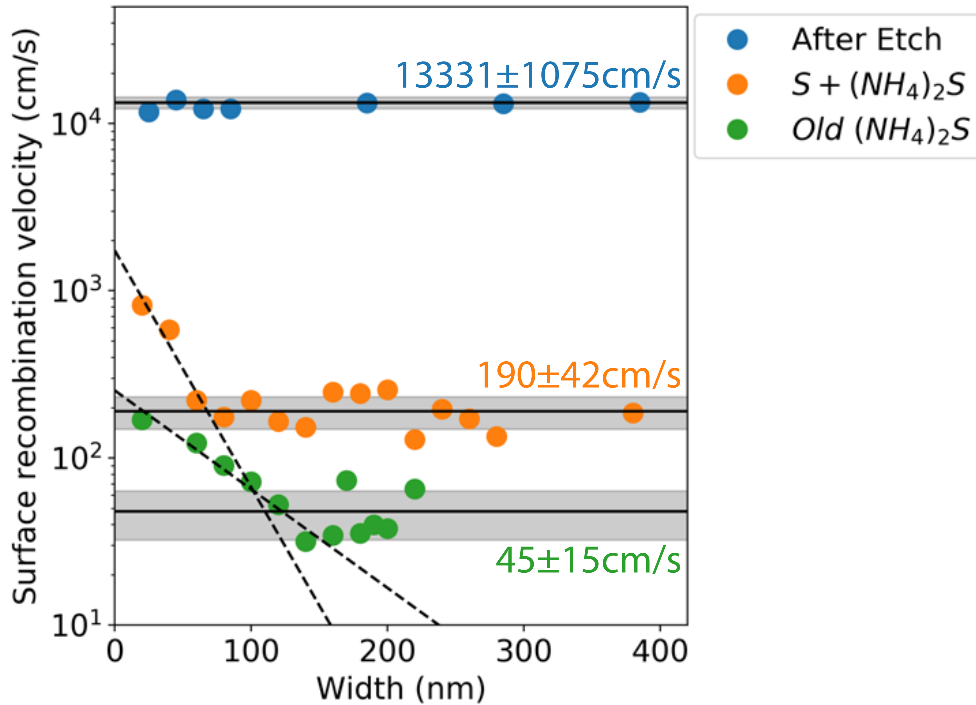

Figure 6.6: Surface recombination velocity as a function of width. At narrow ridge widths the surface recombination velocity increases exponentially (plotted as a dashed line). At large widths the surface recombination is approximately constant (mean plotted as a solid black line with ± a standard deviation).

surface recombination velocities.

Importantly, because the surface is encapsulated, it shows no degradation of surface properties over several months and testing with the CW and femtosecond lasers. This surface passivation, using thin ALD oxides, can potentially enable highly efficient nanoscale devices.

6.5 Thinner oxide

This surface passivation is particularly promising for applications that require thin oxides like FETs or antenna-LEDs, but it still too thick to be practical.

As discussed in Chapter 3.3, the effective oxide width scales as:

$$w_{eff} = w + 2\frac{\epsilon_s}{\epsilon_{ox}}t_{ox} \tag{6.25}$$

We can modify this equation if we have both Al_2O_3 and TiO_2.

$$w_{eff} = w + 2\frac{\epsilon_s}{\epsilon_{Al_2O_3}}t_{Al_2O_3} + 2\frac{\epsilon_s}{\epsilon_{TiO_2}}t_{TiO_2} \tag{6.26}$$

where $\epsilon_s = 12.46$, $\epsilon_{Al_2O_3} = 2.72$, and $\epsilon_{TiO_2} = 5.29$. For a ridge with 20nm width, 6nm Al_2O_3, and 4nm TiO_2 we have an effective ridge width of 94nm! In order to reduce the effective width we need to decrease the oxide thickness.

In order to thin down the oxide, we can either deposit a thinner oxide or attempt to selectively remove it after annealing. These following tests were all conducted on an InGaAs active region, so the decay measurements are limited by the overlap in emission and detector responsivity. The first test we conducted was to follow the same process as the sulfur saturated process, but deposit a bilayer Al_2O_3/TiO_2 (3nm/4nm) instead of a trilayer $Al_2O_3/TiO_2/Al_2O_3$ trilayer chip (3nm/4nm/3nm), the results are shown in Fig. 6.7.

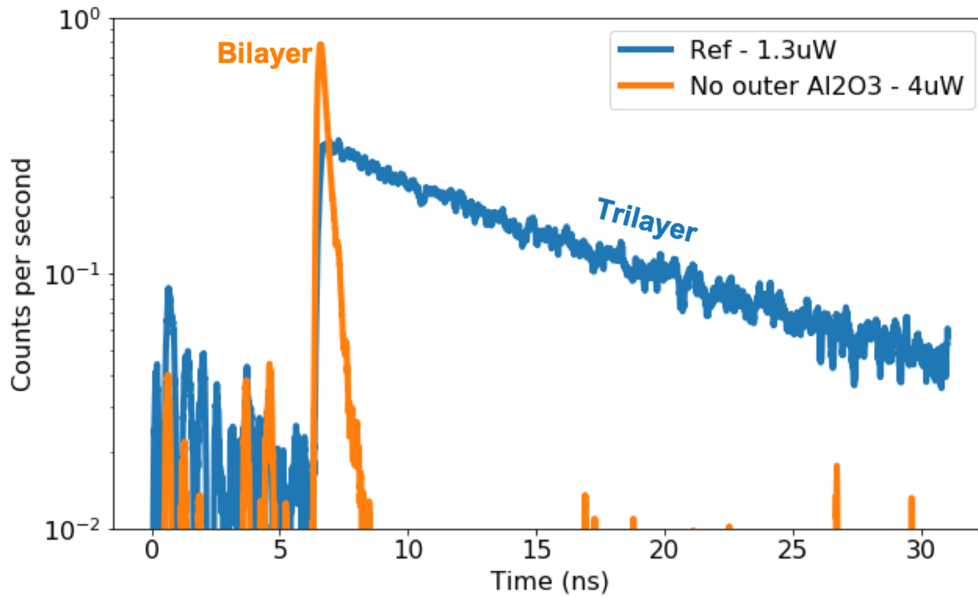

Figure 6.7: Comparison of decay for $Al_2O_3/TiO_2/Al_2O_3$ trilayer chip (3nm/4nm/3nm) and bilayer Al_2O_3/TiO_2 (3nm/4nm) after annealing. Bilayer has significantly shorter lifetime.

We saw a significantly faster decay for the bilayer, demonstrating that this is not compatible with our current surface passivation. From this we concluded that the trilayer plays an important role in surface passivation, and concentrated our efforts on reducing the oxide thickness after surface passivation.

Accordingly, we explored how to remove the Al_2O_3 after N_2 annealing, and found that Al_2O_3 can be selectively wet etched with a 30:1 buffered hydroflouric acid (BHF) [82] with

little to no degradation in the surface recombination velocity. However, the effective width for this structure - assuming a 20nm semiconductor ridge width - is still 66nm.

Our next test was to explore thinner trilayer stacks - this process is shown in Fig. 6.8a-c. The blue curve is a reference after BHF removal. In Fig. 6.8a, we found that almost all the trilayer decay curves had nearly the same profile (after normalization) and further evidence that the bilayer (after outer Al_2O_3 BHF removal) was not significantly different than the trilayers. Then we soaked the samples in 30:1 BHF for 30s and measured the decays again as shown in Fig. 6.8b. All the samples, with the exception of the thinnest oxide, were nearly unchanged. Note, however, that all the bilayers decay slightly faster than the thickest reference bilayer, but because these are InGaAs active regions and the decay is likely radiative limited, it is unclear how significant this change is. At this point in the process, we significantly reduced the effective width of the device - again assuming a 20nm semiconductor ridge width - the effective width was reduced to 48nm, nearly 2x narrower effective width than the trilayer and nearly equivalent to a 3nm $\mathrm{Al_2O_3}$ oxide.

Finally, we deposited an Ag antenna and performed substrate removal as shown in Fig. 6.8c. We found that this did not change the surface recombination velocity significantly. However, after metal deposition we saw that the pumping and collection was significantly more efficient, so to get comparable signal intensity we had to cut the pump power by a factor of 10x. This was an important verification that this surface passivation is compatible with cavity backed slot antenna formation.

6.6 Field effect passivation

While potential mechanisms for surface passivation were not fully explored in this chapter, we believe the surface is passivated by a combination of chemical passivation through sulfur-saturated ammonium sulfide and field effect passivation through a high density of negative fixed charges in the ALD trilayer - specifically in the inner $\mathrm{Al_2O_3/TiO_2}$ layers [72, 75, 83, 84]. Field effect passivation can possibly explain the deviation from the constant surface recombination velocity at narrow ridge widths ($< 160nm$), because narrower ridge widths lead to higher electrostatic doping [83, 84], and both surface recombination and the radiative decay rate from electrostatic doping are proportional to N. So, at the narrowest ridge widths, the decay curve does not give an unambiguous fit of surface recombination.

By placing a large fixed charge on the surface of the semiconductor we can move the Fermi level, which leads to low minority carrier concentration on the surface, which in turn yields a low effective surface recombination velocity. In addition, for nanoscale dimensions, the Fermi level movement can lead to a high majority carrier concentration - effectively leading to electrostatic doping. For doping this modifies the rate equation to:

$$G \approx A'N + BN(P_{ES} + N) + C_n N^2(P_{ES} + N) + C_p N(P_{ES} + N)^2 \tag{6.27}$$

where P_{ES} is the electrostatic doping level and $C_{n,p}$ are the Auger rates for electron and holes. This can then be arranged to be:

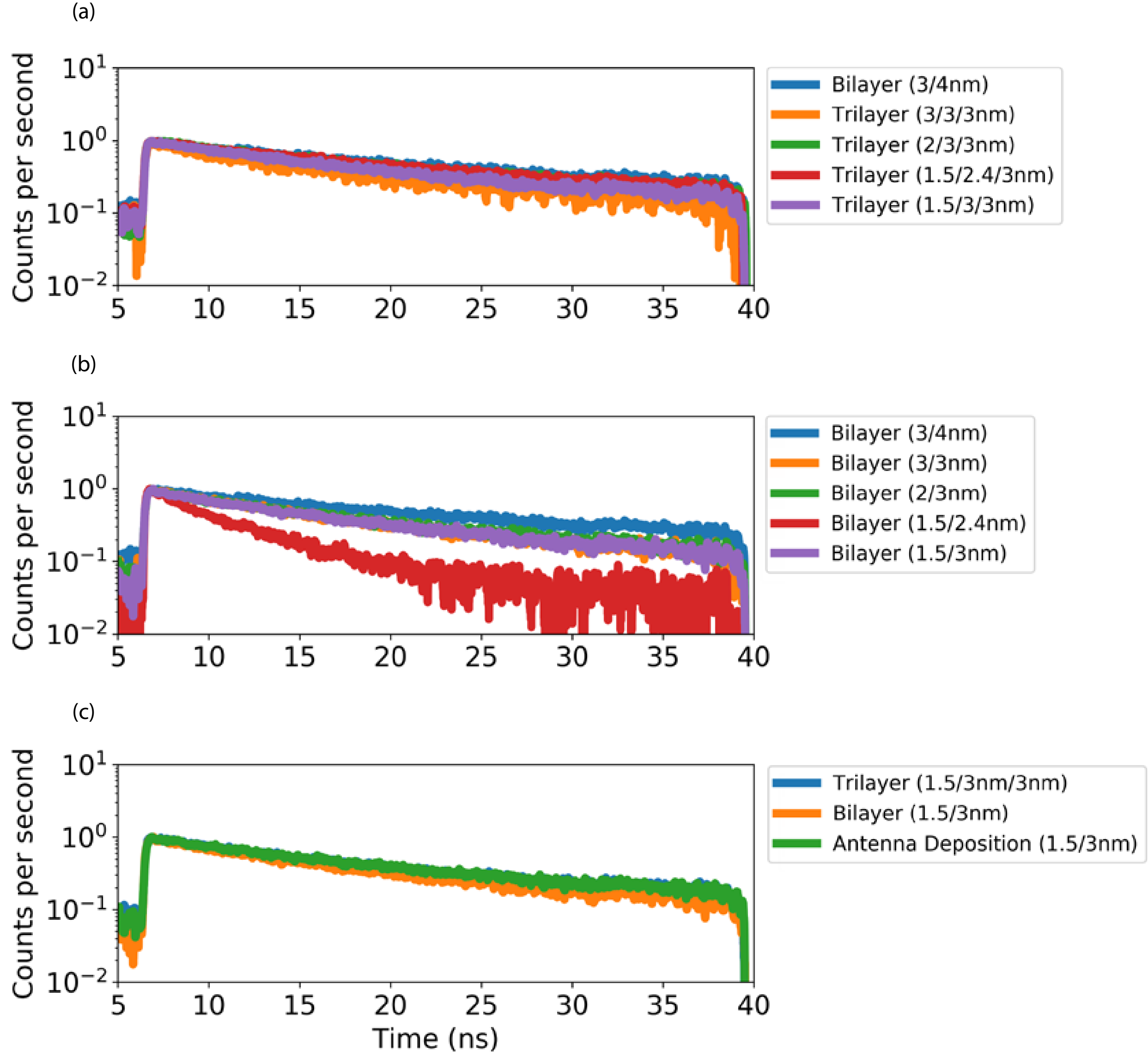

Figure 6.8: Decay for InGaAs active region test chip for different Al_2O_3 and TiO_2 thickness (a) after annealing, (b) after 30:1 BHF removal of outer Al_2O_3, and (c) after antenna deposition and substrate removal for 1.5nm/3nm Al_2O_3/TiO_2 bilayer - the antenna deposition sample was pumped at 10x less power.

$$G \approx (A' + BP_{ES} + C_pP_{ES}^2)N + (B + 2C_pP_{ES} + C_nP_{ES})N^2 + (C_n + C_p)N^3 \tag{6.28}$$

where we see that the decay from electrostatic doping has terms proportional to N,

similar to the A' term. The electrostatic doping and change in the decay are modeled in Fig. 6.9, where we assumed a fixed oxide charge between $Q_{ox} = -(2 \times 10^{12} - 3 \times 10^{13})cm^{-2}$ [72, 83]. For the $S + (NH_4)_2S$ $A' + P_{ES}$ curve we used the same surface recombination velocity from the text ($190cm/s$), but for the Old $(NH_4)_2S$ $A' + P_{ES}$ curve we decreased the surface recombination velocity to $24cm/s$ to get a better fit.

