

Subtractive Photonics in Bulk CMOS

Thesis by
Craig Edward Ives

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended August 31st, 2023

© 2024

Craig Edward Ives
ORCID: 0009-0006-3893-541X

All rights reserved

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Ali Hajimiri, for being an invaluable mentor during my time as a graduate student. He has supported me through all the highs and lows, encouraging me in the challenging times and celebrating with me in the successes. His advice, both in life and work, has always been helpful and insightful, and I have gained very much from his breadth and depth of knowledge. Moreover, he has created a lab in which creativity can flow freely and ambitious ideas are always welcome to the discussion table, which is exactly the kind of environment that nurtures breakthroughs. It's been a great privilege to study in CHIC lab, and I will always be grateful that Ali gave me the opportunity to do so.

I would like to next thank my candidacy and thesis committee members, Professor Azita Emami, Alireza Marandi, Changhuei Yang, and Constantine Sideris. I greatly appreciate their input on my work.

AROUTIN KHACHATURIAN has in many ways served as a second advisor to me during my time in CHIC lab. From the very beginning of my Ph.D., he has been kind, supportive, and instructive. He has always been willing to answer my questions, share his knowledge, and sacrifice his time to help me. I owe him many thanks for all that he has done for me and for all that I've learned from him. REZA FATEMI also tolerated many questions from me and provided much helpful advice and guidance. His willingness to engage in deep philosophical discussion led to many enjoyable late night conversations in our shared office, during which I never failed to learn something new. I am grateful to have been Reza's officemate and co-worker in photonics. ELLIOTT WILLIAMS never attended the photonics meetings or worked directly on a project with me, and yet we managed to forge a wonderful, lasting friendship. I cannot thank Elliott enough for all of his support, advice, camaraderie, and humor. His level-headed approach to work and life has always been refreshing, and I highly value his perspective, whether the conversation is deeply philosophical, personal, technical, or any other kind. MATAN GAL has been a formative influence on me. He was an excellent mentor, partly because of his incorrigible optimism, partly because of his knowledge and skill in the lab, and partly because his utterly rational perspective made it impossible to disagree with him. I learned a lot working with Matan and I am very thankful for all of his help and support. COSTIS welcomed me from the very beginning and was encouraging and supportive throughout the time we overlapped in CHIC lab. We had many great conversations, both in the lab and the

gym, and I owe him many thanks for always being willing to lend an ear and offer sage advice. Austin Fikes has always been a great labmate, and I thank him for his support and help over the years. I would like to thank Parham Khial as well for the interesting conversations and for everything I learned from him. Debjit Sarkar has spent a good amount of time working with me and I would like to thank him for all of the interesting discussions, as well as his patience and sense of humor. I appreciate the tolerance and good company of all my past officemates, Brian Hong, Noyan Akbar, Samir Nooshabadi, and David Baum. I would like to thank all of my CHIC and MICS lab colleagues Alex Pai, M. Reza Hashemi, Behrooz Abiri, Oren Mizrahi, Ailec Wu, Armina Khakpour, Volkan Gurses, Alex Ayling, Jesse Brunet, Alex White, Mohith Manohara, Cole Brabec, Saransh Sharma, and Fatima Aghlmand for all of their help and the wonderful conversations over the years.

I am very grateful to my family for their support and encouragement throughout my life and Ph.D. studies. My wife Rachel undoubtedly deserves the greatest thanks for her endless patience and support. Words cannot express my gratitude.

ABSTRACT

Much of humanity's technological advancement over the last few decades may be attributed to exponentially increasing computing power, the bedrock of which is bulk CMOS technology. Exponentially increasing data rates in communications have also played an important role, facilitated by advancements in fiber optics and integrated photonics. However, efforts to capitalize on the complementary strengths of these two domains by merging them, an idea first envisioned almost 40 years ago, have so far proven inadequate. All previous attempts to integrate photonics in bulk CMOS have required either expensive process modification or resulted in waveguides with high loss.

In this thesis, we discuss our investigations of a new method of integrating photonics into bulk CMOS, which we call the method of subtractive photonics. This method entails forming waveguides out of the back-end interconnect of an electronic chip. The interconnect metal is designed to wrap around dielectric channels such that when the metal is etched away, suspended dielectric waveguides remain. Although this method introduces a large, previously untapped design space, since there are many interconnect layers that can be used in photonic structures, it also introduces certain severe constraints. This thesis explores some of the possibilities this design space opens up, as well as some of the challenges involved in designing photonics in a process intended only for electronics. As part of this exploration, we demonstrate waveguides with an upper bound on loss that is significantly lower than the best previously published waveguide loss for unmodified bulk CMOS. We also demonstrate the first measurements of waveguide loss at visible and near-visible wavelengths in unmodified bulk CMOS, as well as the first measurements of waveguide coupled photodiodes in unmodified bulk CMOS. These proof-of-concept results may pave the way towards fully integrated electronic-photonic systems in unmodified bulk CMOS.

PUBLISHED CONTENT AND CONTRIBUTIONS

- [1] C. Ives and A. Hajimiri, “Subtractive Photonic Waveguide-Coupled Photodetectors in 180nm Bulk CMOS,” in *2023 IEEE International Solid- State Circuits Conference (ISSCC)*, 2023. DOI: 10.1109/ISSCC42615.2023.10067724,
C.I. designed and measured the chips.
- [2] C. Ives, D. Sarkar, and A. Hajimiri, “Subtractive Photonics in Bulk CMOS,” *IEEE Journal of Solid-State Circuits*, pp. 1–14, 2023. DOI: 10.1109/JSSC.2023.3308915,
C.I. designed and measured the chips.
- [3] M. Gal-Katziri, C. Ives, A. Khakpour, and A. Hajimiri, “Optically Synchronized Phased Arrays in CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 57, no. 6, pp. 1578–1593, 2022. DOI: 10.1109/JSSC.2021.3136787,
C.I. designed the CMOS photodiode and TIA, managed the photonics side of the measurements, and contributed photodiode analysis.
- [4] R. Fatemi, C. Ives, A. Khachaturian, and A. Hajimiri, “Subtractive photonics,” *Optics Express*, vol. 29, no. 2, pp. 877–893, 2021. DOI: 10.1364/OE.410139,
C.I. assisted with layout, developed the post-processing procedure, and measured the chips.
- [5] M. Gal-Katziri, C. Ives, A. Khakpour, and A. Hajimiri, “A 28 GHz Optically Synchronized CMOS Phased Array with an Integrated Photodetector,” in *2021 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2021, pp. 111–114. DOI: 10.1109/RFIC51843.2021.9490452,
C.I. designed the CMOS photodiode and TIA and managed the photonics side of the measurements.

TABLE OF CONTENTS

Acknowledgements	iii
Abstract	v
Published Content and Contributions	vi
Table of Contents	vi
List of Illustrations	viii
List of Tables	xix
Chapter I: Introduction	1
1.1 Historical Background	1
1.2 Methods of Photonic Integration	9
1.3 Overview of BEOL Technology	20
Chapter II: The Subtractive Photonics Concept	28
2.1 Introduction to the Concept	28
2.2 TO0	31
2.3 Risk Mitigation	39
2.4 Waveguide Optimization	41
Chapter III: Designs and Experimental Results	61
3.1 Risk Mitigation in 180 nm	62
3.2 Etching Procedure	75
3.3 TO1	75
3.4 TO2	88
3.5 TO3	97
3.6 TO4	116
3.7 TO5	132
3.8 TO6	132
Chapter IV: Photodiodes and Photodiode Couplers	150
4.1 General Considerations	150
4.2 Tapering Waveguide-Photodiode Coupler	159
4.3 Reflecting Waveguide-Photodiode Coupler	166
Chapter V: Optical Reference Distribution	173
Chapter VI: Conclusion	179
6.1 Overview of Thesis Work	179
6.2 Future Outlook	181
Bibliography	184

LIST OF ILLUSTRATIONS

<i>Number</i>	<i>Page</i>
1.1 Brief timeline highlighting events germane to the development of photonics in bulk CMOS.	8
1.2 Tradeoff tree demonstrating the benefits of photonics in bulk CMOS.	20
1.3 Tradeoff tree considering the benefits of subtractive photonics against other methods of integrating photonics into bulk CMOS.	21
1.4 Outline of the aluminum interconnect fabrication process. The steps are detailed in the text.	23
1.5 Outline of the copper interconnect fabrication process. The steps are detailed in the text.	25
2.1 Opposing workflows for developing devices that can be mass manufactured in CMOS foundries.	29
2.2 The method of subtractive photonics. Metal surrounding the waveguide is etched in post-processing, leaving suspended dielectric waveguides.	30
2.3 Terminology for the waveguides used in this thesis.	31
2.4 Depiction of the interconnect layers of the 65 nm process.	32
2.5 Layout of the 65nm test chip.	33
2.6 Example of fitting a photonic structure with curves to the minimum Manhattan grid.	34
2.7 Initial etching tests using Aluminum Etch Type A.	35
2.8 A waveguide fib cut. Top: Diagram depicting the location of the FIB cut relative to the V7 grating coupler. Middle: SEM images of a grating coupler cavity with a FIB opening for exposing the buried waveguide. These are images of the same location but with varying contrast. Bottom: Close-up images of the buried waveguide. SEM images courtesy A. Khachaturian.	36
2.9 Images demonstrating the clean removal of the barrier layer.	37

2.10	Initial etching tests. Top left: Grating couplers after initial etching in AETA. Top right: Grating couplers after tests with chromium and titanium etchants. Bottom left: Grating couplers after 1 hour in a 1:1 H ₂ O ₂ /EDTA solution at 60°C. Bottom right: Grating couplers after 1 hour in a refreshed 1:1 H ₂ O ₂ /EDTA solution at 60°C.	38
2.11	Images of the top, middle, and bottom portions of the V7-FIB structure.	39
2.12	Illustration of etching progress. Left: Test structures after 1 hour in a 1:1 H ₂ O ₂ /EDTA solution at 60°C. Right: Test structures after 1 hour in a refreshed 1:1 H ₂ O ₂ /EDTA solution at 60°C.	51
2.13	Results of expedited etching. Top: Test structures after 1 hour in AETA at 80°C. Bottom left: Test structures after 1 hour in 1:1 H ₂ O ₂ /EDTA at 80°C. Bottom right: Test structures after another hour in AETA at 80°C.	52
2.14	Some results from a three part etching sequence. Top left: Aluminum pad peeling off after 2 hours in AETA at 80°C. Right: Test structures after an additional hour in AETA at 80°C. Bottom left: An exposed V7 grating coupler after 2 hours in H ₂ O ₂ /EDTA at 40°C.	53
2.15	A partially exposed grating coupler demonstrating the nature of the barrier layer. SEM image courtesy A. Khachaturian.	53
2.16	Partially etched grating couplers demonstrating the thinness of the barrier layer. SEM images courtesy A. Khachaturian.	54
2.17	A collapsed grating coupler formed from the V7 layer. SEM image courtesy A. Khachaturian.	54
2.18	Remnants of grating couplers formed using the M7 layer as a ridge and the V6 layer as a slab. Left: A partially etched cavity with the support remnants visible against the back wall of the cavity. Right: A fully etched cavity with the V6 slab remnants visibly collapsed onto the bottom of the cavity. SEM images courtesy A. Khachaturian.	55
2.19	A suspended channel waveguide. SEM images courtesy A. Khachaturian.	55
2.20	The gap between device and system research groups.	56
2.21	The risks that the subtractive photonics project is exposed to, and means of mitigating those risks.	56
2.22	A symmetric slab waveguide, with refractive index variation only in \hat{x} .	57
2.23	The fundamental mode of a 1D waveguide with a 3 μm wide slab at a wavelength of 1550 nm. The index of the core is 1.495 and the index of the cladding is 1.466.	57