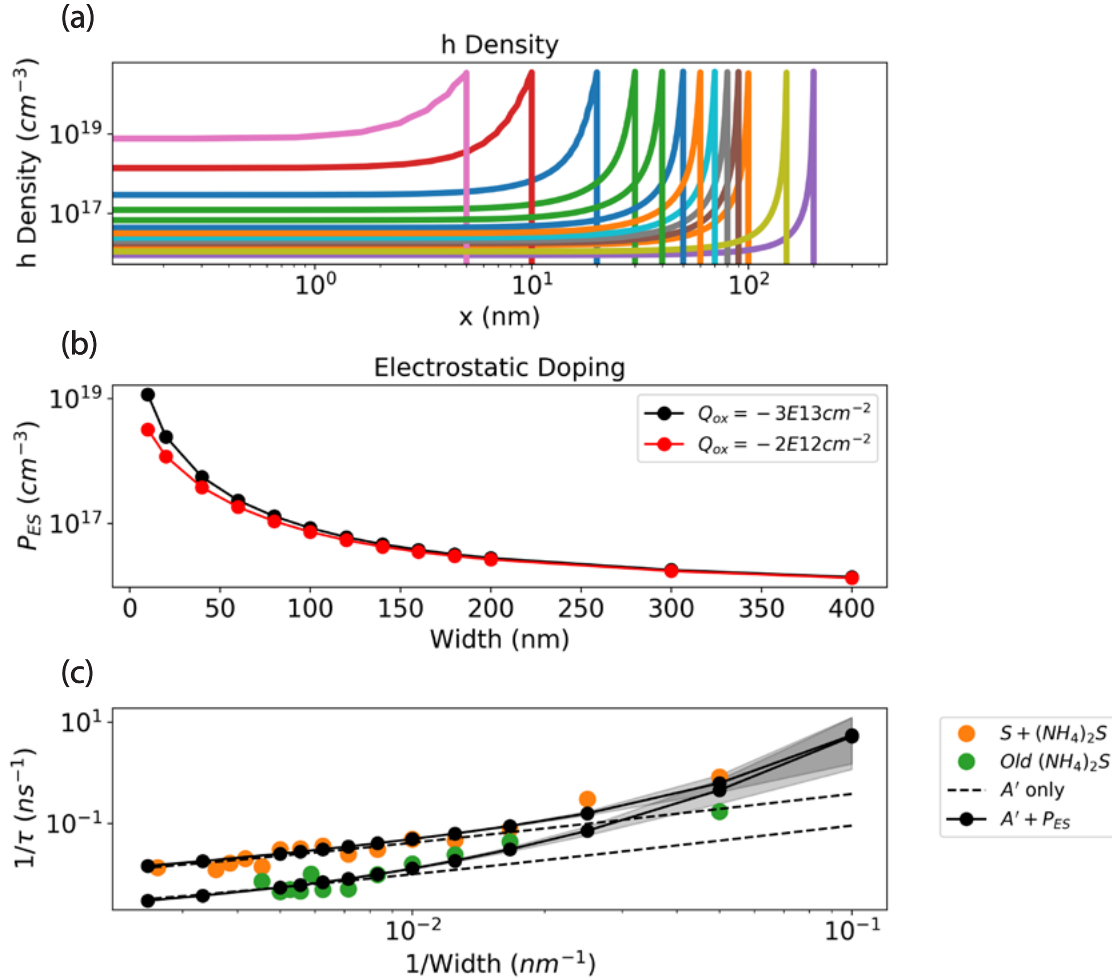

Figure 6.9: (a) h density as a function of position for different widths assuming $Q_{ox} = -3E13$, (b) effective electrostatic doping level, P_{ES}, for $Q_{ox} = -3 \times 10^{13}cm^{-2}$ and $Q_{ox} = -2 \times 10^{12}cm^{-2}$, and (c) $1/\tau$ as a function of 1/width - dashed line is for A' fit only and solid black line includes average between electrostatic doping terms from (b) - gray bars represent $|Q_{ox}| = 2 \times 10^{12} - 3 \times 10^{13}cm^{-2}$ range

We can try to validate this model by looking at the internal quantum efficiency for low pump powers. To model the internal quantum efficiency as a function of generation rate we

need to simulate a few quantities. The generation rate is:

$$G = \eta_{in} \frac{P_{laser}}{\hbar\omega_{in} V_{active}} \tag{6.29}$$

where η_{in} is the absorption efficiency, P_{laser} is the input laser power, $\hbar\omega_{in}$ is laser photon energy, and V_{active} is the active region volume. η_{in} is simulated by using a Gaussian beam with a spot size of $3.8\mu m$ where the electric field is polarized along the length of the device.

The power collected ($P_{collected}$) is:

$$P_{collected} = \eta_{out}\eta_{optics}\hbar\omega_{out} B N^2 V_{active} = Counts \times \hbar\omega_{out} \tag{6.30}$$

where η_{out} is the collection efficiency, η_{optics} is the product of the detector and optical setup efficiencies, and $\hbar\omega_{out}$ is the output photon energy.

Then the internal quantum efficiency is:

$$\eta_{IQE} = \frac{Counts \times \hbar\omega_{out}}{P_{laser}} \frac{1}{\eta_{in}\eta_{out}\eta_{optics}(\frac{\omega_{out}}{\omega_{in}})} \tag{6.31}$$

The collection efficiency, η_{out}, is simulated by placing a dipole in the active region and integrating the farfield power within a numerical aperture of 0.5. The modeling parameters are summarized in Table 6.1.

Parameter summary				
Width	η_{in}	η_{in} ($\theta_{gausian} = 15°$)	η_{out}	η_{optics}
40nm	0.26%	0.11%	11.2%	0.05%
100nm	0.35%	0.30%	16.4%	0.05%
200nm	0.33%	0.29%	16.3%	0.05%

Table 6.1: Summary of efficiencies used to calculate the internal quantum efficiency η_{IQE}. η_{in} and η_{in} ($\theta_{gausian} = 15°$) are the input efficiencies assuming a Gaussian beam profile perfectly aligned to the ridge length and with a $\theta_{gausian} = 15°$ polarization offset. η_{out} is the collection efficiency, and η_{optics} is the product of optical and detector efficiencies.

Finally, we measured new L-L curves for 40nm, 100nm, and 200nm devices and converted these to η_{IQE} in Fig. 6.10. We assumed a surface recombination velocity of 190 cm/s, $B = 10^{-10} cm^3 s^{-1}$, and $C_n = C_p = 8 \times 10^{-29} cm^6 s^{-1}$ [85]. The internal quantum efficiency for the three samples was within the same order of magnitude for the pump power range

measured. If we were seeing increased surface recombination velocity (or even constant surface recombination velocity) we would expect the internal quantum efficiency of the device to decrease as we decrease the width. Because it stayed approximately constant, that was some evidence that we may have electrostatic doping of the device.

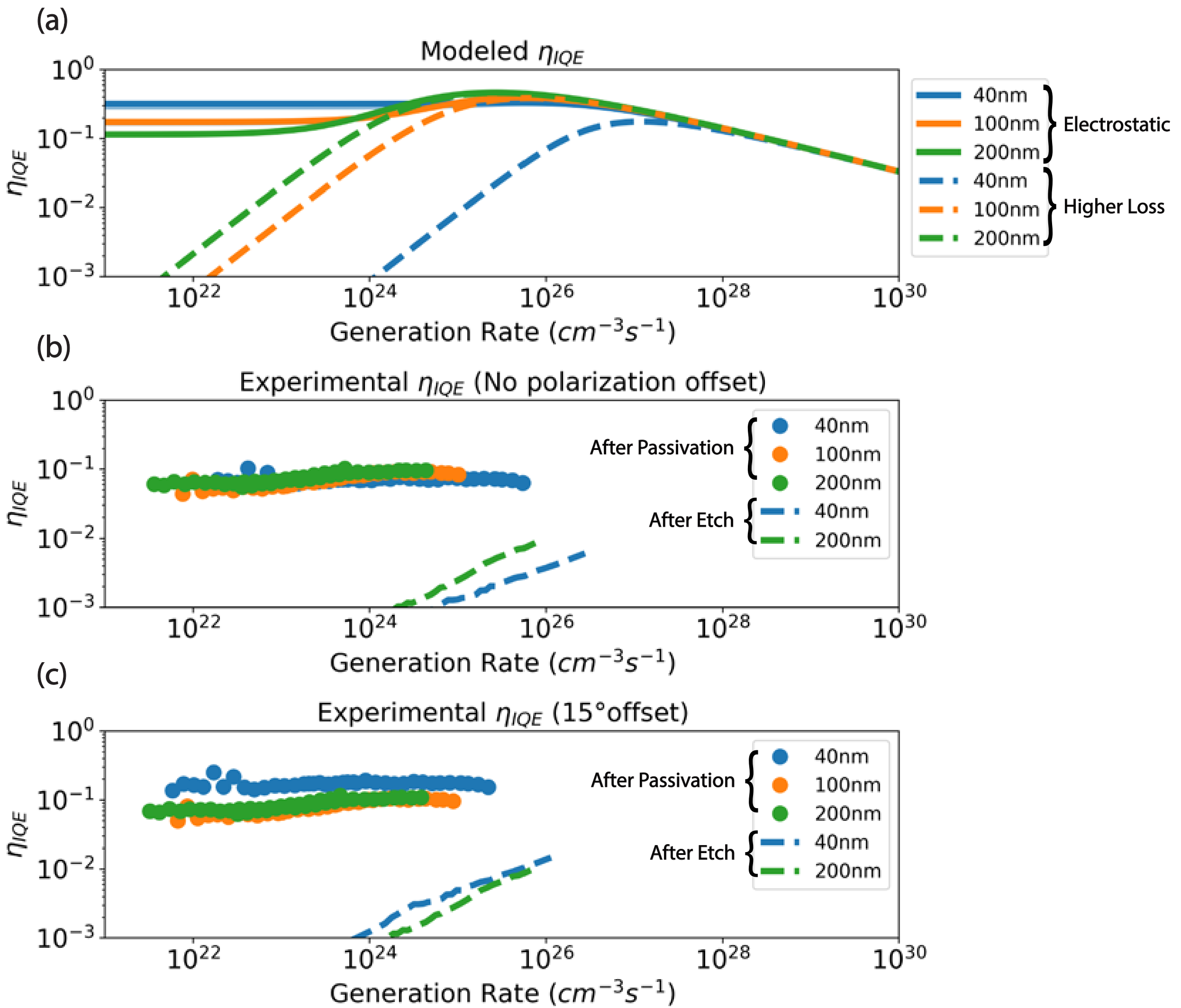

Figure 6.10: (a) Expected η_{IQE} assuming we have electrostatic doping (solid) and assuming decreased lifetime is from higher non-radiative emission (dashed line). Experimental η_{IQE} using adjusted counts and powers from Table 6.1 for after etch and after passivation in $S + (NH_4)_2S$ sample assuming perfectly aligned Gaussian beam (b) and with a 15° polarization offset (c).

We can now look at the model for an antenna coupled device where we assume a surface

recombination velocity of $v_s = 1.3 \times 10^4 cm/s$ for an unpassivated device and $v_s = 190 cm/s$ for the passivated device. Again we are setting $B = 10^{-10} cm^3 s^{-1}$, and $C_n = C_p = 8 \times 10^{-29} cm^6 s^{-1}$, and the electrostatic doping for a 20nm ridge is simulated to be $P_{ES} = 2.5 \times 10^{18} cm^{-3}$. We can then sweep the average enhancement from 1 to 256 to see how the internal quantum efficiency is expected to change using Eq. 6.32. We see for the highest enhancements we have near unity efficiency until we reach the Auger recombination dominant regime.

$$\eta_{IQE} = \frac{F_{avg} B N (P_{ES} + N)}{A' N + F_{avg} B N (P_{ES} + N) + C_n N^2 (P_{ES} + N) + C_p N (P_{ES} + N)^2} \tag{6.32}$$

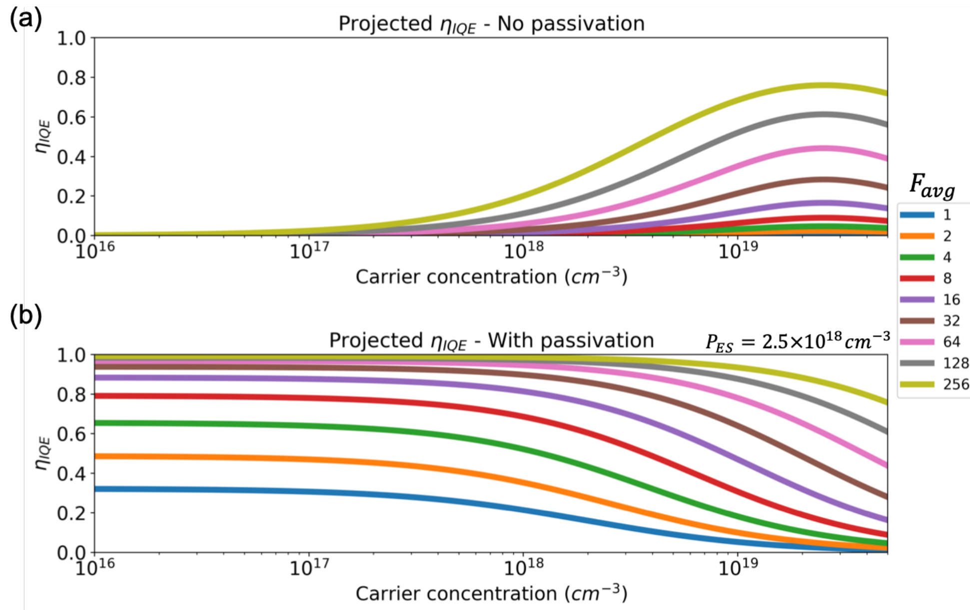

Figure 6.11: Projected η_{IQE} as a function of enhancement for a 20nm width for (a) undoped, unpassivated device, and (b) passivated device with $P_{ES} = 2 \times 10^{18} cm^{-3}$.

From this analysis we can see that the efficiency in the unpassivated device is extremely low when average enhancement is small. However, as we increase the enhancement, the device can be efficient at high carrier concentrations. But, by passivating the device (and inducing electrostatic doping) the internal quantum efficiency can be high even at low carrier concentrations and average enhancement, and as we increase the average enhancement this approaches unity. Additionally, we can see that for higher average enhancement the radiative rate is dominant over Auger recombination for higher carrier concentration. This effectively allows us to user higher carrier concentrations while maintaining high internal quantum efficiency.

We also note that p-doping the active region would have a similar affect on increasing the internal quantum efficiency at low pump powers. However, we would need to test if p-doping is also compatible with this surface treatment, as band bending potentially plays a large role in the surface passivation. If we dope to $N_a = 2 \times 10^{19} cm^{-3}$, this would provide us with high internal quantum efficiency and high speed even with low carrier concentrations. However, the 3dB frequency for the doped active region would be:

$$f_{3dB} = \frac{F_{avg} B_0 p_0}{2\pi} \tag{6.33}$$

This equation is a factor of two smaller than the undoped case, but the benefit is that this rate would be maintained even at low injection. We see that for a doping of $N_a = 2 \times 10^{19} cm^{-3}$, we need an average enhancement of 314 to get 100GHz direct modulation. If we dope the active region slightly higher at $N_a = 3 \times 10^{19} cm^{-3}$ then we would only need an average enhancement of around 200. But we can also see that in order to close the link at the highest datarates we may want to do a small signal modulation under high injection - then, if we bias the device above the doping level, we would retain the factor of 2 and get faster modulation. Finally, the ABC model is appropriate for approximating modulation rate, output optical power, and internal quantum efficiency, but does not account for high injection effects around degeneracy. More sophisticated modeling and experimental verification are still required.

In the next chapter we will combine the waveguide coupling, internal quantum efficiency, and enhancement results in a full link model.

Chapter 7

Link Model

In this chapter we model the link performance using an optical antenna-LED as our transmitter and state-of-the-art components in the receiver. By using a transimpedance amplifier, a series of linear amplifiers, and interleaving StrongArm sense amplifiers, we achieved an end-to-end energy consumption of under 1fJ/bit. We also discuss ways to increase the power of the antenna-LED to meet the system demands.

7.1 Transmitter macro model

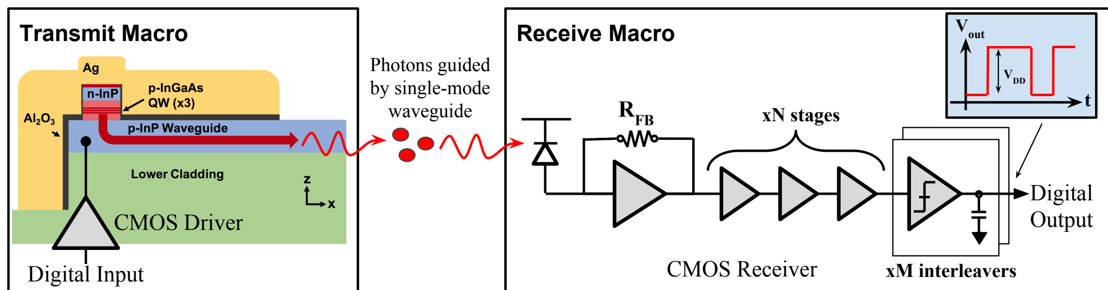

Figure 7.1: Full Link Image

The energy per bit (E/b) of the transmitter can be broken down into two sources. The first is the energy required to modulate the device, and the second is the photon energy consumption. The cavity backed slot antenna-LED is small, fast, efficient, and can operate without a threshold current making it a good candidate for the optical source.

The required energy from modulation comes from both the static and dynamic energy consumption – in other words, from the Joule loss and the parasitic capactive charging of the device.

$$E_{mod,dynamic} = \frac{1}{2}CV^2$$

$$E_{mod,static} \approx \frac{I^2R}{2} \times \frac{1}{DataRate}$$

where $DataRate$ is the data rate in bits/s. This can be calculated by knowing the parasitic capacitance and the L-I-V curve of the device, to first order we can model this by the rate equation.

$$\frac{\eta_{inj}I}{qV_{active}} = AN + F_{avg}B_0N^2 + CN^3$$

where I is the current and η_{inj} is the injection efficiency into the active region (assumed to be ≈ 1). Then we can write the optical power as:

$$P_{optical} = \hbar\omega F_{avg}B_0N^2V_{active}$$

which can be rewritten as:

$$P_{optical} = \frac{\hbar\omega}{q}\eta_{IQE}I$$

therefore the required current is:

$$I = \frac{qP_{optical}}{\hbar\omega\eta_{IQE}}$$

Now we can write an approximation for the modulator energy consumption as:

$$E_{mod,dynamic} \approx \frac{1}{2}C(\frac{qP_{optical}}{\hbar\omega\eta_{IQE}})^2R_{ridge}^2$$

$$E_{mod,static} \approx (\frac{qP_{optical}}{\hbar\omega\eta_{IQE}})^2\frac{R_{total}}{2} \times \frac{1}{DataRate}$$

where R_{ridge} is the series resistance of the ridge, and R_{total} is the total series resistance of the LED ridge, waveguide, and contacts. The main source of parasitic capacitance C is from the waveguide to the silver antenna/n-contact- i.e. the isolation oxide.

The photon energy consumption (E_{TX}) can be calculated from the total optical power required for the receiver and then propagated back to the transmitter by including the various loss mechanisms, namely the antenna efficiency, waveguide coupling efficiency, and photodiode absorption efficiency.

$$E_{TX} = T_{Bit}V_{TX}(I_{Req,noise} + I_{Req,swing})$$

$$V_{TX} = \frac{\hbar\omega}{q}\frac{1}{\eta}$$

$$\eta = \prod \eta_{system}$$

where T_{Bit} is the bit period, V_{TX} is link conversion efficiency in Volts, $I_{Req,noise}$ is the noise sensitivity requirement, $I_{Req,swing}$ is the swing sensitivity requirement, and η is the product of system efficiencies.

Due to high average enhancement, the internal quantum efficiency can be near unity, and, by tuning the radiation resistance, the antenna efficiency can be 64% with a 94% coupling efficiency to a single mode waveguide. However, due to the small active volume, the maximum power is constrained to a few μW for a single quantum well.

7.2 Receiver macro model

The receiver is composed of a photodiode, transimpedance amplifier, N linear amplifier stages, and M time-interleaving StrongArm sense amplifiers [86]. The main constraint placed on the receiver side is that the output signal is a rail-to-rail swing, digital signal. For our model we use an unbiased photodiode to prevent Joule loss, with a capacitance of 600aF and absorption efficiency to 80% [87].

Model Parameters		
Model Inputs	Variable	Value
f_T	Extrinsic unity current-gain frequency	260 GHz [88]
α	Fraction of f_T for self loaded stage	0.8
β	Gain stage input/output cap ratio	0.33
V_{DD}	Supply Voltage	0.8 V
E_{mod}	Modulator energy per bit	100 aJ/bit
V_{TX}	Link conversion efficiency	1.65V
R	Photodiode Responsivity	1A/W [87]
C_{PD}	Photodiode Capacitance	600aF [87]
Sweep Parameters		
Sweep Variable	Description	Range
N	Number of linear amplifiers	0-15
M	Number of interleaving stages	1-16
I_{Bias}	Bias current amplifiers	1μA-100mA

From the CMOS side, the receiver model takes into account not only resistor thermal noise and transistor FET noise, but also the input swing sensitivity required to have an output rail-to-rail signal. All of these metrics combine to yield a topology and data-rate specific E/b both for the receiver side and transmitter side. For each data-rate, an optimization was performed to find the minimum E/b given the sweep parameters and model parameters in Table 7.2. For more information on the design methodology and parameters please refer to [86].

7.3 End-to-end link performance

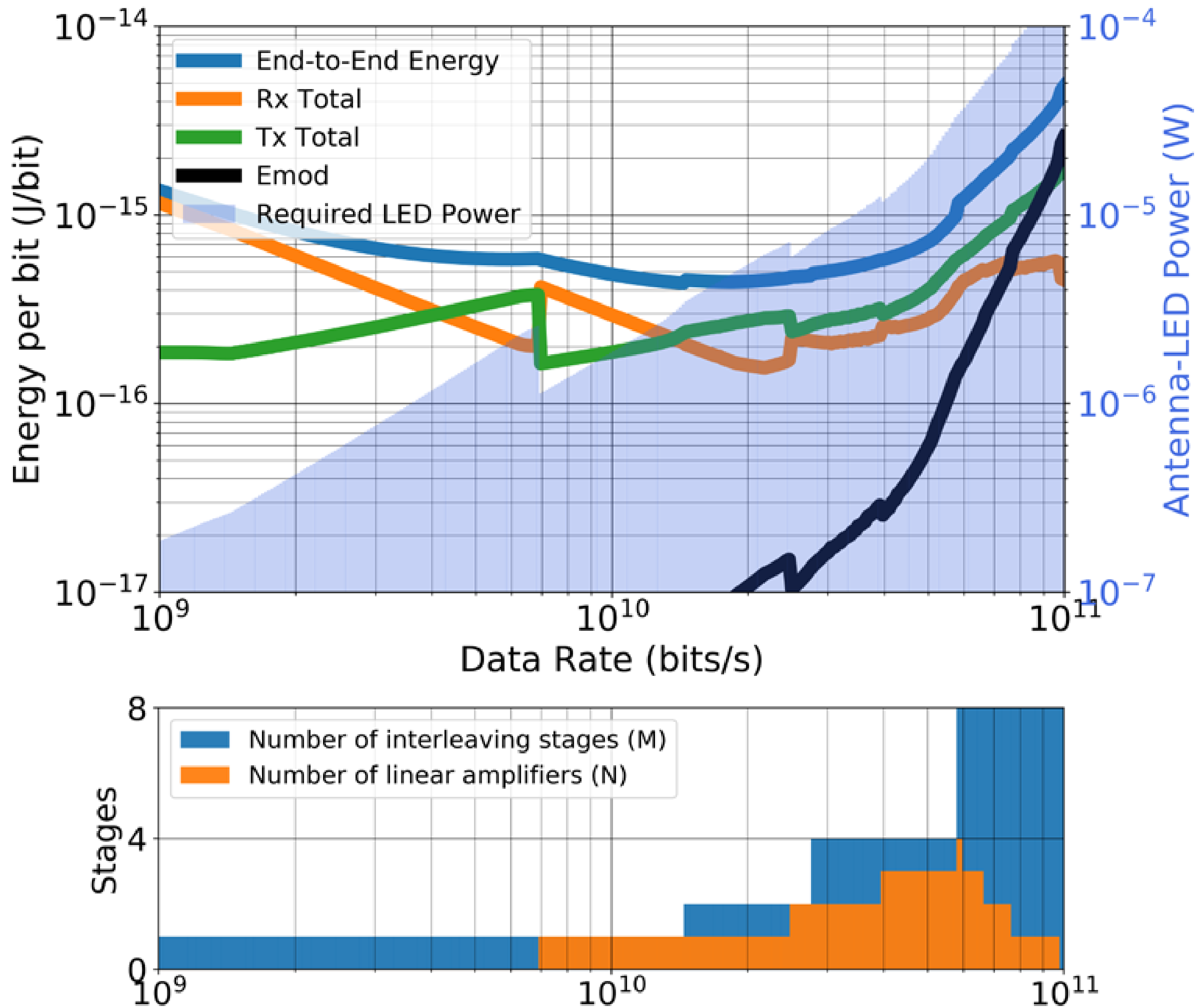

Figure 7.2: Optimal energy per bit breakdown, required LED power, and receiver topology vs data rate, using the parameters in Table 7.2

As shown in Figure 7.2, we are able to achieve under 1fJ/bit up to 60Gbps by using the parameters in Table 7.2 and the amplifier topology in the bottom panel of the figure.

High f_T and α are required to maintain an energy per bit under 1fJ/bit with a minimal LED power. In Figure 7.2, with an f_T of 260 GHz, the maximum self-loaded gain quickly saturates for high data-rates. This forces the optimizer to trade-off, increasing the number

of Rx linear amplifiers (albeit with a power penalty) and increasing Tx energy directly. The optimum E/b attempts to balance the energy expended by the receiver and E_{TX}.

The antenna-LED power imposes a restriction on the maximum attainable data-rate for the link. Let us plug in some reasonable numbers for an antenna-LEDs using a single quantum well active region (6nm thickness), where we assume that we are pumping the device quite hard: $\hbar\omega = 1.28 \times 10^{-19} J$, $B_0 = 10^{-10} cm^3 s^{-1}$, $F_{avg} = 201$, $N = 1 * 10^{19} cm^{-3}$, $V_{active} = 1.56 * 10^{-17} cm^3$ gives us $4.09\mu W$, if we pump this very high at $N = 3 * 10^{19} cm^{-3}$ then we could reach a power of $36.8\mu W$.

Model Performance			
f_T	E/b at 50Gbps	Max rate for $N = 1 \times 10^{19} cm^{-3}$	Max rate for $N = 3 \times 10^{19} cm^{-3}$
100 GHz	4.52 fJ/b	8.1 Gbps	32 Gbps
260 GHz	719 aJ/b	16 Gbps	59 Gbps
500 GHz	421 aJ/b	20 Gbps	90 Gbps
1000 GHz	345 aJ/b	22 Gbps	100 Gbps

Table 7.3 shows performance characteristics for four f_Ts, only 100GHz and 260GHz are currently attainable with current CMOS - 500GHz and 1000GHz represent potential future scaling of f_T. Their respective E/b at 50Gpbs as well as the maximum attainable data-rate given a transmit-side antenna-LED power limit at $N = 1 \times 10^{19} cm^{-3}$ and $N = 3 \times 10^{19} cm^{-3}$ are listed.

7.4 Increasing power

The link model suggests that as we approach 100GHz we will need close to $100\mu W$ of power. Note that this is the power emitted from the device before losses - it does not include the antenna efficiency and waveguide coupling efficiency. The following subsections summarize our attempts at increasing the power. For convenience the optical power equation without antenna efficiency and waveguide coupling efficiency is given below:

$$P_{antenna} = k \times \hbar\omega B_0 V_{active} F_{avg} N^2 \tag{7.1}$$

where k is the effective number of antenna-LEDs.

We see that for the single quantum well we would need $\sim 25\times$ power to close the link using this geometry. The following subsections summarize our attempts at increasing the power.

Increase number of antennas

One of the most obvious ways to increase the power is to simply have an array of devices and then combine them into the single mode waveguide. This could be accomplished by either

having multiple antenna-LEDs on the same waveguide or having multiple antenna-LEDs and using a power combiner to combine them into a single mode waveguide. Unfortunately, we quickly discovered that this is not as simple as it sounds. Due to conservation of energy and time reciprocity, we cannot cascade power-combiners to continually double the power. An intuitive explanation can be found from a simple 2x2 waveguide coupler:

$$\begin{bmatrix} TE_{out,Mode1} \\ TE_{out,Mode2} \end{bmatrix} = \begin{bmatrix} \sqrt{c} & jp\sqrt{1-c} \\ jp\sqrt{1-c} & \sqrt{c} \end{bmatrix} \begin{bmatrix} TE_{in,Top} \\ TE_{in,Bottom} \end{bmatrix} \tag{7.2}$$

where p is ± 1 with the two input modes $TE_{in,Top}$ and $TE_{in,Bottom}$ and two output modes $TE_{out,Mode1}$ and $TE_{out,Mode2}$. Due to energy conservation, any passive 2x2 coupling matrix will take this form (determinant is equal to 1). For incoherent combination, with a random phase relationship between the two branches, we can use superposition to determine the output power. There is no choice of c that would allow us to combine power from $TE_{in,Top}$ and $TE_{in,Bottom}$ in the $TE_{out,Mode1}$ or $TE_{out,Mode2}$ branch. Note that if we have a known phase relationship we can obtain coherent addition (ex. $TE_{in,Top} = -jTE_{in,Bottom}$ to get coherent addition in $TE_{out,Mode1}$).