2.24	Example of a rib waveguide that is single mode with an infinite slab, but multimode with a finite slab. Note that for the case of the finite slab, many more modes are supported than are shown.	58
2.25	Imaging a waveguide facet by scanning a fiber across it.	58
2.26	Definition of slab extension and the two waveguide widths used in mode overlap calculations.	59
2.27	Optimization of slab extension for a V5M5-V4M4 waveguide.	59
2.28	Modeling a bend using an index transformation. Left: A bent waveguide has an index profile given by n_{eq} . Phase velocity increases radially in the bend in order for the wavefront at ϕ_1 to arrive in phase at ϕ_2 . Right: The bend is modeled using a straight waveguide with a refractive index gradient to compensate for the increasing phase velocity.	60
2.29	Bend loss of a V5M5-V4M4 waveguide.	60
3.1	Depiction of the interconnect layers of the 180 nm process.	62
3.2	Example of the distortion caused by modeling the full dielectric stackup. The light blue regions denote low-k dielectric.	62
3.3	Schematic cross section of the M6-V5M5V4 waveguide geometry.	64
3.4	Confined modes of the M6-V5M5V4 waveguide at a wavelength of 1550 nm and a ridge width of 5 μm	64
3.5	Normalized effective index of the confined modes of the M6-V5M5V4 waveguide across ridge widths at a wavelength of 1550 nm.	65
3.6	Illustration of the enclosure rules. The metal layers must extend beyond the vias.	67
3.7	Mitigating DRC errors. Left: Breakdown of the DRC errors in T00. Right: Illustration of the corrugation used on vertical surfaces to mitigate enclosure violations.	67
3.8	Illustration of the gridding algorithm. Left: A polygon corner before gridding. Right: The same polygon corner after gridding. Note the acute angle that must be removed by other means (Fig. 3.9).	68
3.9	Illustration of the polygon rounding procedure. Left: A polygon corner before rounding. Center: The same polygon corner after rounding. Right: The same polygon corner after gridding. Note the absence of acute angles.	69
3.10	An example layout showing waffling on the V5 layer, along with a 3D model of waffling.	70

3.11	Defining waveguide facets. Top: Waveguide layout for dicing followed by polishing. Bottom: Waveguide layout for obtaining a waveguide facet defined by lithography by dicing and then etching.	71
3.12	Placing the dice line. Top: A dice line that is too close to the waveguide facet. Bottom: Choosing the dice line to fit within a lensed fiber's working distance.	72
3.13	Simulated fiber coupling efficiency for all confined modes of a M6-V5M5V4 waveguide and a 1550 lensed fiber with a 5 μm spot diameter. The ridge width of the waveguide is 5 μm	73
3.14	Comparison of dicing and polishing. Left: Waveguides before etching but after dicing. Right: Waveguides before etching but after dicing and polishing.	73
3.15	Designing a test structure for ease of measurement. Top: A test structure with edge couplers at 90° to each other. Bottom: A test structure with measurement fibers lying in the same axis.	74
3.16	Designing a test structure to ensure waveguide coupling. Top: A straight test structure with in-line measurement fibers. Bottom: A S-shaped test structure for ensuring only confined light is measured.	75
3.17	Layout of TO1.	77
3.18	A TO1 chip in mid-etch. Left: A partially etched M6-V5M5V4 waveguide. Middle: Partially etched mechanical test structures. Right: Displaced wafling.	78
3.19	SEM image of an etch gauge. SEM image courtesy A. Khachaturian.	79
3.20	A fully etched M6-V5M5V4 waveguide. Left: Focus on the waveguide facet. Right: Focus on the reflection of the waveguide facet.	79
3.21	SEM images of a fabricated M6-V5M5V4 waveguide. SEM images courtesy A. Khachaturian.	80
3.22	SEM images of waveguide wafling. Left: A M6-V5M5V4 waveguide with a partially removed piece of wafling. Right: Close view of the wafling. SEM images courtesy A. Khachaturian.	81
3.23	Layout and micrograph of the M6-V5M5V4 mechanical test structures.	81
3.24	Measurement of TO1. Top: Experimental setup used for measuring insertion loss of the waveguides. Bottom: Image of 635 nm light propagating in the waveguides.	82

3.25	Measurements of M6-V5M5V4 waveguides with varying path lengths and bend radii. The bottom diagrams indicate the locations of the test structures in TO1.	83
3.26	Schematic cross section of the V5M5V4-M4V3 waveguide geometry.	84
3.27	Confined modes of the V5M5V4-M4V3 waveguide at a wavelength of 1550 nm and a ridge width of 5 μm	84
3.28	Simulated fiber coupling efficiency for all confined modes of a V5M5V4-M4V3 waveguide and a 1550 lensed fiber with a 5 μm spot diameter. The ridge width of the waveguide is 5 μm	85
3.29	Normalized effective index of the confined modes of the V5M5V4-M4V3 waveguide across ridge widths at a wavelength of 1550 nm. . .	85
3.30	SEM images of a fabricated V5M5V4-M4V3 waveguide. SEM images courtesy A. Khachaturian.	86
3.31	Microscope and SEM images of V5M5V4-M4V3 waveguides. SEM images courtesy A. Khachaturian.	87
3.32	Microscope images of structures designed to test the limits of support width for the V5M5V4-M4V3 geometry.	88
3.33	Microscope images of structures designed to test the limits of support length for the V5M5V4-M4V3 geometry.	88
3.34	Layout of TO2.	90
3.35	The design of waffling on all via layers. Left: Exploded view of the via layers used in waveguide construction. Right: Depiction of the metal layers of a waveguide cavity, with waffling on all via layers. . .	91
3.36	The residue of waffling on all layers.	91
3.37	Measuring M6-V5M5V4 waveguides in TO2. Left top: Location of the tested waveguides in TO2. Left bottom: Measurements of insertion loss. Right: Microscope image of the ridge damage causing the excessive insertion loss of the longest waveguide of chip 1.	93
3.38	Upper bounds on waveguide loss derived from the insertion loss of each waveguide.	94
3.39	Schematic cross section of the V5M5-V4M4V3 waveguide geometry used in TO2.	94
3.40	Normalized effective index of the confined modes of the V5M5-V4M4V3 waveguide across ridge widths at a wavelength of 780 nm.	95

3.41	Confined modes of the V5M5-V4M4V3 waveguide at a wavelength of 780 nm and a ridge width of 4 μm	95
3.42	Simulated fiber coupling efficiency for all confined modes of a V5M5-V4M4V3 waveguide and a 780 lensed fiber with a 2 μm spot diameter. The ridge width of the waveguide is 4 μm	96
3.43	Transmission of the fundamental mode of a V5M5-V4M4V3 waveguide at 780 nm, assuming a support period of 100 μm , based on the extension of the waveguide slab beyond the 4 μm ridge width.	96
3.44	Bend loss comparison for several variations of the V5M5-V4M4V3 geometry, where the bend radius is defined relative to the area used by the standard bend. The standard bend refers to a waveguide bend with a slab on both sides of the ridge.	97
3.45	Additional waveguide measurements in TO2. Upper: Locations of the measured waveguides. Lower left: Insertion loss of V5M5-V4M4V3 singlemode waveguides. Lower right: Insertion loss of M6-V5M5V4 waveguides with 300 μm support spacing and 50 μm long supports.	98
3.46	The etching and polymer deposition test structure after etching.	99
3.47	Layout of TO3.	100
3.48	Microscope images of waffling in TO3. Note the transparency of the barrier layer in the lower left, and note the individual vias in the waffling image in the lower right.	102
3.49	Example waveguide cross section.	103
3.50	Microscope images showing peeling of glass overhangs.	103
3.51	Example of an interlayer transition that tapers one layer at a time. Upper left: Transition diagram. Upper right: Side view of an FDTD simulation of the transition. Bottom: Transmission as the interlayer coupler length is varied.	105
3.52	Example of an interlayer transition that tapers two layers at a time. Upper left: Transition diagram. Upper right: Side view of an FDTD simulation of the transition. Bottom: Transmission as the interlayer coupler length is varied.	106
3.53	Transition diagrams depicting the interlayer transitions taped out in TO3.	107
3.54	Microscope images showing waffling residue, peeling of sidewall corrugation, and barrier layers.	108

3.55	Waffling and interlayer couplers. Top: Cross section depicting the problem caused by using the waffling algorithm on an interlayer coupler. Bottom: Images of V5 tapering and the resulting stuck waffling.	109
3.56	Examination of barrier layers on an interlayer coupler.	110
3.57	Close view of a waveguide taper to illustrate the layer dependency of the barrier.	111
3.58	Examination of barrier layers on an interlayer coupler.	112
3.59	Microscope photo demonstrating that the barrier layer is constrained by both the M3 and the V4 layers.	113
3.60	Close view of a waveguide taper to illustrate the layer dependency of the barrier.	114
3.61	Examination of barrier layers on an interlayer coupler. The black boxes in the bottom cross section denote the place where the barrier layer exists.	115
3.62	Layout of TO4.	117
3.63	Structure of the M6-V5M5V4 waveguide used in TO4.	118
3.64	Images of the M6-V5M5V4 spiral. Left: Microscope photo. Right: The spiral illuminated with 635 nm light.	119
3.65	Etching comparison of the M6-V5M5V4 spiral. Left: The spiral after etching in AETA for 4 hours at 80°C. Right: The spiral after additional etching in AETA and H ₂ O ₂ /EDTA.	119
3.66	Block diagram of the spiral measurement setup.	119
3.67	Edge coupling setup with a 1550 nm lensed fiber.	120
3.68	Upper bound on the loss of the M6-V5M5V4 waveguide for various wavelengths and chip samples. The dashed lines refer to the waveguide loss reported in Orcutt <i>et al.</i> [47] and Hung <i>et al.</i> [55].	121
3.69	Table of the values plotted in Fig. 3.68.	121
3.70	Block diagram of the V5M5-V4M4V3 spiral measurement setup.	122
3.71	A beam with a central mass suspended above an electrode.	122
3.72	Elastic force and electrostatic force demonstrating the end of stability at the pull-in voltage.	123
3.73	Cross section of the coupling portion of the MEMS directional coupler.	124
3.74	The V5 and M2 layers of the MEMS directional coupler. Both the tapeout geometry and its complement are shown to highlight both the metal and dielectric structures.	125
3.75	A visualization of the oxide structure left after metal etching.	126