This precludes us from combining incoherent light into an indistinguishable mode. However, we can still get more power into the output waveguide by altering either coupling to a different spatial mode $TE_{out,Mode2}$ or different wavelengths. The first method is not very practical, because even with zero insertion loss we would need more than 25 modes to get sufficient power in the bus waveguide - this would require a much higher complexity and footprint, and a larger photodetector would come with increased capacitance. The second method is to use different wavelengths - this can be seen by adding an explicit wavelength dependence:

$$\begin{bmatrix} TE_{out,Mode1}(\omega) \\ TE_{out,Mode2}(\omega) \end{bmatrix} = \begin{bmatrix} \sqrt{c(\omega)} & jp\sqrt{1-c(\omega)} \\ jp\sqrt{1-c(\omega)} & \sqrt{c(\omega)} \end{bmatrix} \begin{bmatrix} TE_{in,Top}(\omega_1) \\ TE_{in,Bottom}(\omega_2) \end{bmatrix} \tag{7.3}$$

Now we can get power combination by designing a coupling element where $c(\omega_1) = 1$ and $c(\omega_2) = 0$. For this toy example we can get twice as much power in the output waveguide. This idea was explored in [89], where if we have the same material spectrum then we can design high Q antennas to get pseudo wavelength selectivity. But for the same active region (material spectrum) this provides limited scaling of the power in the output waveguide by $1.5-3$ times. The same principles discussed for the 2x2 waveguide coupler apply for putting multiple antennas on the same waveguide - effectively it is reduced to the same problem.

Increase the active region volume

As shown in Chapter 3, one of the easiest ways to increase the power is to increase the active region volume, but this comes with a trade-off for modulation speed. In Table 7.1, we compare the maximum expected f_{3dB} and P_{rad} for different active region designs using the averaging from Table 3.1 and waveguide coupling structure from Table 4.1 and pumping to $N = 3E19$.

Active Region Design				
Design	F_{avg}	Volume	f_{3dB}	Power
Single quantum well	201	$1.58 \times 10^{-17} cm^3$	$\approx 192GHz$	$37\mu W$
Multiple quantum well (3x)	163	$4.75 \times 10^{-17} cm^3$	$\approx 156GHz$	$90\mu W$
Double heterostructure	68	$3.17 \times 10^{-16} cm^3$	$\approx 65GHz$	$247\mu W$
Double heterostructure ($w = 14nm$)	94	$2.26 \times 10^{-16} cm^3$	$\approx 90GHz$	$247\mu W$

Table 7.1: The projected average enhancement F_{avg}, active region volume V, f_{3dB}, and output power for different active region designs biased to a carrier concentration of $N = 3 \times 10^{19} cm^{-3}$ for a 20nm ridge width.

Switching to the double heterostructure would allow us to get significantly more power out of the active region, but comes with a significant speed trade-off. If we compare the maximum datarate for these powers we see that the single quantum well can reach 59Gbps, the multiple quantum well can reach up to 79Gbps, and the double heterostructure has enough power for 100Gbps. However, the double heterostructure with a width of 20nm lacks sufficient speed to create an efficient link at 100Gbps, so in the last row we provide the metrics for a 14nm ridge that has an f_{3dB} of 90GHz. However, for slow rates $< 80Gbps$, or for future scaling of f_T where less optical power will be required, the multiple quantum well has a good power-speed trade-off, making it a more optimal choice for active region design.

7.5 Integration and outlook

The final consideration is how we integrate the link with CMOS. So far, the analysis in this chapter has assumed perfect integration with CMOS - in other words, that the distance between the transistors and the transmitter/photodetectors is negligible. However, in reality there will be some additional integration capacitance. If we integrate this on top of the metal stack with a through oxide via then we pick up at least an additional 1.5fF on the transmitter and receiver [90], and if they are placed further away, requiring a through silicon via or additional routing, then we acquire at least 10fF of additional capacitance [91]. A comparison for the energy per bit assuming using 260GHz f_T at 50Gbps is shown in Fig. 7.3.

In this chapter we have shown the potential for $< 1fJ/bit$ optical interconnects using existing technology at 50Gbps; to reach higher data rates with low energy consumption would require FETs with higher f_T. However, these links can only achieve low energy consumption

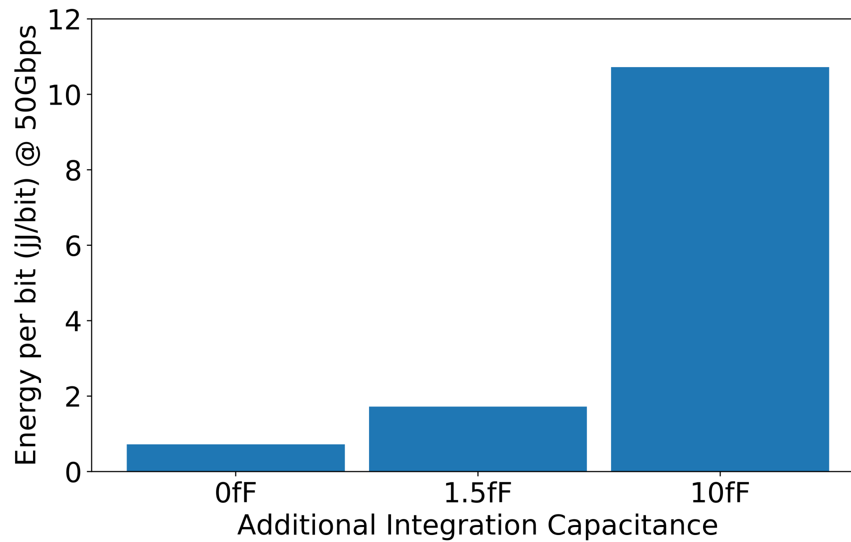

Figure 7.3: Energy per bit at 50Gbps with 260GHz f_T for additional integration capacitances. 1.5fF roughly corresponds to the self capacitance in a through oxide via.

if they are integrated closely with the transistors. In other words, we lose the benefit of having on-chip optical interconnects the further they are from our transistors. Creating these links will not be trivial, it will require a close collaboration between device, RF, system, and process engineers to fully realize the potential of the on-chip optical interconnects.

Chapter 8

Conclusion

In this dissertation we focused on the cavity backed slot antenna as a promising geometry for creating energy efficient, high speed, nanoscale on-chip optical interconnects. We presented theory for how to achieve efficient single mode waveguide coupling of the cavity backed slot antenna, including applying inverse design which allowed us to utilize a coupled cavity effect to decrease the metal loss. We fabricated optically-pumped waveguide coupled devices, which showed good agreement with theory and presented a pathway for achieving electrical injection. To improve the efficiency and therefore the power output, we developed a surface passivation process that achieved a record low surface recombination velocity, paving the way for efficient nanoscale devices. Additionally, we discussed how to account for all the dipoles in the device figures of merit, including showing that Lorentz reciprocity can be used to quickly probe the electromagnetic response of these devices, allowing for a significant reduction in computational resources with high accuracy. Finally, using our device figures of merit, the system was modeled and the receiver was optimized for each datarate. Using existing technology models, we showed it was possible to achieve $< 1fJ/bit$ energy consumption at high data rates. More work will need to be done to demonstrate such a link, including: design and demonstration of high speed electrical injection, integration with a low capacitance photodetector, further system design, and close integration with CMOS, but with the right collaboration of engineers this task should be feasible. The optical antenna-LED has been shown to be a very promising device, with a great deal of potential due to its high speed and efficiency at the nanoscale size.

Appendix A

Fabrication Notes

This appendix details some of the fabrication decisions; discussing some of the failed samples and what we did to mitigate that problem. While these have not been rigorously tested, we lay out some of our theories behind the failure and what we did to adjust the process.

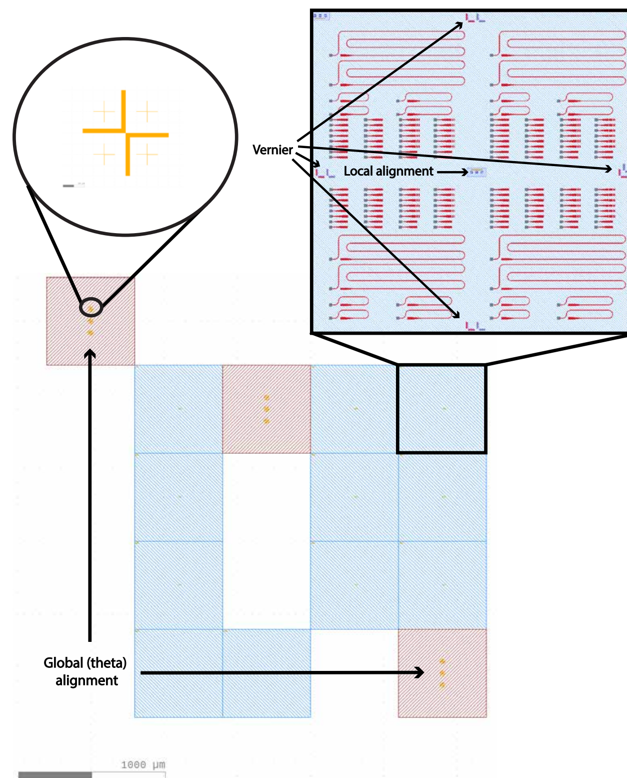

Figure A.1: GDS file layout of waveguide coupling design highlight global, local, and Vernier alignment marks.

A.1 Alignment Marks

The alignment mark mask consists of a pair of global theta alignment marks to correct for angle offsets, these should be placed at the extreme edges of your pattern to reduce angle errors. As shown in Fig. A.1, they consist of a large mark for easy location in the EBL SEM and a smaller mark to perform alignment on. The small mark has a similar structure to the large mark, but with an arm length of $5\mu m$ and width of 100nm. These are deposited using Cr/Au which has a high density which provides high contrast in the SEM with surrounding materials even when buried under 100's of nanometers of resist. The global (theta) alignment command in CABL is R2.

Then each 600μm field was manually aligned using a local alignment mark in the center of the field (CABL command R23). This way we have a global theta correction as well as local position correction. Through this combination we achieved very small alignment offsets ($< 20nm$) near the center of the field. Potentially due to theta alignment or deflection, alignment offsets grew towards the edge of the field. This could potentially be mitigated by using smaller fields.

A.2 Nanoscale Vernier Marks

To measure the alignment offset we added nanoscale vernier marks with a bar width of 300nm and spacing of 300nm for the top set and 320nm of the bottom set - allowing us to measure offsets on the order of $10 - 20nm$ in the scanning electron microscope. An example is shown in Fig. A.2. The y-offset in this image is less than the sensitivity of the Vernier marks ($< 10nm$) and the x-offset is estimated to be $\approx 20nm$.

A.3 Adhesion

We found that expired HSQ resist is particularly susceptible to poor adhesion, especially on Al_2O_3. For the LED-ridge patterning we found that we could mitigate this by exposing the sample to an 80W O2 plasma in a reactive ion etcher. Two potential explanations for improved adhesion are that the O2 plasma introduces nanoscale surface roughness which creates more surface area or the formation of an oxide that HSQ can better adhere to.

For the waveguide patterning, we found that a TiO_2 layer acts as a very effective adhesion layer. Without this layer the patterns would detach and be completely removed or float and reattach to other parts of the sample sometimes in random orientations. With these process modifications the adhesion issue was effectively solved for our process, allowing us to extend the lifetime of the expensive HSQ resists significantly.

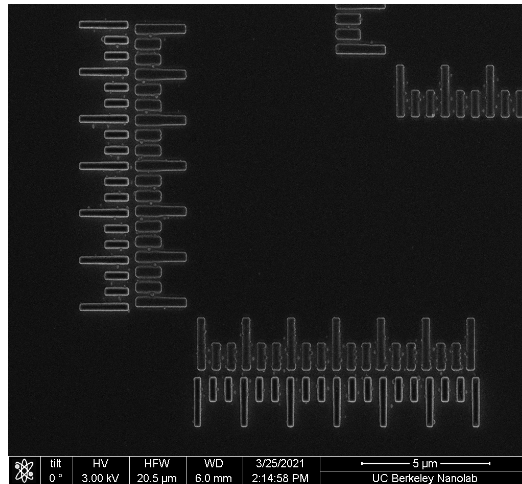

Figure A.2: SEM of vernier marks showing no y-offset and an x-offset of +20nm for the waveguide layer relative to the LED ridges. This vernier pair was on the right edge of the 600μm field.

A.4 High Contrast EBL

For our standard HSQ resist we use a $\approx 140nm$ thick XR-1541-6 (6% formulation). This can easily yield straight sidewalls with a number of different exposures or developers. However, when we use a less dilute flowable oxide (FOX) at $\approx 450nm$ we quickly run into problems with electron forward scattering in EBL. In order to mitigate this problem we switched our process from the 50keV acceleration voltage tool to the 100keV acceleration voltage to have less forward scattering; however, as shown in Fig. A.3, this problem was not fully resolved by changing the acceleration voltage alone. We still see a significant area around the waveguide ($\approx 30\mu m$) where the resist has not cleared - the thickness of this layer is nearly 180nm which makes it nontrivial to clear in the dry etch step.

So, we explored different processes for the resist development step - largely inspired from the work in [92]. We found that using a lower concentration ammonium hydroxide (NaOH) combined with a dilute salt for several minutes gave us nearly vertical sidewalls. Specifically, we used a 1:5 concentration of MF 351 Developer:DI (based on NaOH and boron sodium oxide), developed for 8 minutes – this process provided repeatable high contrast FOX-15 development. Higher concentrations of MF 351 led to quicker development but at the expense of a lower contrast process.

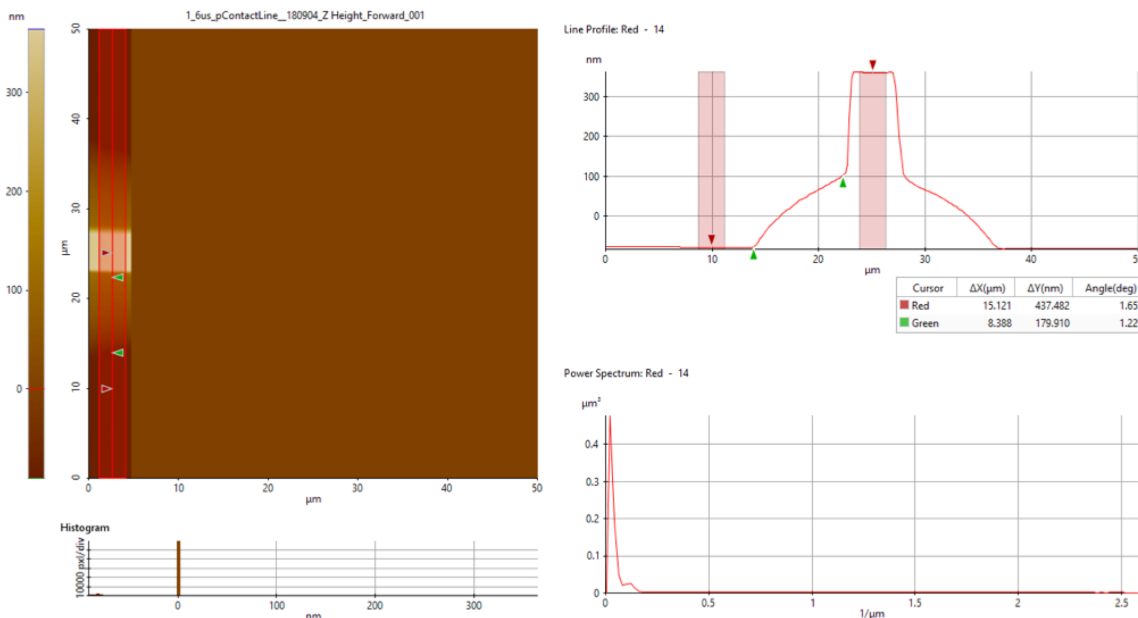

Figure A.3: Atomic force micrograph (AFM) taken after FOX-15 exposure in Crestec 130keV and development for 20s in OPD 4262 (dilute TMAH). Profile in top right clearly shows evidence of forward scattering near waveguide edges causing resist to be under-developed in that region.

A.5 Silver agglomeration

We found that one of the disadvantages of using silver is its tendency to agglomerate under low heat ($\gtrapprox 150C$) [93, 94] - this effect was shown for our devices in Fig. A.4. In planar films it was found that agglomeration can be prevent by depositing thicker films [93], but even for thick films ($\geq 250nm$) the silver morphology can still change [94] at temperature as low as 200C leading to higher loss.

Some of the prevailing methods to prevent agglomeration are capping with Al_2O_3 [95, 96], which has been found to lock in the silver grains preventing agglomeration or morphology changes temperature up to at least 400C [95] and extends their chemical stability [96]. Finally, in [97] it was shown that a metal layer can serve as an effective adhesion and encapsulation layer.

However, when we tested the Al_2O_3, as shown in Fig.A.4b, we found that the pump down of the chamber at 200C for 5 minutes was long enough that the silver agglomerated before encapsulation. This could perhaps be mitigated by a quicker pump down or lower deposition temperature. But, in an attempt to improve the quality of our film we decided

to encapsulate the silver in-situ by sputtering Nb immediately following silver sputtering. We found that this encapsulation method was able to prevent agglomeration at least up to 200C, which would leave us sufficient thermal budget to complete our processing. We saw no noticeable agglomeration in Fig.A.4c after a hotplate anneal at 200C for over an hour.

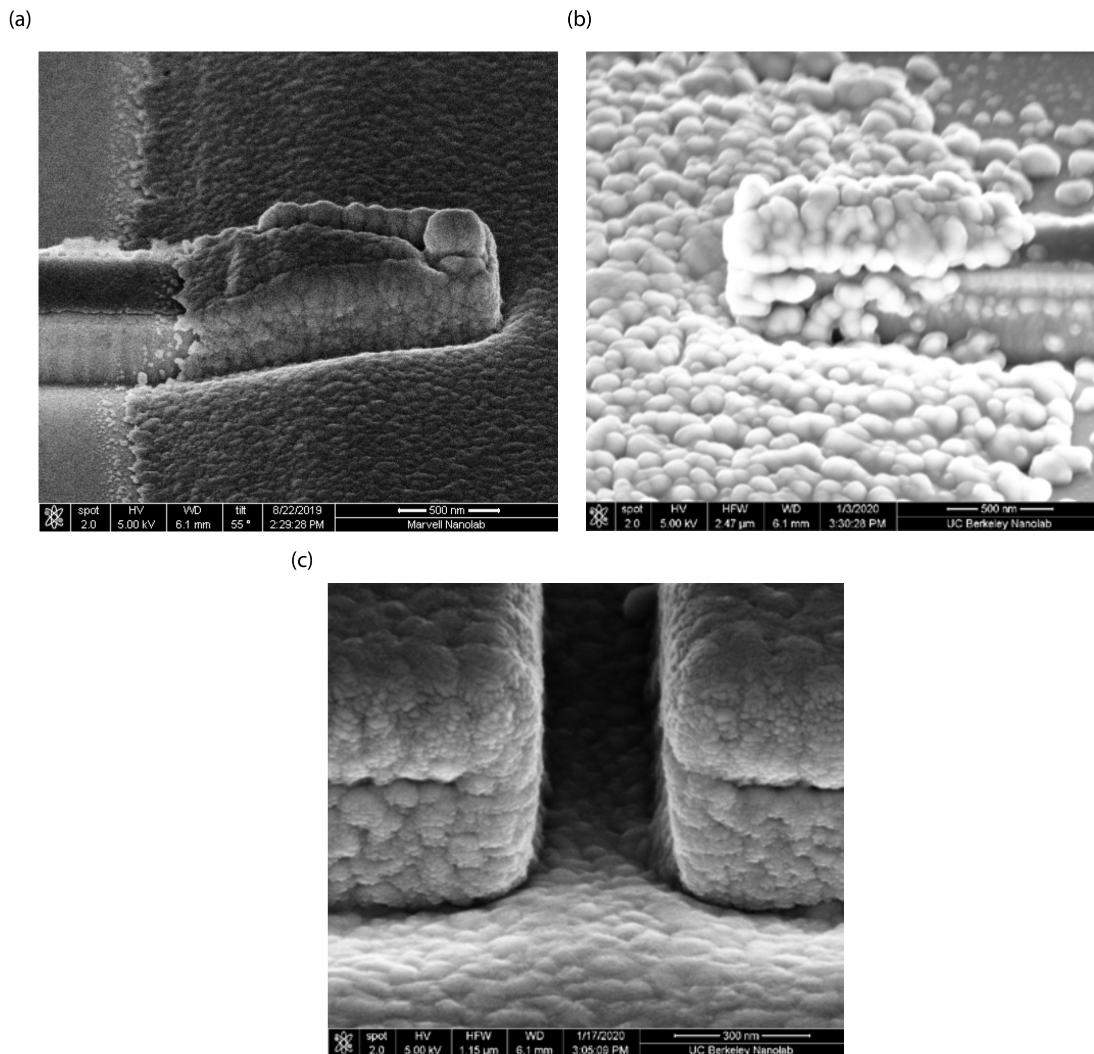

Figure A.4: Comparison of waveguide coupled LEDs (a) after liftoff, (b) after atomic layer deposition capping with Al_2O_3 at 200C showing agglomeration, and (c) sample capped with 70nm Nb and heated on hotplate to 200C for more than an hour showing no obvious signs of agglomeration.

A.6 Substrate removal

It is a little unusual to take such care to protect the surface then re-expose it in step (Fig. 5.2j) - in a fully optimized process this would not happen. However, this was done because a sample failed during substrate removal. We believe this failure was due to having a poor interface oxide between our InGaAsP etch stop and our InP waveguide. In this sample, the waveguide appeared to be intact after the wet etch removal of the substrate, but it was not until removal of the etch stop that we discovered that the InP waveguide was actually removed leading to an indent in the oxide where the waveguide was supposed to be. An example of this is shown in Fig. A.5 - where part of the InP was unintentionally etched during substrate removal.

As shown in the main process flow in Fig. 5.2, this was mitigated by removing the ALD oxide on the surface and redepositing a new thin Al_2O_3 to cover both the LED ridge and waveguide sidewalls under the metal region. In an optimized process we would want to change this step - starting with testing whether we need the Al_2O_3 on the sidewall of the waveguide under the Ag, or if this is unnecessary.

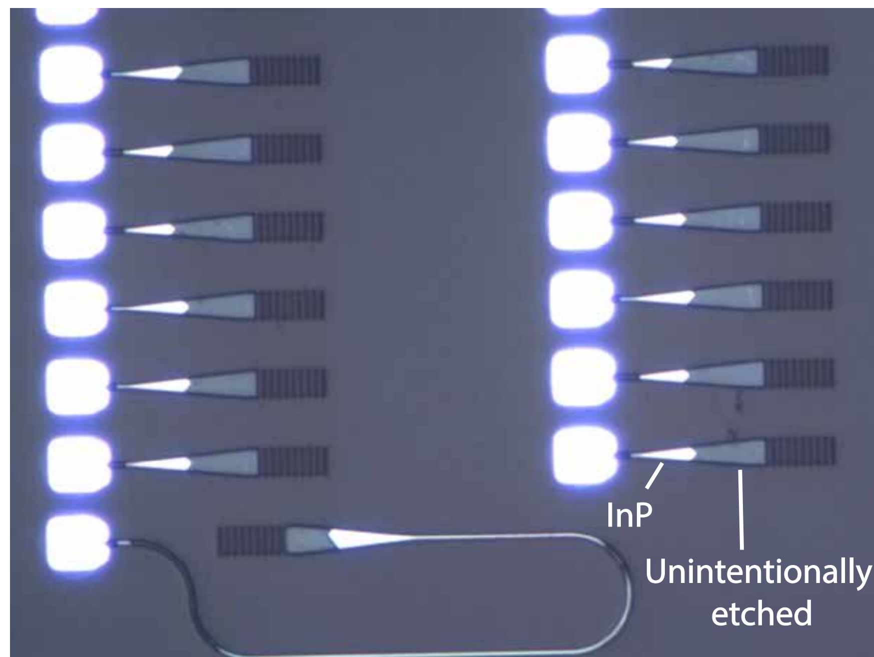

Figure A.5: Optical micrograph after substrate and etch stop removal for a device with no additional Al_2O_3 deposited showing InP waveguide was unintentionally etched.

Appendix B

Time Correlated Single Photon Counting

This appendix will describe the theory behind time correlated single photon counting (TC-SPC) as well as the practical implementation. Compared to other lifetime measurement techniques like pump-probe, streak cameras, or up-conversion, TCSPC is attractivedue to the large dynamic range, relatively low cost, and good timing resolution (<20ps).

B.1 Theory

TCSPC is based on the repetitive, precisely timed registration of single photons. These photons are timed to a corresponding excitation pulse - typically a laser, and dectector capable of measuring single photons, such as: a Photomultiplier Tube (PMT), Micro Channel Plate (MCP), a Single Photon Avalanche Diode (SPAD), or a superconducting nanowire. As shown in Fig. B.1, the difference between laser sync and received photon can be recorded and compiled in a histogram. The curve that is built up from the individual events represents the time decay that would be obtained from a typical single time-resolved measurement [98].

In order for the decay to be accurate the measurement needs to be single photon - i.e. the probability for detecting more than one photon per cycle needs to be low in order to avoid pile-up. The condition to avoid pile-up is:

$$CR_{adjusted} < 5\% \times f_{laser} \tag{B.1}$$

where $CR_{adjusted}$ is the adjusted count rate, and f_{laser} is the repetition frequency of the laser - and ideally, the count rate should be kept to $< 1\%$ the repetition rate. This condition can be achieved by inserting neutral density filters to cut down the signal until that equation is satisfied.

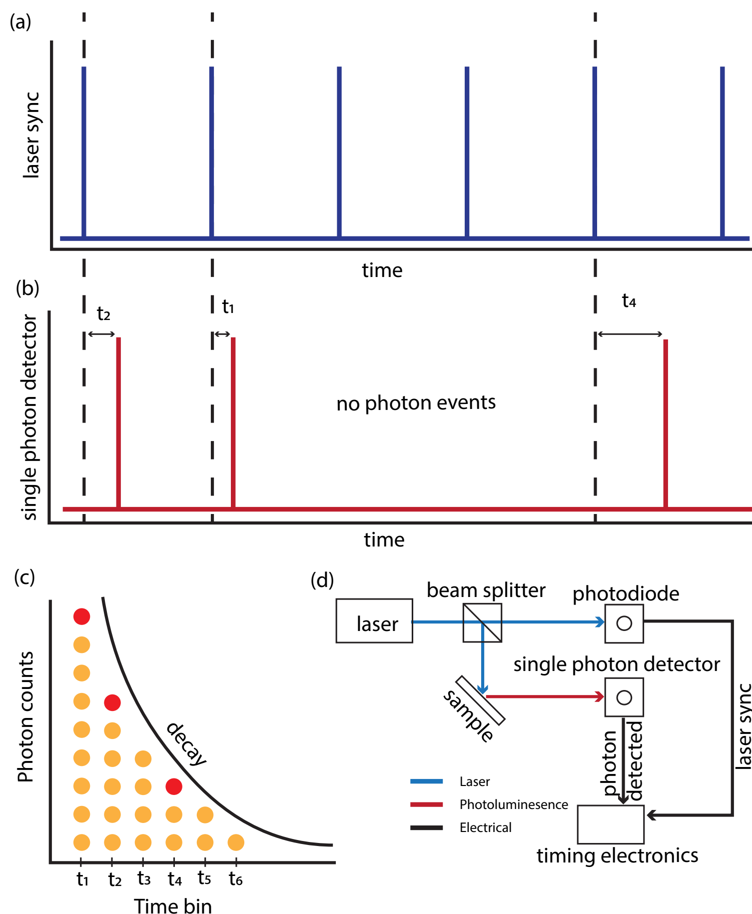

Figure B.1: (a) Laser pulse at fixed repetition rate. (b) Three different photon arrival times, binned at t_1, t_4, and t_2. (c) Histogram built up from binned photon arrival times. (d) Basic implementation of TCSPC. Drawing inspired by [28].

B.2 Avalanche photodiode

A SPAD is a p-n junction that is biased beyond the breakdown voltage in a meta-stable state so that when a single photon excites an electron-hole pair it causes an avalanche event - this has been historically called Geiger mode in the literature. The avalanche is quenched by biasing below breakdown, held-off for a short time period (known as the hold-off time), then the bias is restored in order to detect another photon.