3.76	Micrograph of the transition region of the MEMS modulator. Note that the waffling has not been removed from the waveguides.	127
3.77	Cross section of the coupling portion of the MEMS directional coupler.	128
3.78	Optical coupling lengths of the class 1 MEMS modulator based on gap size.	128
3.79	Displacement of the electrodes of the class 1 MEMS modulator with applied voltage.	129
3.80	Optical coupling lengths of the class 2 MEMS modulator based on gap size.	129
3.81	Displacement of the electrodes of the class 2 MEMS modulator with applied voltage.	130
3.82	Optical coupling lengths of the class 3 MEMS modulator based on gap size.	130
3.83	Displacement of the electrodes of the class 3 MEMS modulator with applied voltage.	131
3.84	Microscope image of a MEMS modulator bent upward.	132
3.85	Demonstration of waveguide ridges removed with the surrounding waffling, both on modulators (top) and photodetectors (bottom). . . .	133
3.86	Layout of TO6.	134
3.87	System diagram of the proposed coherent receiver.	135
3.88	Mode confinement of the V5M5-V4M4 waveguide at 780 nm.	136
3.89	Comparison of waveguide geometries. Top: Cross section of the V5M5-V4M4V3 waveguide used in the spiral of TO4, which had adequate ridge adhesion. Bottom: Cross section of the V5M5-V4M4 waveguide used in the MEMS modulators of TO4, which had insufficient ridge adhesion.	137
3.90	Modes of the V5M5-V4M4 waveguide with a ridge width of 3 μm at 780 nm.	137
3.91	Simulated fiber coupling efficiency of a 780 nm lensed fiber with a spot size of 2 μm coupling to the confined modes of a V5M5-V4M4 waveguide.	138
3.92	A visualization of the V5M5-V4M4 directional coupler, with the inset image showing a close view of the shared slab.	139

3.93	MMI splitter optimization. Top: A view of the EME simulation setup using the full dielectric stack. Middle: EME simulation of the MMI splitter at its optimal length of 138 μm . Bottom: Plot of the EME optimization of the MMI splitter.	141
3.94	A visualization of the V5M5-V4M4 MMI splitter.	142
3.95	System diagram of the coherent receiver using an MMI hybrid. . . .	143
3.96	MMI hybrid optimization. Top: A view of the EME simulation setup using the full dielectric stack. Middle: EME simulations of the MMI hybrid at its optimal length of 1.11 mm. Bottom left: Plot of the EME optimization of the MMI hybrid length. Bottom right: Mode convergence of the EME simulation at the optimal length of 1.11 mm.	144
3.97	A visualization of the quadrature output of the V5M5-V4M4 MMI hybrid.	145
3.98	Microscope image of V5M5-V4M4 waveguides after barrier layer etching. Most of the waveguides show no evidence of a ridge, while the one remaining ridge has slipped from its designed position. . . .	146
3.99	Microscope images of waveguides submerged in etchant. This chip underwent 3.5 hours of AETA etching at 80°C to remove the aluminum, and then 3 hours of H ₂ O ₂ /EDTA etching at room temperature with no stirring. These images show several waveguides on this chip while they are still in the H ₂ O ₂ /EDTA. The red arrows indicate the position of the remaining ridges.	147
3.100	Microscope images of a waveguide region throughout the etching process. Upper left (step 1): Before etching. Upper right (step 2): After 9 hours of AETA at 80°C. Lower left (step 3): After 2 hours of H ₂ O ₂ /EDTA at room temperature. Lower right (step 4): After 6 hours of H ₂ O ₂ /EDTA at room temperature.	148
3.101	Waveguide geometry comparison. Top: V5M5-V4M4V3 waveguide used in TO4 - no ridge problems. Middle: V5M5-V4M4 waveguide used in TO6 - ridge removed during post-processing. Bottom: V5M5V4-M4V3 waveguide used in TO6 - no ridge problems.	149
4.1	TCAD simulation of a bulk CMOS photodiode.	150
4.2	Responsivity of each of the three CMOS photodiode junctions across the wavelengths of 635, 780, and 850 nm.	154
4.3	Estimated junction bandwidths due to transit time, capacitance, and diffusion current.	156

4.4	Comparison of the response of each junction of a CMOS photodiode to increasing reverse bias.	157
4.5	3D model of the implemented tapering waveguide-photodiode couplers.	159
4.6	Location of the CMOS photodiode in the tapering coupler.	160
4.7	Tapers in the photodiode coupler. Top: Ridge taper. Bottom: Slab taper.	160
4.8	Coupling region between the waveguide and the photodiode isolation glass.	161
4.9	Dimensions of the small photodiode and tapering coupler.	162
4.10	Measurements of the small photodiode. Top: Die photo with the small photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.	163
4.11	Setup for imaging the waveguide facets by scanning a lensed fiber. . .	164
4.12	V5M5-V4M4V3 waveguide facet imaged using the small photodiode.	164
4.13	Setup for measuring the responsivity of the small photodiode independent of the waveguide.	165
4.14	Determining coupling loss for the small photodiode. Top: FDTD simulation of coupling into the small photodiode. Bottom: Comparison of waveguide coupling and vertical coupling.	166
4.15	Dimensions of the large photodiode and tapering coupler.	167
4.16	Measurements of the large photodiode. Top: Die photo with the large photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.	168
4.17	V5M5-V4M4V3 waveguide facet imaged using the large photodiode.	168
4.18	Determining coupling loss for the large photodiode. Top: FDTD simulation of coupling into the large photodiode. Bottom: Comparison of waveguide coupling and vertical coupling.	169
4.19	Dimensions of the segmented photodiode and tapering coupler. . . .	170
4.20	Measurements of the segmented photodiode. Top: Die photo with the segmented photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.	171
4.21	Reflecting coupler simulations. Top: FDTD simulation setup and cross section. Middle: Transmission of the reflecting coupler across wavelengths. Bottom: Die photo of the reflecting coupler and connected waveguide.	172
5.1	An example of an RF phased array with both optical (red) and electrical (blue) reference distribution.	173

5.2	Overview of the optically synchronized phased array. An optical timing signal (red) is distributed to subarrays, each of which receive the reference, amplify it, and distribute it to eight TX channels. . . .	174
5.3	The 65 nm bulk CMOS photodiode. Top: Layout and die photo. Bottom: Measurement of the photodiode responsivity with bias. . . .	175
5.4	The injection locked TIA. Note that V_r provides reverse bias to the photodiode.	176
5.5	Two views of the board setup used to couple the fiber to the RFIC. The copper pipe is held in place by the lower board, and does not touch the upper board. The fiber passes through the pipe and then through a hole cut in the upper board that is filled with index matching gel. The fiber is then fixed to the lower board and the pipe.	176
5.6	Beam pattern measurement setup.	177
5.7	Measured beam patterns of the phased array module using optical reference distribution.	177
5.8	Demonstration of synchronization of two phased array modules. . . .	178

LIST OF TABLES

<i>Number</i>	<i>Page</i>
1.1 Comparison of Bulk CMOS Integration Methods	19
2.1 List of tapeouts.	31
3.1 Classes of MEMS directional couplers.	128
4.1 Assumed doping concentrations	151
4.2 Depletion region widths	152
4.3 Absorption coefficients and penetration depths in silicon (from [96]).	152
4.4 Ratio of ionization coefficients for each junction under an applied voltage V_A	157

Chapter 1

INTRODUCTION

1.1 Historical Background

The purpose of this thesis is to study methods of integrating photonics into bulk CMOS, focusing especially on suspended dielectric waveguides in the back end of bulk CMOS processes. As an intertwining of integrated photonics and integrated electronics, this work has its roots in both fields. Some important milestones are highlighted in Fig. 1.1.

The Groundwork for Silicon Electronics and Photonics

In December 1947, the point-contact transistor was demonstrated by John Bardeen and Walter Brattain [1], which was closely followed by the proposal for the bipolar junction transistor by William Shockley in January 1948 [2]. An important development in coherent optical technology closely followed the invention of the transistor, which was the first demonstration of holography by Dennis Gabor in 1948 [3]. A decade later in 1958, Jack Kilby demonstrated the first integrated circuit [4], closely followed in 1959 by Robert Noyce's proposal for a practical integrated circuit using deposited metal interconnect [5]. The next year, 1960, the MOSFET was demonstrated [6].

Although holography had been demonstrated over a decade earlier, the age of coherent optics did not take off until 1960, when the first laser was demonstrated by Maiman [7], [8]. This achievement set off a flurry of research activity in optics, with the first semiconductor lasers demonstrated in 1962 [9]–[11], p-n junctions analyzed as waveguides by Yariv in 1963 [12], glass fiber proposed as the medium for long distance communication in 1966 [13], and extensive research on thin film optics culminating in a proposal for integrated optics published in 1969 [14]. In the world of electronics, Carver Mead invented the MESFET in 1965 [15], [16], and Intel was founded in 1968 [17], beginning a relentless march in the improvement of digital electronics that is still ongoing. As a result, by the end of the 1960's, the basic ideas underlying the work of this thesis had been set out.

Three years after Intel's founding, in 1971, the first commercial microprocessor was released, the Intel 4004 [17], [18]. That same year, capitalizing on recent developments in both III-V electronics and photonics, Yariv considered uniting the

two fields through monolithic integration on a single chip [19]. Later in the decade, as silicon technology progressed, the need arose for an efficient method of designing large, complex digital systems. This need was met by Carver Mead, who created a method to automate chip layout, and authored an influential textbook on very large scale integrated (VLSI) circuits with Lynn Conway in 1978 [20]. As the size of digital systems continued to grow, it was foreseen that optical interconnects could benefit VLSI systems, and several options for implementing these optical interconnects were proposed by Goodman in 1984 [21]. One of these options was the use of optical waveguides on a VLSI chip, which seems to be the first suggestion of monolithic integration of photonics and silicon electronics.

The Dawn of Silicon Photonics

38 years ago, in the year 1985 (which was 38 years after the invention of the transistor), another revolution began in photonics, albeit one that started slowly and quietly, with the publication of the first paper on silicon photonics by Richard Soref and Lorenzo [22]. The authors demonstrated a rib waveguide whose index contrast was formed by growing a lightly doped waveguiding layer on top of a heavily doped substrate, and then etching a ridge. In this paper the authors also make the first suggestion of modulating light using the plasma dispersion effect.

The following year Soref and Lorenzo followed with a much more comprehensive paper [23], in which they clearly stated their motivation for pursuing waveguides in silicon: "The project was motivated by two considerations: 1) many of the processes developed for the Si electronic circuit industry can be applied to Si optical devices, and 2) high-speed Si electronic circuits can be combined monolithically with Si guided-wave devices in an optoelectronic integration." These reasons constitute two of the three typically cited justifications for silicon photonics:

- CMOS infrastructure can be leveraged to improve yields and reduce manufacturing costs.
- The high index contrast enables small photonic structures, and thus large photonic systems.
- There is potential for the monolithic integration of photonics with silicon electronics.