The primary source of noise in InGaAs SPADs are from thermally generated carriers. This can be reduced by using a Peltier cooler to lower the temperature. In Si this is usually sufficient to run in so-called "free-running" mode; however, InGaAs SPADs need another modification to keep the dark counts low. The InGaAs SPADs are gated - or only biased past breakdown for a short time frame around the laser sync. When a photon is detected

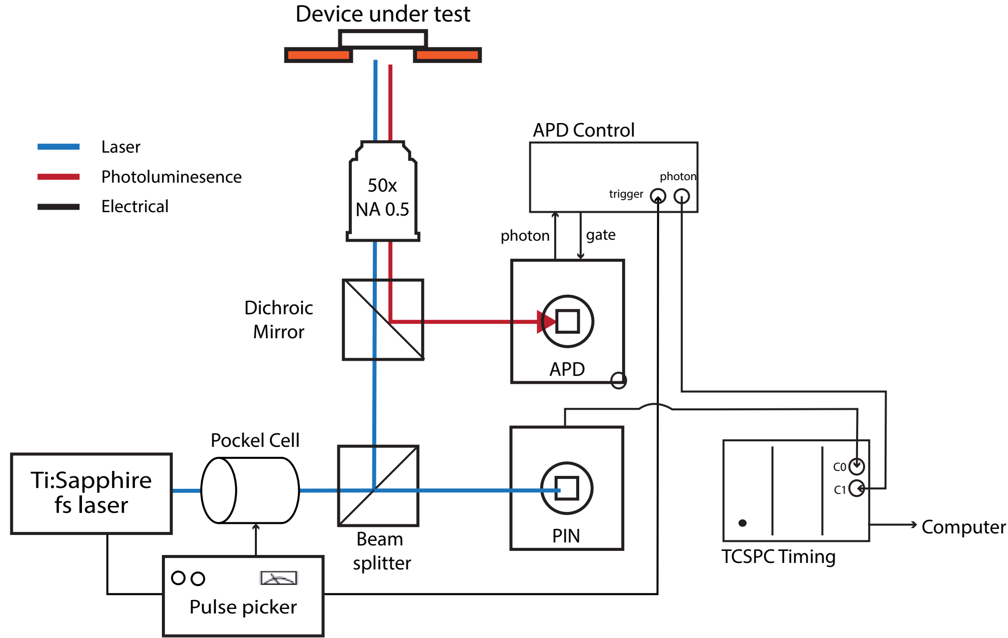

Figure B.2: Block diagram of the TCSPC optical setup at near-IR wavelengths. Drawing inspired by [28].

and the circuit is quenched, the gates during the hold-off time will be skipped leading to a longer dead time. As shown in Fig. B.2, the addition of the SPAD gating increases the complexity of the electronics slightly, we need to add an additional sync signal between the laser pulse picker and the SPAD gate.

We can calculate the effective count rate adjusting for the hold-off time and the gate time:

$$CR_{adjusted} = \frac{CR_{measured}}{1 - CR_{measured} \times T_{HO}} \frac{1}{T_{ON}} \tag{B.2}$$

where $CR_{measured}$ is the measured count rate, T_{HO} is the hold-off time, and T_{ON} is the gate time. The adjusted count rate can now be plugged into Eq. C.1 to compare to the laser repetition rate.

Biasing

There is an important trade-off around biasing. At larger excess bias (voltage past breakdown), the photon detection efficiency and dark count rates increase - leading to a lower signal to noise ratio. However, at large excess biases the timing jitter is reduced, leading to more accurate timing measurements for fast decays. So, for short lifetimes where timing is important we should use a large excess bias ($\approx 7V$) and hold-off time ($\approx 40\mu s$) to decrease

jitter while maintaining low dark counts. And, for long lifetimes where we have low photons we should use a low excess bias to increase the signal to noise ratio.

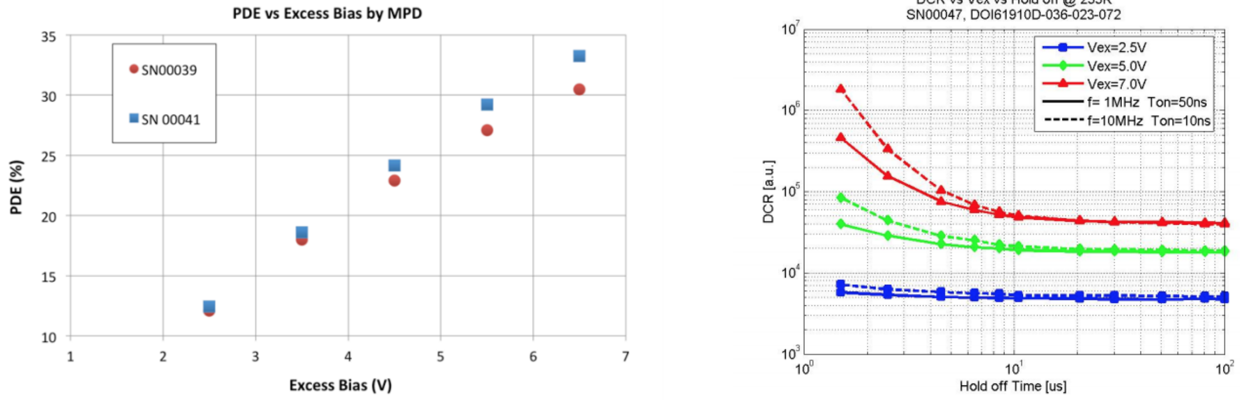

Figure B.3: Biasing trade-off - left: Photon detection efficiency as a function of excess bias and right: dark count rate as a function of hold-off time with three excess biases plotted. Reproduced from [99].

Afterpulsing and background subtraction

When a photon or dark count causes an avalanche event there is a probability of filling traps that can be released triggering additional pulses. This effect can be seen in the right side of Fig. B.3, if the detector is not held-off for a long enough period then it can increase the dark counts - this is shown for the $T_{HO} \approx 1 - 10\mu s$ range.

For extremely long lifetimes in our surface passivation devices we found that we had to account for the change of background with the addition of photons. We used the lowest noise setting ($V = 2.5V$ and $T_{HO} = 15\mu s$) we measured the change in the dark counts as a function of the increased counts (both increased dark counts and photons). This process is shown in Fig. B.4. The R^2 value for the best fit is 0.9989 - indicating that a linear fit is a good model for this photon range and detector settings.

B.3 Setup details

The sample was excited with the output of a Ti:Sapphire femtosecond laser (Coherent Ultra II). Part of the laser pulse was split off and sent to a fast PIN photodiode (Picoquant TDA 200) to trigger the TCSPC timing module (Picoquant Picoharp 300) and establish a time reference. A pulse picker (Conoptics) was used to reduce the repetition rate of the output of a Ti:Sapphire femtosecond laser from 80 MHz to 4MHz MHz to ensure the period between

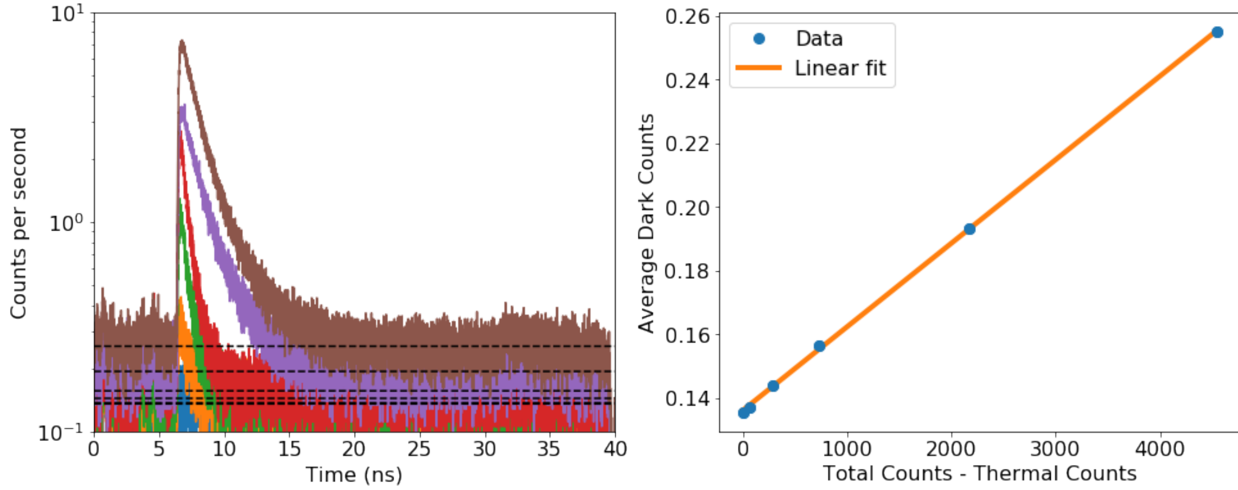

Figure B.4: Background subtraction for MPD InGaAs SPAD with $V = 2.5V$ and $T_{HO} = 15\mu s$ with laser rep rate of 4MHz. The power dependent dark count rate was determined by measuring a short lifetime device which decayed to the noise floor quickly. The dark count average was taken between 24-28ns with a 4ps bin (1000 samples). On the right we plot the dark count average as a function of the total counts minus the thermal dark counts (no photons).

successive laser pulses was much longer than the decay time of the sample. The sync output of the pulse picker was used trigger the gate and turn-on the InGaAs APD (Micro Photon Devices) in preparation for the arrival of a photon. The gating signal needs to be precisely timed to occur the same time that the laser pulse arrives at the sample; this was achieved through a variable length coaxial cable delay. If a photon was detected by the APD, the gating signal was immediately turned off. Otherwise after 40ns, if no photon was detected, the gate was turned off until the next laser pulse and the cycle was repeated again.

B.4 Auto-alignment

We found that the Pockel's cell had around a 10-16dB suppression, so the signal is essentially re-pumped at the full laser rep rate of 80MHz with a signal that is 16dB lower than the main peak. Experimentally, we found that the second laser re-pump would affect the signal decay (at 25ns). So instead of measuring long decays (>25ns) in a single shot we stitched multiple measurements together in 25ns windows. The optimal delay was found by minimizing the least squares error between the signals, as shown in Eq. B.3.

$$\min_{t_s} \frac{1}{n} \sum (y_1(t) - y_2(t - t_s)) \tag{B.3}$$

This process is shown in Fig. B.5, where three background-subtracted decay curves are collected the three different powers (shown in blue, orange, and green). The time delay between the blue and orange curves, t_{s1}, is calculated. We then interpolate to remove redundant time points and calculate the time delay between the combined curve and the green curve, t_{s2}. Again, the data is interpolated to remove redundant points and now the total decay curve can be fit.

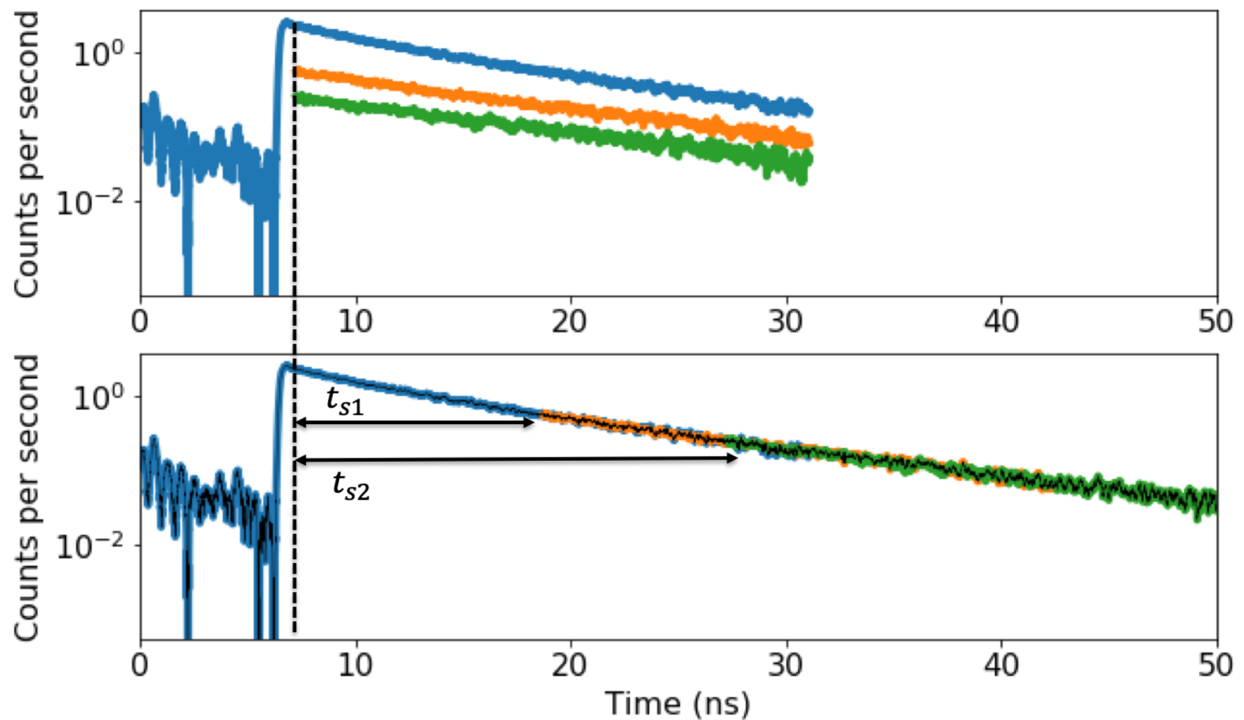

Figure B.5: Top: three background-subtracted decay curves for three separate pump powers. Bottom: optimal time delay calculated from least squares analysis.

Appendix C

Inverse Design

In this work we used the Berkeley Photonic Inverse Design package, originally described in [42]. The inverse design optimization problem that was solved can be written as the following:

$$\max_{\theta} \sum_{\omega} c_{\omega} T_{\omega}(x, r) : \text{Radius of Curvature} \geq 100\text{nm} \quad (9)$$

where θ denotes the optimization parameter space – which in this case is the interface between InP and Ag in the metal-optic waveguide coupler region, ω is an index that defines the frequency bandwidth of the optimization, T is the Poynting vector evaluated at positions r in the waveguide for electric and magnetic fields abbreviated by vector x, and c is a user-defined weight chosen for each frequency index. Finally, we included an optimization constraint on the radius of curvature to ensure fabricability. A brief discussion of the limitations of our inverse design implementation follow.

The objective function that was used in inverse design does not give individual control over our figures of merit, F_{avg}, η_{antenna}, and η_{WC}. Consequently, we included the weights, c, in the objective function to provide this control. An additional limitation comes in reference to Fig. 4.8(c) where the length of the metal along the coupler section sidewalls is not perturbed. Because it is undesirable to have metal along the sidewalls of the coupler section (XY plane) with a different length than the metal on top of the waveguide (XZ plane), the metal on top of the waveguide effectively constrained the designable region. Therefore, we used several metal lengths as initial conditions for inverse design optimization.

Lastly, one of the most important considerations for our choice of the waveguide coupler structure in Fig. 4.8 was its compatibility with top-down fabrication. In other words, because the entire ridge must share the same etch mask, it must also share the same 2D cross-sectional shape in the XY plane. Therefore, a geometrical constraint is required in the inverse design optimization to maintain the conformal nature of the ridge which is composed of several materials. Such a constraint was unavailable in our basic implementation of inverse design. We imposed this constraint ad hoc by updating the SOG-Ag and SiO_2-Ag interfaces

every three iterations to match the changing InP-Ag interface, but no significant convergence issues were encountered.

Bibliography

[1] Nicola Jones. "How to stop data centres from gobbling up the world's electricity". en. In: *Nature* 561.7722 (Sept. 2018). Number: 7722 Publisher: Nature Publishing Group, pp. 163–166. DOI: 10.1038/d41586-018-06610-y.

[2] *Cisco Annual Internet Report - Cisco Annual Internet Report (2018–2023) White Paper*. en.

[3] A. F. Benner et al. "Exploitation of optical interconnects in future server architectures". In: *IBM Journal of Research and Development* 49.4.5 (July 2005). Conference Name: IBM Journal of Research and Development, pp. 755–775. ISSN: 0018-8646. DOI: 10.1147/rd.494.0755.

[4] Chen Sun et al. "Single-chip microprocessor that communicates directly using light". en. In: *Nature* 528.7583 (Dec. 2015). Number: 7583 Publisher: Nature Publishing Group, pp. 534–538. ISSN: 1476-4687. DOI: 10.1038/nature16454.

[5] Nir Magen et al. "Interconnect-Power Dissipation in a Microprocessor". en. In: (), p. 7.

[6] D. A. B. Miller. "Device requirements for optical interconnects to silicon chips". In: *Proceedings of the IEEE* 97.7 (July 2009), pp. 1166–1185. ISSN: 0018-9219. DOI: 10.1109/JPROC.2009.2014298.

[7] Igor L. Markov. "Limits on fundamental limits to computation". en. In: *Nature* 512.7513 (Aug. 2014). Number: 7513 Publisher: Nature Publishing Group, pp. 147–154. ISSN: 1476-4687. DOI: 10.1038/nature13570.

[8] L. Schaper and D. Amey. "Improved Electrical Performance Required for Future MOS Packaging". In: *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* 6.3 (Sept. 1983). Conference Name: IEEE Transactions on Components, Hybrids, and Manufacturing Technology, pp. 283–289. ISSN: 1558-3082. DOI: 10.1109/TCHMT.1983.1136189.

[9] David A. B. Miller. "Attojoule Optoelectronics for Low-Energy Information Processing and Communications". en. In: *Journal of Lightwave Technology* 35.3 (Feb. 2017), pp. 346–396. ISSN: 0733-8724, 1558-2213. DOI: 10.1109/JLT.2017.2647779.

[10] Jan Rabaey, Anantha Chandrakasan, and Borivoje Nikolic. *Digital Integrated Circuits: A Design Perspective*. Pearson, 2002.

[11] J. D. Meindl. "Low power microelectronics: retrospect and prospect". In: *Proceedings of the IEEE* 83.4 (Apr. 1995). Conference Name: Proceedings of the IEEE, pp. 619–635. ISSN: 1558-2256. DOI: 10.1109/5.371970.

[12] Shun Lien Chuang. *Physics of Photonic Devices, 2nd Edition — Wiley*. en-us. 2nd. 2009.

[13] Suguru Yamaoka et al. "Directly modulated membrane lasers with 108 GHz bandwidth on a high-thermal-conductivity silicon carbide substrate". en. In: *Nat. Photonics* 15.1 (Jan. 2021), pp. 28–35. ISSN: 1749-4885, 1749-4893. DOI: 10.1038/s41566-020-00700-y.

[14] K. Ding and C. Z. Ning. "Fabrication challenges of electrical injection metallic cavity semiconductor nanolasers". en. In: *Semicond. Sci. Technol.* 28.12 (Nov. 2013), p. 124002. ISSN: 0268-1242. DOI: 10.1088/0268-1242/28/12/124002.

[15] E. Zielinski et al. "Excitonic transitions and exciton damping processes in InGaAs/InP". en. In: *Journal of Applied Physics* 59.6 (Mar. 1986), pp. 2196–2204. ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.336358.

[16] Edward Mills Purcell. "Proceedings of the American Physical Society". en. In: *Phys. Rev.* 69.11-12 (1946), pp. 674–674. ISSN: 0031-899X. DOI: 10.1103/PhysRev.69.674.

[17] Erwin K. Lau et al. "Enhanced modulation bandwidth of nanocavity light emitting devices". EN. In: *Opt. Express, OE* 17.10 (May 2009). Publisher: Optical Society of America, pp. 7790–7799. ISSN: 1094-4087. DOI: 10.1364/OE.17.007790.

[18] Constantine Balanis. *Antenna Theory Analysis and Design*. 3rd. 1982.

[19] W. Shockley. "Currents to Conductors Induced by a Moving Point Charge". In: *Journal of Applied Physics* 9.10 (Oct. 1938). Publisher: American Institute of Physics, pp. 635–636. ISSN: 0021-8979. DOI: 10.1063/1.1710367. URL: https://aip.scitation.org/doi/abs/10.1063/1.1710367 (visited on 05/13/2021).

[20] S. Ramo. "Currents Induced by Electron Motion". In: *Proceedings of the IRE* 27.9 (Sept. 1939). Conference Name: Proceedings of the IRE, pp. 584–585. ISSN: 2162-6634. DOI: 10.1109/JRPROC.1939.228757.

[21] Michael S. Eggleston et al. "Optical antenna enhanced spontaneous emission". en. In: *Proceedings of the National Academy of Sciences* 112.6 (Feb. 2015), pp. 1704–1709. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1423294112.

[22] René Kullock et al. "Electrically-driven Yagi-Uda antennas for light". en. In: *Nat Commun* 11.1 (Dec. 2020), p. 115. ISSN: 2041-1723. DOI: 10.1038/s41467-019-14011-6.

[23] Thang B. Hoang et al. "Ultrafast spontaneous emission source using plasmonic nanoantennas". eng. In: *Nat Commun* 6 (July 2015), p. 7788. ISSN: 2041-1723. DOI: 10.1038/ncomms8788.

[24] Jiani Huang et al. "Tailored Emission Spectrum of 2D Semiconductors Using Plasmonic Nanocavities". en. In: *ACS Photonics* 5.2 (Feb. 2018), pp. 552–558. ISSN: 2330-4022, 2330-4022. DOI: 10.1021/acsphotonics.7b01085.

[25] David Arbel et al. "Light emission rate enhancement from InP MQW by plasmon nano-antenna arrays". en. In: *Opt. Express* 19.10 (May 2011), p. 9807. ISSN: 1094-4087. DOI: 10.1364/OE.19.009807.

[26] Michael S. Eggleston et al. "Ultrafast spontaneous emission from a slot-antenna coupled WSe_2 monolayer". In: *ACS Photonics* 5.7 (July 2018), pp. 2701–2705. DOI: 10.1021/acsphotonics.8b00381.

[27] S. A. Fortuna et al. "Large spontaneous emission rate enhancement from an electrically-injected nanoLED coupled to an optical antenna". In: *2015 IEEE Photonics Conference (IPC)*. Oct. 2015, pp. 172–173. DOI: 10.1109/IPCon.2015.7323683.

[28] Seth Fortuna. "Integrated Nanoscale Antenna-LED for On-Chip Optical Communication". PhD thesis. EECS Department, University of California, Berkeley, Aug. 2017. URL: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-144.html.

[29] Jung-Hwan Song et al. "Fast and bright spontaneous emission of Er3+ ions in metallic nanocavity". In: *Nat Commun* 6 (May 2015). ISSN: 2041-1723. DOI: 10.1038/ncomms8080.

[30] Yingbo Zhao, Vivian Wang, and Ali Javey. "Molecular Materials with Short Radiative Lifetime for High-Speed Light-Emitting Devices". en. In: *Matter* 3.6 (Dec. 2020), pp. 1832–1844. ISSN: 2590-2385. DOI: 10.1016/j.matt.2020.09.009.

[31] Kevin C. Y. Huang et al. "Antenna electrodes for controlling electroluminescence". en. In: *Nature Communications* 3 (Aug. 2012), p. 1005. ISSN: 2041-1723. DOI: 10.1038/ncomms1985.

[32] Johannes Kern et al. "Electrically driven optical antennas". en. In: *Nature Photonics* 9.9 (Sept. 2015), pp. 582–586. ISSN: 1749-4893. DOI: 10.1038/nphoton.2015.141.

[33] Michael S. Eggleston and Ming C. Wu. "Efficient coupling of an antenna-enhanced nanoLED into an integrated InP waveguide". In: *Nano Lett.* 15.5 (May 2015), pp. 3329–3333. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.5b00574.

[34] V. Dolores-Calzadilla et al. "Waveguide-coupled nanopillar metal-cavity light-emitting diodes on silicon". en. In: *Nature Communications* 8 (Feb. 2017), p. 14323. ISSN: 2041-1723. DOI: 10.1038/ncomms14323.

[35] Myung-Ki Kim et al. "Engineering of metal-clad optical nanocavity to optimize coupling with integrated waveguides". EN. In: *Opt. Express, OE* 21.22 (Nov. 2013), pp. 25796–25804. ISSN: 1094-4087. DOI: 10.1364/OE.21.025796.

[36] Young-Ho Jin, Byoung Jun Park, and Myung-Ki Kim. "Extreme field enhancement in nano-gap plasmonic cavity via 90% efficient coupling with silicon waveguide". EN. In: *Opt. Express, OE* 24.22 (Oct. 2016), pp. 25540–25547. ISSN: 1094-4087. DOI: 10.1364/OE.24.025540.

[37] Andrew Michaels and Eli Yablonovitch. "Inverse design of near unity efficiency perfectly vertical grating couplers". EN. In: *Opt. Express, OE* 26.4 (Feb. 2018), pp. 4766–4779. ISSN: 1094-4087. DOI: 10.1364/OE.26.004766.

[38] Bing Shen et al. "An integrated-nanophotonics polarization beamsplitter with 2.4 2.4 μm^2 footprint". en. In: *Nature Photonics* 9.6 (June 2015), pp. 378–382. ISSN: 1749-4893. DOI: 10.1038/nphoton.2015.80.

[39] Louise F. Frellsen et al. "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides". EN. In: *Optics Express* 24.15 (July 2016), pp. 16866–16873. ISSN: 1094-4087. DOI: 10.1364/OE.24.016866.

[40] S. Bhargava and E. Yablonovitch. "Lowering HAMR near-field transducer temperature via inverse electromagnetic design". In: *IEEE Transactions on Magnetics* 51.4 (Apr. 2015), pp. 1–7. ISSN: 0018-9464. DOI: 10.1109/TMAG.2014.2355215.

[41] Alexander Y. Piggott et al. "Fabrication-constrained nanophotonic inverse design". En. In: *Scientific Reports* 7.1 (May 2017), p. 1786. ISSN: 2045-2322. DOI: 10.1038/s41598-017-01939-2.

[42] Christopher M. Lalau-Keraly et al. "Adjoint shape optimization applied to electromagnetic design". EN. In: *Opt. Express, OE* 21.18 (Sept. 2013), pp. 21693–21701. ISSN: 1094-4087. DOI: 10.1364/OE.21.021693.

[43] Andrew Michaels and Eli Yablonovitch. "Leveraging continuous material averaging for inverse electromagnetic design". EN. In: *Opt. Express, OE* 26.24 (Nov. 2018), pp. 31717–31737. ISSN: 1094-4087. DOI: 10.1364/OE.26.031717.

[44] Jesse Lu and Jelena Vučković. "Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes". EN. In: *Optics Express* 20.7 (Mar. 2012), pp. 7221–7236. ISSN: 1094-4087. DOI: 10.1364/OE.20.007221.

[45] Y. Elesin et al. "Time domain topology optimization of 3D nanophotonic devices". In: *Photonics and Nanostructures - Fundamentals and Applications* 12.1 (Feb. 2014), pp. 23–33. ISSN: 1569-4410. DOI: 10.1016/j.photonics.2013.07.008.

[46] Hugo M. Doeleman, Ewold Verhagen, and A. Femius Koenderink. "Antenna–cavity hybrids: matching polar opposites for Purcell enhancements at any linewidth". In: *ACS Photonics* 3.10 (Oct. 2016), pp. 1943–1951. DOI: 10.1021/acsphotonics.6b00453.

[47] Domenico D'Agostino et al. "Low-loss passive waveguides in a generic InP foundry process via local diffusion of zinc". EN. In: *Opt. Express, OE* 23.19 (Sept. 2015). Publisher: Optical Society of America, pp. 25143–25157. ISSN: 1094-4087. DOI: 10.1364/OE.23.025143.

[48] A. Vardi et al. "Nanoscale Mo Ohmic Contacts to III–V Fins". In: *IEEE Electron Device Letters* 36.2 (Feb. 2015). Conference Name: IEEE Electron Device Letters, pp. 126–128. ISSN: 1558-0563. DOI: 10.1109/LED.2014.2386311.

[49] M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh. "Empirical low-field mobility model for III–V compounds applicable in device simulation codes". en. In: *Journal of Applied Physics* 87.6 (Mar. 2000), pp. 2890–2900. ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.372274.

[50] Sadao Adachi. "III-V Ternary and Quaternary Compounds". en. In: *Springer Handbook of Electronic and Photonic Materials*. Ed. by Safa Kasap and Peter Capper. Springer Handbooks. Cham: Springer International Publishing, 2017, pp. 1–1. ISBN: 978-3-319-48933-9. DOI: 10.1007/978-3-319-48933-9_30.

[51] Zhihang Tong et al. "Surface Improvement of InAlAs/InGaAs InP-Based HEMT Through Treatments of UV/Ozone and TMAH". In: *IEEE Journal of the Electron Devices Society* 8 (2020). Conference Name: IEEE Journal of the Electron Devices Society, pp. 600–607. ISSN: 2168-6734. DOI: 10.1109/JEDS.2020.3000493.