Soref and Lorenzo were evidently not concerned about the benefits of high index contrast, since their waveguides 1) used a rib geometry for lateral confinement, and

2) used a difference in doping concentration for vertical confinement. As a result, the first waveguides in silicon exhibited low index contrast. Despite the poor optical confinement, however, it is clear that monolithic integration with electronics has been a fundamental part of the vision of silicon photonics from the very beginning.

In the same paper, Soref and Lorenzo consider many of the most pertinent aspects of the silicon photonics platform. They consider amorphous silicon and polysilicon as waveguide materials, but note that these materials are "far less desirable for waveguiding than single-crystal Si" [23] due to the high optical loss. The options for single-crystal silicon include "epitaxial silicon-on-insulator (SOI) technology, ion implantation to form a low-index buried layer, and epitaxial growth of silicon on silicon," with the authors ultimately choosing epitaxial growth of silicon for their first demonstrations [23]. Their demonstration waveguides use a multimode rib structure, and exhibit a loss of 15-20 dB/cm. Although far superior to the expected losses of polysilicon waveguides, this loss still left much to be desired, especially with the benefit of hindsight. The achievement of these loss values represented a stepping stone to better results, however, which is a precedent one might keep in mind when considering the results of this thesis.

The authors also consider five possible methods of optical modulation, introducing them with the following telling comments: "The crystal lattice of Si is centrosymmetric; therefore, Si does not exhibit the linear electrooptic effect (the Pockels effect). For this reason, some people would argue that Si is not an active material, and is therefore not genuinely useful in integrated optics. We do not share this pessimistic view. On the contrary, there are several optical switching mechanisms that have the potential to provide active switching directly in silicon." [23] ¹ The listed mechanisms include the plasma dispersion effect using electrical generation of carriers for modulation; the plasma dispersion effect using optical generation of carriers; the Kerr effect, which is extremely fast, but also weak in silicon; the real part of electroabsorption due to the Franz-Keldysh effect, termed electrorefraction, which is a method possible in a SiGe platform; and the method of acousto-optic modulation.

Although Soref and Lorenzo consider the possible benefits of a SiGe platform, which would provide a photodetection mechanism in the infrared, they also mention the possibility of detection of 1.3 μm light through Schottky photodiodes, using the

¹It is interesting to note the parallels between the lack of Pockels effect in silicon, which led to pessimism regarding its usefulness as a photonic material, and the lack of intrinsic high speed modulation mechanisms in glass, which potentially explains the dearth of research into glass waveguides on CMOS.

mechanism of internal photoemission. For the waveguides explored in this thesis, which are transparent in both the visible and near IR, and for which no deposition of germanium can be relied upon to facilitate photodetection, the possibility of detection in the near IR using Schottky photodiodes is prescient. Section 6.2 explores this further.

In the following year, 1987, Soref and Bennett published their oft-cited paper on electro-optic effects in silicon [24]. This paper has served as the bedrock of silicon photonic design ever since it was published, mainly due to the inclusion of a simple linear relation between changes in carrier concentration and changes in refractive index and absorption, paving the way to high-speed plasma dispersion modulators in later years. Moreover, in a subtle footnote Soref and Bennett point out that submicron channel waveguides enable both singlemode operation at $1.3 \mu\text{m}$ and "practical voltages" for carrier depletion modulators. The authors also devote space to other modulation mechanisms, previously noted in the 1986 Soref and Lorenzo paper [23], [24].

To summarize, in these three papers Soref, Lorenzo, and Bennett:

- Propose leveraging CMOS infrastructure to improve yields and reduce manufacturing costs.
- Propose monolithic integration of silicon waveguides with silicon electronics.
- Propose multimode and singlemode silicon waveguides using silicon-on-silicon, silicon-on-insulator, and silicon-on-sapphire technologies.
- Consider and reject polysilicon waveguides due to likely high losses.
- Consider waveguides made in a SiGe alloy material system.
- Demonstrate rib waveguides in silicon-on-silicon with losses of 15-20 dB/cm.
- Propose five modulation mechanisms in silicon, investigating plasma dispersion in detail.
- Propose detection in the near IR using Schottky photodiodes.

Parallel Developments in the Era of Silicon Photonics

In the same year (1987) that Soref and Bennett published their landmark paper, the Taiwan Semiconductor Manufacturing Company (TSMC) was founded [25]. The

creation of TSMC kicked off the present era of fabless semiconductor companies, in which design is separated from manufacturing, allowing both processes to be independently optimized. As future developments would later show, this would prove to be a monumental change in the semiconductor industry.

Almost a decade after silicon photonics began, polysilicon waveguides began to be investigated in earnest at MIT in 1994 [26], with initial waveguide losses clocking in at 77 dB/cm [27], [28]. This work had two motivations: 1) polysilicon waveguides could be deposited on the back-end of a CMOS chip to avoid using expensive area that could otherwise be devoted to transistors, and 2) waveguide deposition could enable multilayer photonic systems [26], [27]. Subsequent effort led to a reduction in waveguide loss to 15 dB/cm in 1996 [29], matching the loss of the first waveguides demonstrated by Soref and Lorenzo [23].

Another revolution in the world of integrated electronics occurred soon after, with IBM publishing the first demonstration of a chip with dual damascene copper interconnect in 1997 [30], the culmination of an industry-wide effort to develop a next-generation interconnect [31]. Most prominently for this thesis, TSMC introduced its 180 nm process with aluminum interconnect in 1998 [32]. TSMC's 65 nm process with copper interconnect would debut in 2005 [33].

In 2001, the photonics company Luxtera was founded to realize the full potential of silicon photonics by developing a monolithic electronic-photonic SOI process [34], [35]. Intel was also studying silicon photonics at this time, and in 2004 Intel demonstrated the first silicon photonic modulator with a bandwidth greater than 1 GHz [36]. This development marked a sea-change in the silicon photonics community, with many research groups subsequently publishing modulator designs with bandwidths greater than 1 GHz [37]. The fruits of that race can be seen in contemporary silicon photonic modulators, which can reach bandwidths of 50 GHz [38].

In 2006, two papers were published at SPIE dedicated to exploring the possibilities of electronic-photonic integration. The first was part of a DARPA funded effort to develop an electronic-photonic integrated circuit (EPIC), and demonstrated crystalline SOI waveguides with a loss of 0.35 dB/cm at 1550 nm, as well as deposited polysilicon waveguides with a loss of 4 dB/cm at 1550 nm [39]. The second explored three different methods of integration: hybrid integration, monolithic FEOL integration, and monolithic BEOL integration [40], coming to similar conclusions as those discussed in Section 1.2. The deposited silicon nitride waveguides exhibited a loss of

2 dB/cm at 1300 nm and 6 dB/cm at 1550 nm, while the deposited amorphous silicon waveguides exhibited a loss of 5 dB/cm at 1300 nm and 4 dB/cm at 1550 nm [40].

The Dawn of Photonics in Bulk CMOS

Not long after the 1 GHz modulation bandwidth barrier was broken in SOI, efforts were underway to integrate photonics into unmodified bulk CMOS processes. The first results were published in 2008, which demonstrated the fabrication of waveguides in a foundry bulk CMOS process using the polysilicon layer normally used for transistor gates [41]. However, waveguide loss was not recorded. In the same year, a method for etching the silicon substrate underneath the polysilicon waveguides using a XeF₂ gas etch was also demonstrated [42], a method which would prove crucial in the subsequent development of "zero-change" photonics using commercial processes [43].

In a parallel development, Intel demonstrated a back-end deposited photonic platform in 2010 [44] with an eye towards photonic interconnect for microprocessors. The platform used silicon nitride waveguides, polycrystalline germanium photodetectors, electro-optic polymer modulators, and dual-damascene copper interconnect. Just like the polysilicon waveguides investigated at MIT a decade and a half earlier, this back-end photonic platform had the advantage of waveguides that did not compete with transistors for area, with the potential for multilayer waveguiding as well. Moreover, silicon nitride waveguides are low loss (~1 dB/cm in this work [44]) and transparent across the visible and near IR.

In the same year, an entirely different method of integrating photonics into bulk CMOS debuted, that of deep trench isolation (DTI). In this method, a thick oxide layer is created underneath deposited waveguides to prevent coupling to the substrate. In the first demonstration of this method, published in 2010 [45], a layer of amorphous silicon was deposited over the deep trench, and then crystallized using solid phase epitaxy. The crystallized silicon was then selectively etched to form a waveguide with a loss of 6.1 dB/cm [45].

Still yet another method of creating waveguides in bulk CMOS was demonstrated in 2010, called Oxidized Silicon-On-Insulator (OxSOI). In this method, silicon pillars are first formed from the bulk substrate. A silicon nitride cap is then deposited to protect the upper portion of the pillar from oxidation. A long oxidation step then creates a thick silicon dioxide layer underneath the silicon channel. Waveguides fabricated using this method had losses of 2.92 dB/cm [46].

The following year, 2011, saw the first publication of waveguide loss for waveguides fabricated in an unmodified bulk CMOS process, with results of 55 dB/cm in a commercial 28 nm process [47]. This paper used a XeF₂ gas etch to locally etch the substrate under the polysilicon waveguide, without which the loss would have been prohibitive. These results were improved upon in 2012 through the addition of an annealing step in an emulated DRAM process, which resulted in polysilicon waveguides with 6.2 dB/cm loss [48].

In 2013, the method of DTI with polysilicon waveguides was demonstrated in a modified memory process, with waveguide loss of 18 dB/cm [49]. The next year, in 2014, the method of DTI with crystallized silicon waveguides was used in a 65 nm memory process, and the waveguide loss was reduced to 3 dB/cm (from the previous 6.1 dB/cm) through further optimization [50]. The same year, DTI was used in a modified 180 nm memory process to form deposited polysilicon waveguides with 10.5 dB/cm loss [51], [52]. A few years later in 2018, the combination of DTI and polysilicon waveguides was used in a 65 nm bulk CMOS process [53].

A very different method of photonic integration in bulk CMOS was also published in 2014, in which germanium-silicon quantum wells are grown on a bulk silicon substrate. This method enables modulators and detectors to be fabricated along with the waveguides, which exhibit a loss of 2 dB/cm for wavelengths above 1440 nm [54].

2020 saw the publication of an improvement in waveguide loss in unmodified bulk CMOS, dropping the previous record of 55 dB/cm down to 38 dB/cm [55]. The authors did this by implementing a subwavelength grating waveguide in order to minimize the amount of polysilicon used in the waveguide. Although an improvement, the loss of 38 dB/cm is still unacceptably high for large photonic systems, which generally require losses around 1 dB/cm. This target loss provided the motivation for the subject of this thesis, the method of subtractive photonics.