[52] Yeon Sik Jung. "Study on texture evolution and properties of silver thin films prepared by sputtering deposition". In: *Applied Surface Science* 221.1 (Jan. 2004), pp. 281–287. ISSN: 0169-4332. DOI: 10.1016/S0169-4332(03)00888-2.

[53] Mark W. Knight et al. "Influence of excitation and collection geometry on the dark field spectra of individual plasmonic nanostructures". en. In: *Opt. Express* 18.3 (Feb. 2010), p. 2579. ISSN: 1094-4087. DOI: 10.1364/OE.18.002579.

[54] Min Hu et al. "Dark-field microscopy studies of single metal nanoparticles: understanding the factors that influence the linewidth of the localized surface plasmon resonance". en. In: *J. Mater. Chem.* 18.17 (2008), p. 1949. ISSN: 0959-9428, 1364-5501. DOI: 10.1039/b714759g.

[55] J. J. Mock et al. "Shape effects in plasmon resonance of individual colloidal silver nanoparticles". en. In: *The Journal of Chemical Physics* 116.15 (Apr. 2002), pp. 6755–6759. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1462610.

[56] Jack J. Mock, David R. Smith, and Sheldon Schultz. "Local Refractive Index Dependence of Plasmon Resonance Spectra from Individual Nanoparticles". en. In: *Nano Lett.* 3.4 (Apr. 2003), pp. 485–491. ISSN: 1530-6984, 1530-6992. DOI: 10.1021/nl0340475.

[57] Adam D. McFarland and Richard P. Van Duyne. "Single Silver Nanoparticles as Real-Time Optical Sensors with Zeptomole Sensitivity". en. In: *Nano Lett.* 3.8 (Aug. 2003), pp. 1057–1062. ISSN: 1530-6984, 1530-6992. DOI: 10.1021/nl034372s.

[58] C. Huang et al. "Gain, detuning, and radiation patterns of nanoparticle optical antennas". en. In: *Phys. Rev. B* 78.15 (Oct. 2008), p. 155407. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.78.155407.

[59] Yajie Jiang, Supriya Pillai, and Martin A. Green. "Realistic Silver Optical Constants for Plasmonics". en. In: *Scientific Reports* 6.1 (July 2016). Number: 1 Publisher: Nature Publishing Group, p. 30605. ISSN: 2045-2322. DOI: 10.1038/srep30605.

[60] Yifang Chen et al. "Fabrication of ferromagnetic nanoconstrictions by electron beam lithography using LOR/PMMA bilayer technique". en. In: *Microelectronic Engineering*. Proceedings of the 32nd International Conference on Micro- and Nano-Engineering 84.5 (May 2007), pp. 1499–1502. ISSN: 0167-9317. DOI: 10.1016/j.mee.2007.01.165.

[61] Guillaume Crosnier et al. "Subduing surface recombination for continuous-wave operation of photonic crystal nanolasers integrated on Silicon waveguides". EN. In: *Opt. Express, OE* 23.21 (Oct. 2015). Publisher: Optical Society of America, pp. 27953–27959. ISSN: 1094-4087. DOI: 10.1364/OE.23.027953.

[62] Gary Shambat et al. "Nanobeam photonic crystal cavity light-emitting diodes". In: *Appl. Phys. Lett.* 99.7 (Aug. 2011). Publisher: American Institute of Physics, p. 071105. ISSN: 0003-6951. DOI: 10.1063/1.3625432.

[63] Myung-Ki Kim, Amit M. Lakhani, and Ming C. Wu. "Efficient waveguide-coupling of metal-clad nanolaser cavities". EN. In: *Opt. Express, OE* 19.23 (Nov. 2011), pp. 23504–23512. ISSN: 1094-4087. DOI: 10.1364/OE.19.023504.

[64] Hyunseok Kim et al. "Monolithic InGaAs Nanowire Array Lasers on Silicon-on-Insulator Operating at Room Temperature". In: *Nano Lett.* 17.6 (June 2017). Publisher: American Chemical Society, pp. 3465–3470. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.7b00384.

[65] Nicolas M. Andrade et al. "Inverse design optimization for efficient coupling of an electrically injected optical antenna-LED to a single-mode waveguide". EN. In: *Opt. Express, OE* 27.14 (July 2019). Publisher: Optical Society of America, pp. 19802–19814. ISSN: 1094-4087. DOI: 10.1364/OE.27.019802.

[66] Kengo Nozaki et al. "InGaAs nano-photodetectors based on photonic crystal waveguide including ultracompact buried heterostructure". EN. In: *Opt. Express, OE* 21.16 (Aug. 2013). Publisher: Optical Society of America, pp. 19022–19028. ISSN: 1094-4087. DOI: 10.1364/OE.21.019022.

[67] X. Zhao, A. Vardi, and J. A. del Alamo. "Sub-Thermal Subthreshold Characteristics in Top–Down InGaAs/InAs Heterojunction Vertical Nanowire Tunnel FETs". In: *IEEE Electron Device Letters* 38.7 (July 2017). Conference Name: IEEE Electron Device Letters, pp. 855–858. ISSN: 1558-0563. DOI: 10.1109/LED.2017.2702612.

[68] A. Alian et al. "InGaAs tunnel FET with sub-nanometer EOT and sub-60mV/dec sub-threshold swing at room temperature". In: *Appl. Phys. Lett.* 109.24 (Dec. 2016). Publisher: American Institute of Physics, p. 243502. ISSN: 0003-6951. DOI: 10.1063/1.4971830.

[69] Lu Zhou et al. "Brief Review of Surface Passivation on III-V Semiconductor". en. In: *Crystals* 8.5 (May 2018), p. 226. DOI: 10.3390/cryst8050226.

[70] X. Miao and X. Li. “Scalable Monolithically Grown AlGaAs–GaAs Planar Nanowire High-Electron-Mobility Transistor”. In: *IEEE Electron Device Letters* 32.9 (Sept. 2011). Conference Name: IEEE Electron Device Letters, pp. 1227–1229. ISSN: 1558-0563. DOI: 10.1109/LED.2011.2160248.

[71] E. Yablonovitch et al. “Electron-hole recombination at the Si-SiO2 interface”. In: *Appl. Phys. Lett.* 48.3 (Jan. 1986). Publisher: American Institute of Physics, pp. 245–247. ISSN: 0003-6951. DOI: 10.1063/1.96570. URL: https://aip.scitation.org/doi/10.1063/1.96570 (visited on 05/11/2021).

[72] Dongchul Suh, Duk-Yong Choi, and Klaus J. Weber. “Al $_2$ O $_3$ /TiO $_2$ stack layers for effective surface passivation of crystalline silicon”. en. In: *Journal of Applied Physics* 114.15 (Oct. 2013), p. 154107. ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.4825258.

[73] Wenbo Ji et al. “Dip Coating Passivation of Crystalline Silicon by Lewis Acids”. en. In: *ACS Nano* 13.3 (Mar. 2019), pp. 3723–3729. ISSN: 1936-0851, 1936-086X. DOI: 10.1021/acsnano.9b01038.

[74] Der-Hsien Lien et al. “Electrical suppression of all nonradiative recombination pathways in monolayer semiconductors”. en. In: *Science* 364.6439 (May 2019). Publisher: American Association for the Advancement of Science Section: Report, pp. 468–471. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aaw8053.

[75] L. E. Black et al. “Effective Surface Passivation of InP Nanowires by Atomic-Layer-Deposited Al $_2$ O $_3$ with PO $_x$ Interlayer”. en. In: *Nano Letters* 17.10 (Oct. 2017), pp. 6287–6294. ISSN: 1530-6984, 1530-6992. DOI: 10.1021/acs.nanolett.7b02972.

[76] E. Yablonovitch et al. “Nearly ideal electronic properties of sulfide coated GaAs surfaces”. en. In: *Appl. Phys. Lett.* 51.6 (Aug. 1987), pp. 439–441. ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.98415.

[77] Koushik Banerjee et al. “Study of Short- and Long-Term Effectiveness of Ammonium Sulfide as Surface Passivation for InAs/GaSb Superlattices Using X-Ray Photoelectron Spectroscopy”. en. In: *Journal of Elec Materi* 39.10 (Oct. 2010), pp. 2210–2214. ISSN: 0361-5235, 1543-186X. DOI: 10.1007/s11664-010-1298-x.

[78] Min-Su Park et al. “Surface passivation and aging of InGaAs/InP heterojunction phototransistors”. In: *Journal of Applied Physics* 121.23 (June 2017). Publisher: American Institute of Physics, p. 233105. ISSN: 0021-8979. DOI: 10.1063/1.4986633.

[79] LA Coldren, SW Corzine, and MI Masanovic. *Diode lasers and photonic integrated circuits*. 1995.

[80] A. Higuera-Rodriguez et al. “Ultralow Surface Recombination Velocity in Passivated InGaAs/InP Nanopillars”. In: *Nano Lett.* 17.4 (Apr. 2017), pp. 2627–2633. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.7b00430.

[81] Yuchen Xing et al. "A novel model on time-resolved photoluminescence measurements of polar InGaN/GaN multi-quantum-well structures". en. In: *Scientific Reports* 7.1 (Mar. 2017). Number: 1 Publisher: Nature Publishing Group, p. 45082. ISSN: 2045-2322. DOI: 10.1038/srep45082.

[82] M.M. Winterkorn et al. "ATOMIC LAYER DEPOSITED ETCH STOP LAYERS FOR HYDROFLUORIC ACID". en. In: *2016 Solid-State, Actuators, and Microsystems Workshop Technical Digest*. Hilton Head, South Carolina, USA: Transducer Research Foundation, May 2016, pp. 133–136. ISBN: 978-1-940470-02-3. DOI: 10.31438/trf.hh2016.38.

[83] Deniz Turkay, Cagil Koroglu, and Selcuk Yerci. "Analysis of Field-Effect Passivation in Textured and Undiffused Silicon Surfaces". en. In: *Phys. Rev. Applied* 12.3 (Sept. 2019), p. 034026. ISSN: 2331-7019. DOI: 10.1103/PhysRevApplied.12.034026.

[84] Fa-Jun Ma et al. "Modelling and Simulation of Field-effect Surface Passivation of Crystalline Silicon-based Solar Cells". en. In: *Energy Procedia* 15 (2012), pp. 155–161. ISSN: 18766102. DOI: 10.1016/j.egypro.2012.02.018.

[85] R. S. Tuley and R. J. Nicholas. "Band gap dependent thermophotovoltaic device performance using the InGaAs and InGaAsP material system". en. In: *Journal of Applied Physics* 108.8 (Oct. 2010). Publisher: American Institute of PhysicsAIP, p. 084516. ISSN: 0021-8979. DOI: 10.1063/1.3488903.

[86] K. T. Settaluri et al. "First Principles Optimization of Opto-Electronic Communication Links". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 64.5 (May 2017). Conference Name: IEEE Transactions on Circuits and Systems I: Regular Papers, pp. 1270–1283. ISSN: 1558-0806. DOI: 10.1109/TCSI.2016.2633942.

[87] Kengo Nozaki et al. "Forward-biased nanophotonic detector for ultralow-energy dissipation receiver". In: *APL Photonics* 3.4 (Apr. 2018). Publisher: American Institute of Physics, p. 046101. DOI: 10.1063/1.5022074.

[88] Tetsuya Suemitsu. "InP and GaN high electron mobility transistors for millimeter-wave applications". en. In: *IEICE Electron. Express* 12.13 (2015), pp. 20152005–20152005. ISSN: 1349-2543. DOI: 10.1587/elex.12.20152005.

[89] Sean Hooten et al. "nanoLED Wavelength Division Multiplexer Analysis". en. In: *Conference on Lasers and Electro-Optics*. San Jose, California: OSA, 2019, FW3C.6. ISBN: 978-1-943580-57-6. DOI: 10.1364/CLEO_QELS.2019.FW3C.6.

[90] Chuan Xu and Kaustav Banerjee. "Physical Modeling of the Capacitance and Capacitive Coupling Noise of Through-Oxide Vias in FDSOI-Based Ultra-High Density 3-D ICs". In: *IEEE Transactions on Electron Devices* 60.1 (Jan. 2013). Conference Name: IEEE Transactions on Electron Devices, pp. 123–131. ISSN: 1557-9646. DOI: 10.1109/TED.2012.2227966.

[91] Ioannis Savidis et al. “Electrical modeling and characterization of through-silicon vias (TSVs) for 3-D integrated circuits”. en. In: *Microelectronics Journal* 41.1 (Jan. 2010), pp. 9–16. ISSN: 00262692. DOI: 10.1016/j.mejo.2009.10.006.

[92] Joel K. W. Yang and Karl K. Berggren. “Using high-contrast salty development of hydrogen silsesquioxane for sub-10-nm half-pitch lithography”. In: *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena* 25.6 (Nov. 2007). Publisher: American Institute of Physics, pp. 2025–2029. ISSN: 1071-1023. DOI: 10.1116/1.2801881.

[93] H. C. Kim, T. L. Alford, and D. R. Allee. “Thickness dependence on the thermal stability of silver thin films”. en. In: *Applied Physics Letters* 81.22 (Nov. 2002), pp. 4287–4289. ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.1525070.

[94] Jing Lv et al. “Thermal stability of Ag films in air prepared by thermal evaporation”. en. In: *Applied Surface Science* 253.17 (June 2007), pp. 7036–7040. ISSN: 0169-4332. DOI: 10.1016/j.apsusc.2007.02.058.

[95] Lingwei Ma et al. “Silver Nanorods Wrapped with Ultrathin Al2O3 Layers Exhibiting Excellent SERS Sensitivity and Outstanding SERS Stability”. en. In: *Scientific Reports* 5.1 (Oct. 2015). ISSN: 2045-2322. DOI: 10.1038/srep12890.

[96] G. Baraldi et al. “Preventing the Degradation of Ag Nanoparticles Using an Ultrathin a-Al $_2$ O $_3$ Layer as Protective Barrier”. en. In: *The Journal of Physical Chemistry C* 117.18 (May 2013), pp. 9431–9439. ISSN: 1932-7447, 1932-7455. DOI: 10.1021/jp401421m.

[97] Daniel Adams and T. L. Alford. “Encapsulated silver for integrated circuit metallization”. en. In: *Materials Science and Engineering: R: Reports* 40.6 (June 2003), pp. 207–250. ISSN: 0927-796X. DOI: 10.1016/S0927-796X(03)00025-1.

[98] Michael Wahl. “Time-Correlated Single Photon Counting”. en. In: (), p. 14.

[99] MPD. *InGaAs SPAD - gated*. June 2014. URL: http://www.micro-photon-devices.com/Docs/Datasheet/InGaAs_Datasheet.pdf.