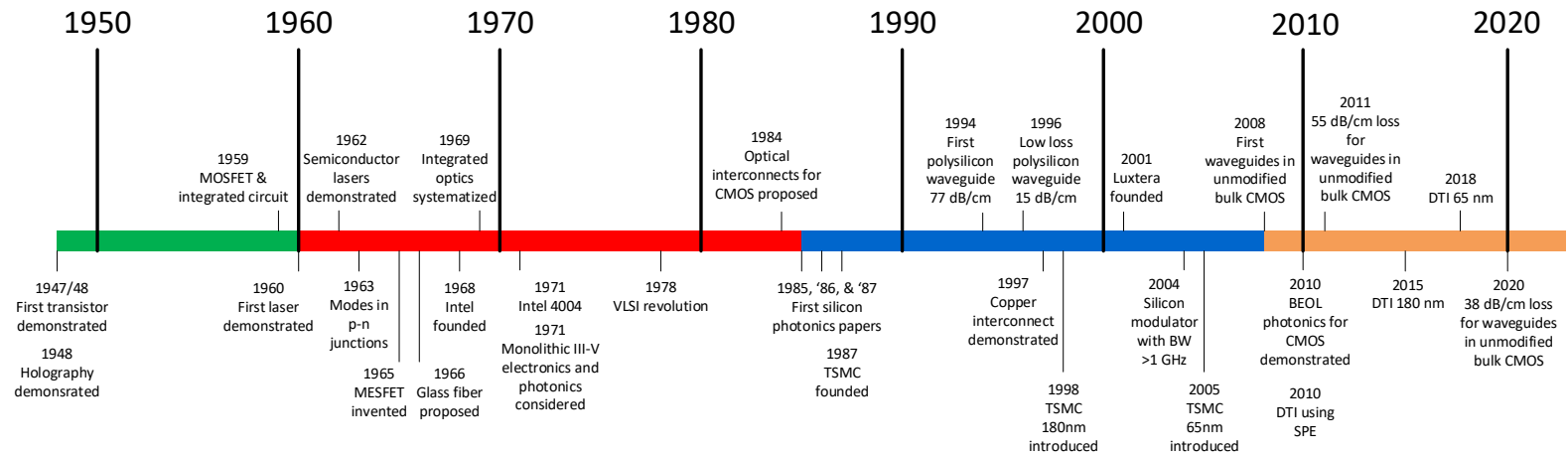


Figure 1.1: Brief timeline highlighting events germane to the development of photonics in bulk CMOS.

1.2 Methods of Photonic Integration

While the historical background provides important temporal context for the development of subtractive photonics, it is also important to consider the purely technological justifications for pursuing such a research agenda. The decision tree leading to subtractive photonics is depicted in Figures 1.2 and 1.3.

Beginning with a generic need for an electronic-photonic system, the designer is immediately confronted with the decision of whether to use integrated photonics or discrete photonics. Just as integrated electronics has a number of advantages over discrete electronics, integrated photonics provides a number of advantages over discrete photonics. Most notably, integrated photonics is well suited for the mass manufacturing of large and complex photonic systems.

Assuming that the designer needs to manufacture a complex electronic-photonic system at scale, the next choice is whether to use silicon or indium phosphide (InP) integrated photonics, which are the two dominant platforms for integrated photonics today [56]. As stated in the previous section, silicon photonics possesses the following advantages over InP:

- CMOS infrastructure can be leveraged to improve yields and reduce manufacturing costs.
- The high index contrast enables small photonic structures, and thus large photonic systems.
- The monolithic integration of photonics with silicon electronics is possible.

In contrast, InP offers the following advantages over silicon:

- Lasers and optical amplifiers can be monolithically integrated.
- Modulators can be created using the Pockels effect or electroabsorption [38].
- The monolithic integration of InP photonics with InP electronics is possible, albeit on a small scale [57].

It is important to note that the third advantage of silicon photonics is not an advantage *per se*, as electronics can also be monolithically integrated with InP photonics. However, the sheer scale of the electronic systems that can be made in silicon, compared with what has been demonstrated in InP, renders silicon a

far more advantageous material for electronic-photonics integrated circuit (EPIC) development [58].

A driving assumption in the line of reasoning leading to subtractive photonics is that the designer wishes to create large-scale systems. On this basis alone, silicon is far superior to InP, both in the world of electronics and the world of photonics, and doubly so in the world of monolithic electronic-photonics systems. Setting aside yields, the authors of one paper on electronic-photonics integration emphatically state that "The high index contrast of Si/SiO₂ dictates that this be the waveguide materials system for dense E-P integration." [39] However, just as there are severe limitations on the electronics that may be integrated in an InP photonics platform, we will see that there are severe limitations on the photonics that may be integrated in a silicon electronics platform.

Having chosen silicon due to the need to create large systems, the designer now faces the very important question of whether to use monolithic integration or some form of hybrid integration. Monolithic integration offers the advantages of:

- Reduced parasitics, which potentially enables higher performance.
- Reduced packaging costs, since there is only one chip.
- Increased design flexibility, since the interface between chips is not a limiting factor in system design.

In contrast, the main advantage of hybrid integration is that optimized electronics and photonics are **available now**. The experience of Luxtera, a company founded on the premise of monolithic integration, provides a compelling case study.

In [59], the authors note that "[Monolithic integration] definitely provided, in particular, the minimal amount of electrical parasitics at electrical/optical interfaces and maximized the benefit of very close integration of these two domains for overall system performance," confirming the advantages of monolithic integration listed above. But the authors continue: "However, this came with significant drawbacks. In particular, development of the technology platform was rendered more complex because the use of a custom substrate made the transistors different from existing SOI or bulk CMOS technologies, this requiring full recharacterization of their performance and the creation of custom device models." Since the photonics requirements were driving the choice of SOI wafer dimensions, new models needed to be developed for

the transistors, a process of essentially reinventing the wheel. "Process integration, in particular of the 100% Ge module, required careful tuning so as not to disrupt the performance of the transistors." Again, workarounds had to be found to prevent the photonics from interfering with the electronics. "Finally, and probably the worst drawback in the long term, the transistor technology was essentially frozen in time to what had been selected at the start of development... [product requirements] would have required redeveloping a new technology with improved transistors every time we wanted to improve electrical performance, which made no business sense." The tight coupling of photonics and electronics made development unnecessarily complicated, which, above and beyond the near-term difficulties, made it prohibitively laborious to advance further technologically in the long-term.

It is important to note here that the business difficulties Luxtera faced with their monolithic platform was mainly due to the fact that they chose an SOI platform, which entailed limitations due to the fact that the waveguides are fabricated on the front-end-of-the-line (FEOL). We will see below that the limitations Luxtera encountered may be overcome using a back-end-of-the-line (BEOL) monolithic platform, although Luxtera chose instead to resolve their difficulties using hybrid integration. Most likely this is due to the fact that, as mentioned above, optimized electronics and photonics processes were immediately available, and because of the challenges facing BEOL integration, which will be discussed below. The authors of [60] note that in Luxtera's new hybrid integration approach, "The small diameter of the micro-bumps and the intimate proximity of the photonic and electronic dies result in sufficiently small parasitics allowing minimal penalties in receiver sensitivity and transmitter power efficiency." Moreover, with recent hybrid integration techniques such as direct hybrid bonding [61] or through-oxide vias [62] that significantly reduce parasitics and provide very small chip-to-chip interconnect pitch, the advantage of monolithic integration is tenuous. It may be argued that packaging costs are lower with monolithic integration since only one chip is used, but this cost advantage might be offset by any process modification or post-processing required to actually implement the photonics. Consequently, limiting costs must be a primary consideration in the development of monolithic EPIC processes.

It may be supposed that the main reason Luxtera chose a SOI platform over a bulk CMOS platform is that the SOI platform has received far more attention than bulk CMOS, as evidenced by the timeline shown in the previous section. SOI silicon photonics has enjoyed decades of intense research by many research groups all over

the world. In contrast, the integration of photonics into bulk CMOS has only been studied by a handful of groups. This is unfortunate for the following reasons:

- Bulk CMOS dominates the electronics industry, providing tremendous incentive to develop new technology within the constraints of this platform.
- Older bulk CMOS nodes provide very low cost tapeouts, with fast turnaround and high shuttle availability, which is beneficial for academic research labs and startup companies.
- Newer bulk CMOS nodes provide the very highest performance digital electronics.

Similar to hybrid integration, the advantage of using a SOI monolithic platform is that they are **available now**. Currently, there are three commercial options:

1. GlobalFoundries 45 nm dedicated SiP process (45CLO), which offers silicon waveguides with 1.36 dB/cm loss at 1310 nm, silicon nitride waveguides with 0.35 dB/cm loss at 1310 nm, 35 GHz MZI bandwidth, 50 GHz Ge photodiode bandwidth, 280/230 GHz NFET/PFET f_T , v-grooves for passive fiber alignment, and other high-performance features [63].
2. The "zero-change" method applied to a commercial RFSOI process [43]. This option became essentially obsolete with the advent of the GlobalFoundries 45CLO process. Moreover, it required post-processing to etch the substrate locally underneath waveguides, and cannot be implemented in an RFSOI node below 32 nm, as the buried oxide becomes too thin [64].
3. IHP dedicated EPIC process using BiCMOS and SiP [65]. Unfortunately, this is a research foundry and so production cannot be scaled to commercial levels [66], limiting applications.

It may be observed from the above discussion that although the GlobalFoundries 45CLO process is a very appealing option, even appearing to embody all the aspirations that motivated the development of silicon photonics in the first place, bulk CMOS still possesses the advantages of market dominance, process variety, and above all, high-performance digital.

Assuming the designer wants to take advantage of all that bulk CMOS has to offer, the next choice is between BEOL and FEOL integration techniques. Without delving into the specifics of the techniques yet, the advantages of BEOL integration are:

- Flexibility in the choice of electronics process, since the fabrication of the photonics is decoupled from that of the electronics.
- Expensive FEOL area that is optimized for transistor fabrication is not used by the photonics. Photonic components, whose sizes are generally limited by the wavelength of light, consume significantly more area than electronic components, so BEOL integration saves substantial cost.
- Multilayer photonics is possible since the fabrication of the photonics is decoupled from that of the electronics.
- Visible wavelengths may be used with glass, SiN, or TiO₂ waveguides.

In contrast, the appeal of FEOL integration lies mainly in the fact that with silicon or polysilicon waveguides, high speed modulation is **immediately available** using the plasma dispersion effect, since the FEOL (poly)silicon may be doped along with transistors. In other words, no additional time needs to be spent developing an effective modulation mechanism.

Based on the foregoing discussion, we can see that the difficulties Luxtera encountered in their SOI FEOL process might have been mitigated using BEOL techniques, since the photonics is decoupled from the electronics, but the lack of an intrinsic mechanism for high-speed modulation in BEOL techniques still hampers their adoption (with the exception of BEOL polysilicon waveguides).

At this point, it is worthwhile to take a closer look at the options available using both the BEOL and FEOL techniques. BEOL techniques involve the creation of photonic components after all the interconnect for the CMOS chip has been completed, either through deposition or post-processing. FEOL techniques involve using waveguides at the same level as transistors, which may be implemented using the "zero-change" method, deep trench isolation (DTI), oxidized SOI (OxSOI), or SiGe quantum well waveguides.

The "zero-change" method entails forming waveguides from the polysilicon layer normally reserved for transistor gates. For waveguides, this polysilicon sits atop a thin layer of SiO₂ called shallow trench isolation (STI), which is normally used to

electrically isolate transistors [67]. Since the STI is normally only a few hundred nanometers thick, light propagating in the waveguides will couple into the silicon substrate underneath, causing unacceptably high losses [42], [47]. To sidestep this problem, the substrate must be locally etched underneath the waveguide using vertical channels that reach to the top of the chip. The etching is done using a XeF_2 gas etch in post-processing, and leaves the polysilicon waveguide supported by the suspended STI. This technique has the advantage that it is **available now** to use in existing bulk CMOS processes, but it unfortunately results in waveguides with high loss. As noted in the timeline of the previous section, this method at first yielded a waveguide loss of 55 dB/cm [47], which was later improved to 38 dB/cm using a subwavelength grating waveguide [55]. These loss values are prohibitively high in the context of large-scale and complex photonic systems, and considering that it will be difficult to remove more polysilicon material than a subwavelength grating already does, it does not appear that losses can be lowered any further without process modification. In [48], the authors implemented just such a process modification by adding an annealing step to the process after the polysilicon gate deposition, which reduced the roughness of the top surface of the polysilicon from 5 nm RMS to 0.3 nm RMS and yielded waveguide loss of 6.2 dB/cm.

The deep trench isolation method entails modifying a bulk CMOS process by adding a module for the deposition of relatively thick oxide trenches in the photonic portions of the chip. In one variant, after the creation of the deep oxide trenches, a layer of amorphous silicon is deposited that covers both the oxide trench and the surrounding crystalline silicon. Solid phase epitaxy is then used to crystallize the amorphous silicon, using the crystalline silicon surrounding the oxide trench as a seed. An etch step then defines waveguides on the trench by removing the surrounding crystalline silicon. Waveguide losses as low as 3 dB/cm have been achieved in this platform [50]. In another variant, polysilicon waveguides are deposited on top of the deep oxide trenches, initially yielding a waveguide loss of 10.5 dB/cm [51]. This method was next demonstrated in a modified 65 nm process, in which those wafers fabricated without electronics achieved a waveguide loss of 8 dB/cm at 1310 nm, while those wafers fabricated with electronics suffered waveguide loss of 21 dB/cm at 1310 nm [53]. The waveguide loss values attained in both variants of the DTI method are significantly better than those obtained using the "zero-change" method, but significant process modification is required.

The oxidized SOI (OxSOI) method entails creating waveguides by oxidizing the

silicon surrounding a channel. The method begins with a bulk silicon wafer, which is then etched to form silicon pillars or ridges. A silicon nitride cap is deposited on the ridges to create a region of silicon protected from oxidation. The silicon is etched further, heightening the ridges. A long oxidation step causes silicon dioxide to form underneath the capped silicon and form a buried oxide layer several micron thick. Waveguides fabricated using this method have a loss of 2.92 dB/cm [46], which is the lowest loss of proposed FEOL methods. However, the 16 hour oxidation period adds significantly to fabrication time, and as with DTI, significant process modification is required.

The SiGe quantum well method entails depositing layers of germanium-silicon alloy to form waveguides, modulators, and detectors. An 8 μm thick graded buffer is first deposited on the bulk silicon that transitions from Si_1Ge_0 to $\text{Si}_{0.17}\text{Ge}_{0.83}$. The waveguide itself, deposited on top of the buffer, uses a germanium-silicon alloy of $\text{Si}_{0.16}\text{Ge}_{0.84}$ 1.5 μm thick, which is partially etched 1 μm to form rib waveguides. To form modulators and detectors, a series of SiGe quantum wells are deposited on top of the waveguides. The waveguide loss decreases significantly for wavelengths longer than 1400 nm, with a measured low of 2 dB/cm at 1440 nm [54]. Although this method is attractive for its low waveguide loss and demonstration of integrated modulators and photodetectors, it requires a large number of additional process steps, beyond even the DTI and OxSOI methods.

BEOL deposition may be used with a number of different waveguide materials, most notably polysilicon [68], amorphous silicon [40], titanium dioxide [69], [70], silicon nitride [44], or a combination thereof [68], with the restriction that all deposition must be accomplished below 450°C [44], [68]. Polysilicon and amorphous silicon provide the highest index contrast, and the loss of polysilicon waveguides can get as low as 4 dB/cm at 1550 nm [39]. Amorphous single-mode silicon waveguides have been demonstrated with losses of 5 dB/cm at 1300 nm and 4 dB/cm at 1550 nm [40]. Silicon nitride waveguides have been demonstrated with a loss of 1 dB/cm at 1310 nm, together with electro-optic polymer modulators and polycrystalline germanium photodetectors [44]. Deposited titanium dioxide waveguides have demonstrated losses of 4 dB/cm at 1550 nm [69], 0.68 dB/cm at 1010 nm, and 7.8 dB/cm at 633 nm [70].

Surveying these waveguide options, it can be seen that waveguide loss is quite similar across all the materials in the infrared. Despite their advantage of high index contrast, silicon waveguides cannot be used in the visible, which may be a limiting

factor for some applications. In addition, the deposition of polysilicon requires an additional annealing or crystallization step [48], [68] to reduce losses, a cost which is potentially offset by the ability to dope polysilicon and form modulators, a luxury that amorphous silicon cannot provide due to the low carrier mobility and carrier lifetime [68]. Consequently, the use of amorphous silicon, silicon nitride, or titanium dioxide waveguides requires the deposition of an additional material to perform modulation, which may include liquid crystal, transparent conducting oxides, electro-optic polymer, etc. [38]. Waveguide-coupled photodetectors may be created using deposited polycrystalline germanium [44], or detectors on the surface of the silicon may be used either for visible wavelengths [71] or infrared wavelengths [72]. Assuming that high speed modulators and detectors can be built, it seems that FEOL integration methods have no intrinsic advantage at all over BEOL integration methods. The main motivation for current FEOL methods therefore seems to be expediency.

The deposition requirements of the BEOL waveguide material systems motivate the search for a solution that is **available now** to be implemented. Although post-processing will no doubt be a necessity for this solution, ideally it would entail only simple steps that could be accomplished at an outsourced assembly and test (OSAT) vendor. OSAT vendors can open up dielectric layers, deposit metal, and etch metal to form redistribution layers, which are abilities that can be used to shape dielectric structures. The method of subtractive photonics capitalizes on these abilities by designing waveguides using the back end dielectrics that are enclosed by metal. In post processing the metal is etched away, leaving the suspended waveguide [73]. Electrical pads may be enclosed in dielectric during the etching process and then opened after the etching is done. Both of these steps can be accomplished at an OSAT vendor, making the method of subtractive photonics amenable to mass manufacturing.

Subtractive photonics preserves all of the advantages of BEOL integration listed above, as well as the generic disadvantage of a lack of an intrinsic mechanism for high speed modulation, which might be compensated in same way using deposited materials. Subtractive photonics also entails two other disadvantages: 1) the low index contrast between glass and air implies that photonic structures will be relatively large, limiting the complexity of photonic systems built in this platform, and 2) since the waveguides must undergo etching as well as permanent suspension in air, there will be certain mechanical constraints on the size of the waveguides. It is interesting to observe that both of these disadvantages might be solved using high

index deposited materials, which together with the need for high speed modulation provides motivation for a switch from suspended dielectric waveguides to channel waveguides using deposited materials.

Having surveyed the field of possibilities, we are now in a position to address some common criticisms of monolithic electronic-photonics integration. In [74], Shekhar provides some exemplary arguments based on the assumption of a SOI monolithic silicon photonics platform. Shekhar notes that "1) maintaining a high performance of both the CMOS transistors and Ge PDs is complicated in an SOI process; 2) the fastest monolithic SiP process is currently based on a 45-nm CMOS SOI process, where the digital circuits are still not as energy efficient as in advanced fin field-effect transistor processes; and 3) the size difference between photonic and electronic devices raises cost considerations for a monolithic implementation. Even a small optical modulator fits in approximately $100 \mu m^2$. This size is tiny compared to that for discrete optics. But it is still 10,000 times the area of a transistor. Transistors also have 10 or more metal interconnects to route the electronics - a luxury photonic circuits do not have for optical waveguides." [74] Furthermore, Shekhar notes that "[hybrid integration] enables us to pick and choose the best processes for both [photonics and electronics], making this the preferred solution for many applications." [74]

Immediately it can be seen that choosing BEOL monolithic integration resolves most of these objections. With decoupling of photonics fabrication from electronics fabrication, the performance of CMOS transistors becomes independent of the performance of the photodiodes. With back end deposition, any electronics may be used, even the most energy efficient CMOS processes, so the electronics process may be optimized for the application. The size difference between electronics and photonics becomes irrelevant since the photonics is not consuming area that is optimized for electronics. Since waveguides are already being deposited instead of formed from wafer material, there is an immediate possibility of multilayer photonics, similar to the multilayer interconnect available to electronics.

Now it may be observed that the last objection, that of separate optimization of electronics and photonics, has not been fully answered. It may be true that the electronics may be fully optimized for the application, but what about the photonics? This objection returns to the main constraint on BEOL integration, which is the difficulty of obtaining high speed modulation and photodetection with post-processing temperatures remaining below $450^\circ C$. However, before an attempt can be made to solve this problem, the waveguides that form the foundation of these active devices

must be studied. Although deposited materials are a tantalizing prospect for BEOL integration, they are farther away from mass production than suspended dielectric waveguides, so it is prudent to explore the design space of the latter as much as possible. This exploration forms the bulk of the work in this thesis.

Table 1.1: Comparison of Bulk CMOS Integration Methods

Ref	Year	Method	Method Type	Process	Waveguide Material	Loss (dB/cm)	Wavelength (nm)
[47]	2011	Zero-change	Unmodified FEOL	28nm bulk CMOS	Polysilicon	55	1550
[55]	2020	Zero-change	Unmodified FEOL	90nm bulk CMOS	SWG polysilicon	38	1550
[51]	2014	DTI	Modified FEOL	180nm bulk CMOS	Polysilicon	10.5	1280
[53]	2018	DTI	Modified FEOL	65nm bulk CMOS	Polysilicon	21*	1300
[50]	2014	DTI	Modified FEOL	65nm bulk DRAM	Crystalline Si	20*	1550
[48]	2012	Zero-change	Modified FEOL	Emulated DRAM	Polysilicon	6.2	1550
[46]	2010	OxSOI	Modified FEOL	Photonics only	Crystalline Si	2.92	1550
[54]	2014	Deposition	Modified FEOL	Photonics only	SiGe	2	>1440
[44]	2010	Deposition	BEOL	Photonics only	Silicon nitride	1	1310
[40]	2006	Deposition	BEOL	Photonics only	Amorphous Si	5/4	1300/1550
[68]	2013	Deposition	BEOL	Photonics only	ELA polysilicon	65	1550

*Comparison is made only to waveguides fabricated with electronics.

SWG: Subwavelength grating

DTI: Deep trench isolation

OxSOI: Oxidized silicon-on-insulator

ELA: Excimer laser annealed

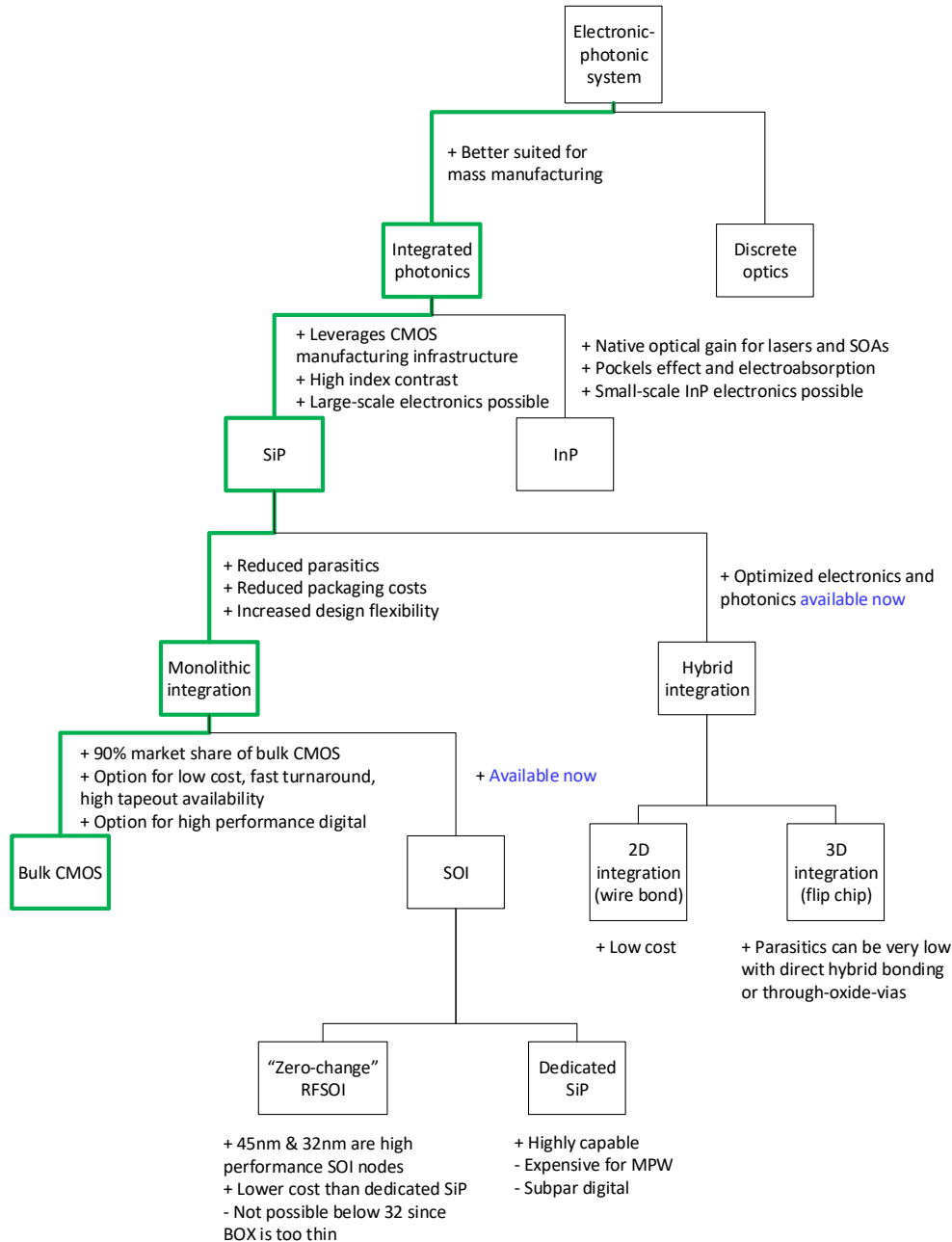


Figure 1.2: Tradeoff tree demonstrating the benefits of photonics in bulk CMOS.

1.3 Overview of BEOL Technology

Since the method of subtractive photonics uses the back end dielectric layers to form waveguides, it is essential to understand how these layers are fabricated, in order to understand their properties. When CMOS manufacturing began, simple SiO₂ interlayer dielectric (ILD) and aluminum metal lines provided sufficient performance.

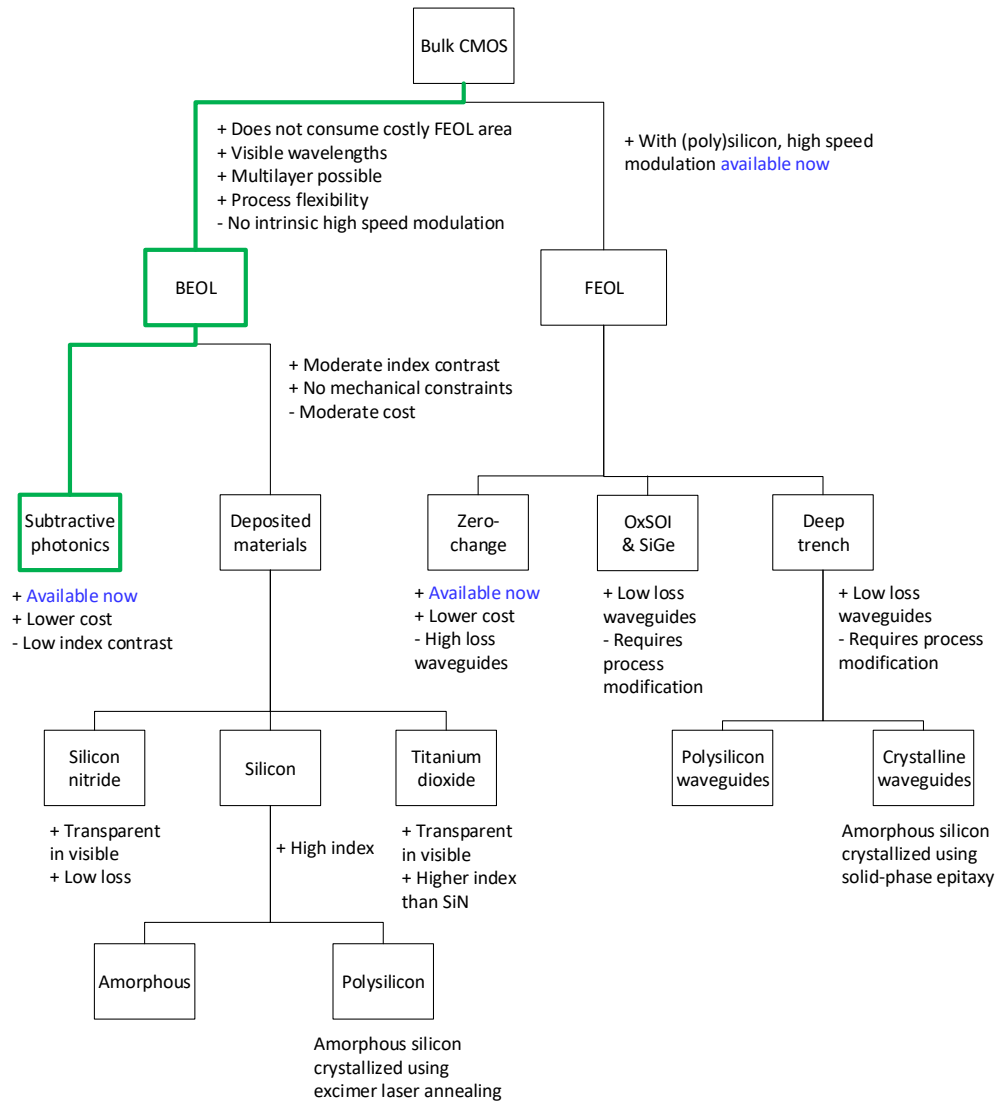


Figure 1.3: Tradeoff tree considering the benefits of subtractive photonics against other methods of integrating photonics into bulk CMOS.

But researchers recognized that around the 250 nm node [75], the RC delay due to the interconnect would begin to act as a bottleneck on system performance. The RC time constant of an interconnect line surrounded on all sides by neighboring lines can be written as [31]

$$RC = 2\rho k\varepsilon_0 \left(\frac{4L^2}{p^2} + \frac{4L^2}{t^2} \right). \quad (1.1)$$

Here ρ is the resistivity of the interconnect line, k is the dielectric constant of the ILD, L is the length of the line, p is the horizontal pitch of the interconnect lines, and t is the vertical pitch of the interconnect lines. The resistivity of the line is affected by the resistivity of the metal, barrier layer, and glue layer, as well as the relative thicknesses of each of these. Clearly the ILD, metal, barrier layer, and glue layer all must be chosen judiciously in order to achieve a specified RC delay.

In order to reduce the dielectric constant of the ILD, researchers began studying alternatives to SiO_2 . Materials with a dielectric constant lower than SiO_2 at the frequencies used in CMOS systems are called low- k dielectric materials. In the 180 nm process used in this work, the low- k ILD is mostly likely fluorosilicate glass (FSG/SiOF/fluorine doped glass). It is likely, and not certain, because the manufacturer does not explicitly specify the low- k material used. However, in an RF process the manufacturer does specify the dielectric constant associated with each layer for the purposes of electromagnetic (EM) simulations, and together with a study of the literature, this dielectric constant can be used to determine the material used. At the beginning of the search for low- k materials, researchers were looking for a "drop-in" replacement for SiO_2 that could be used in the near-term. Fluorosilicate glass can be processed in much the same way as SiO_2 , but has a dielectric constant of 3.6 [31] (against 4.2 for SiO_2) due to the Si-F bonds, which made it an excellent fit for the 180 nm node. For the other process used in this work, 65 nm, the ILD is most likely organosilicate glass (OSG/SiCOH/carbon doped glass). Again, this is based on the literature and the given dielectric constant, since SiCOH has a dielectric constant of about 2.8 [76].

Before 1997, manufacturers used aluminum interconnect in CMOS chips because of aluminum's abundance, ease of deposition, ease of etching, good adhesion to SiO_2 , and low resistivity [31]. In the 180 nm process used in this thesis, the interconnect is an aluminum-copper alloy with tungsten vias. The small amount of additional copper (typically between 0.5% and 4% [31]) is used to improve electromigration resistance, and tungsten is used because of its ability to form high aspect ratio vias.

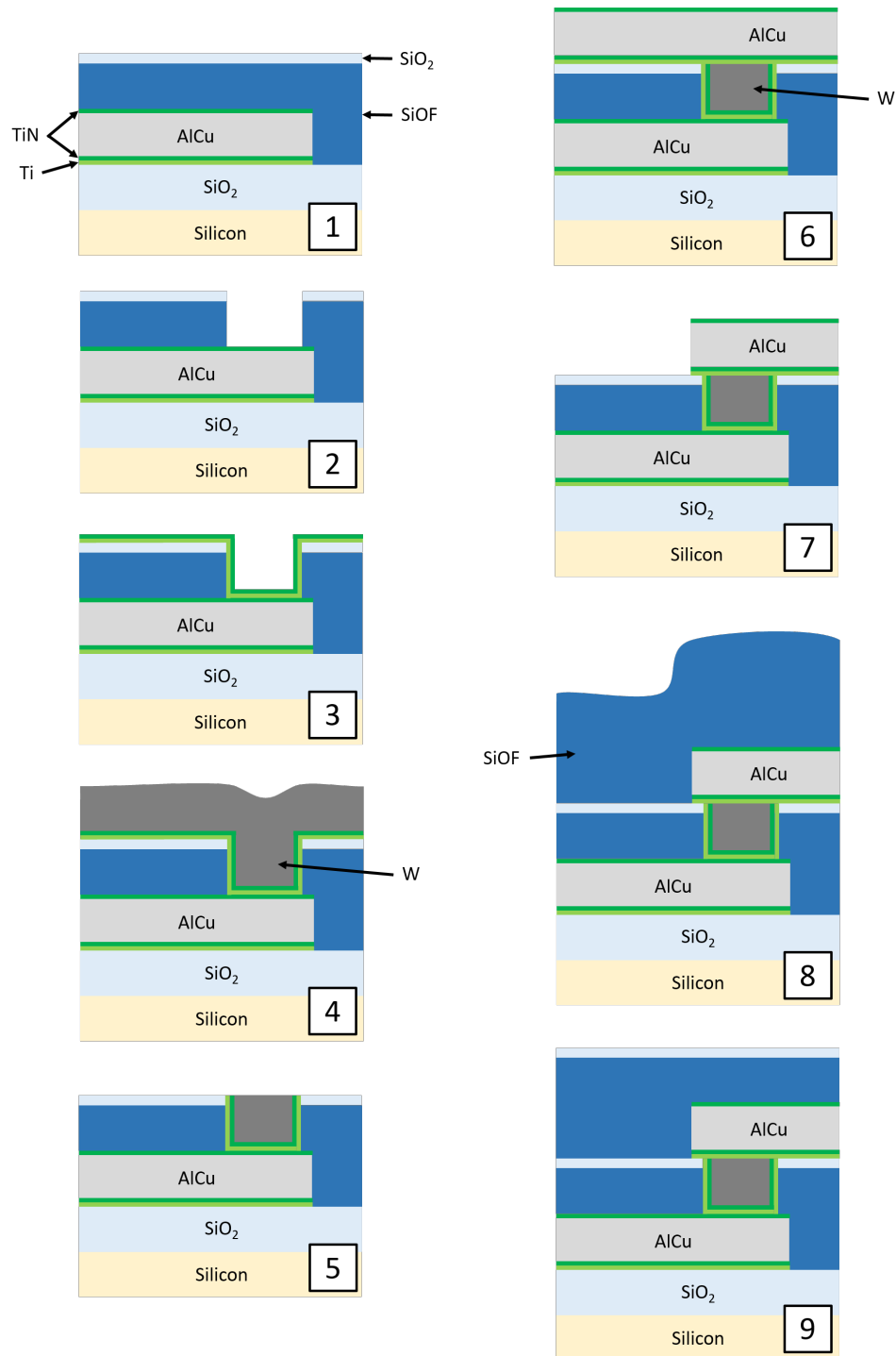


Figure 1.4: Outline of the aluminum interconnect fabrication process. The steps are detailed in the text.

An example of the typical process for depositing a single interconnect layer is shown in Fig. 1.4, following [67], [76] and assuming that a line has already been deposited.

1. Via creation starts with deposited dielectric, in this case FSG with a SiO_2 cap.
2. Resist is deposited and patterned (not shown), allowing the dielectric to be opened for the vias.
3. Titanium is deposited to prevent the formation of silicon nitride during titanium nitride deposition. Titanium nitride is deposited to provide adhesion between the tungsten and dielectric.
4. Tungsten is deposited using chemical vapor deposition.
5. Chemical-mechanical polishing (CMP) is used remove excess tungsten and barrier layer.
6. Titanium is deposited to act as a glue layer between the subsequent titanium nitride and the surrounding dielectric. The titanium also serves to provide a parallel alternative current path in case a void forms in the AlCu line due to electromigration. Titanium nitride is deposited to provide a diffusion barrier between titanium and aluminum. Next the AlCu is deposited, after which follows another layer of titanium nitride. The titanium nitride serves both as an anti-reflection coating for subsequent lithography, as well as an etch stop for the next via.
7. Resist is deposited and patterned (not shown), enabling a subtractive etch of the Ti/TiN/AlCu/TiN underneath.
8. FSG is deposited over the whole chip.
9. The FSG is planarized using CMP, and a SiO_2 cap is deposited.

In 1997, IBM introduced a copper dual damascene interconnect process as a replacement for aluminum interconnect. Copper has lower resistivity than aluminum and better electromigration resistance, but it suffers from worse adhesion to dielectrics and cannot be dry etched. Gold also has lower resistivity than aluminum, but compared to copper, it has higher resistivity, worse electromigration performance, and poor adhesion to dielectrics. Silver has the lowest resistivity of all metals, but it corrodes easily, has poor adhesion to dielectrics, and worse electromigration performance than both copper and gold [77].

Copper's low resistivity and superior electromigration properties rendered it the interconnect material of choice, and it is the interconnect metal of the 65 nm process

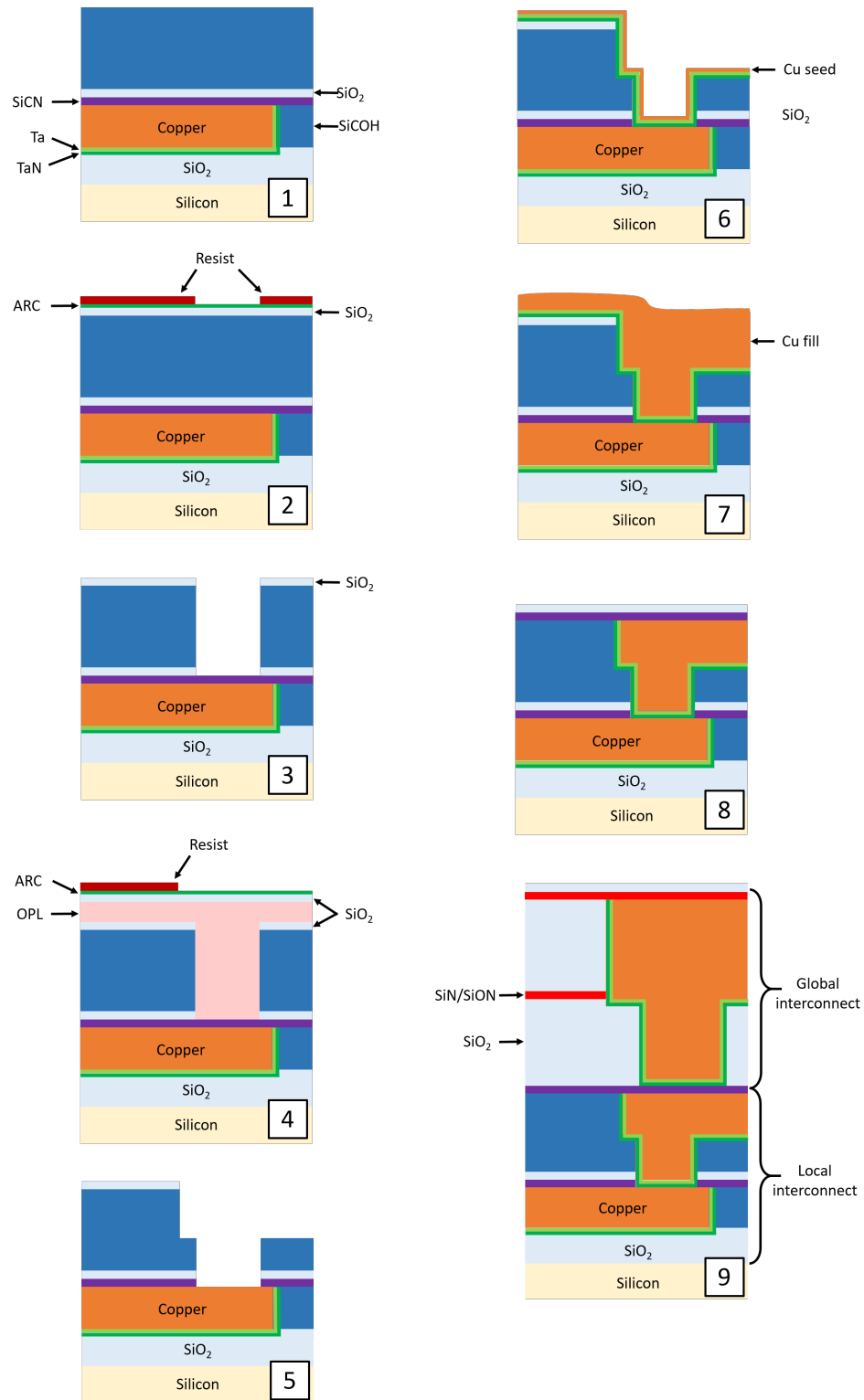


Figure 1.5: Outline of the copper interconnect fabrication process. The steps are detailed in the text.

used in this thesis. The dual damascene process is called "dual" because the vias and trenches are filled with metal in one step, and "damascene" because it is similar to the damascene process in metallurgy, in which metal is inlaid into a cavity.

The liner in a copper interconnect process cannot be titanium based, as it is in aluminum interconnect, since titanium diffuses into copper and increases its resistivity [75]. However, tantalum and tantalum nitride are both effective diffusion barriers for copper. Ta and TaN are usually combined together into a bilayer for two reasons. Both materials exhibit high resistivity independently, but Ta grown on TaN exhibits low resistivity. This is especially important because the Ta provides a parallel current path in case a void appears in a copper line due to migration. Furthermore, TaN exhibits good adhesion to SiO_2 but poor adhesion to Cu, while Ta exhibits good adhesion to Cu but poor adhesion to SiO_2 [75]. This implies that the optimal liner configuration is $\text{SiO}_2/\text{TaN}/\text{Ta}/\text{Cu}$.

The copper itself is usually deposited using electroplating. Compared to other deposition techniques, electroplating offers films with better electromigration resistance and better step coverage (or gap fill). It is also a more cost-effective technique [31]. However, electroplating requires a seed layer, which is usually deposited using conventional means on top of the Ta liner.

There are a variety of ways to implement copper dual damascene interconnect, and one example for a single interconnect layer is shown in Fig. 1.5, following [75] and assuming that a line has already been deposited.

1. Via creation starts with deposited dielectric, in this case OSG.
2. A SiO_2 hardmask, anti-reflection coating (ARC), and resist are deposited.
3. The dielectric stack is patterned for the vias.
4. An organic planarization layer (OPL) is deposited to planarize the chip surface for subsequent lithography. Again, a SiO_2 hardmask, anti-reflection coating (ARC), and resist are deposited.
5. The dielectric stack is patterned for the trenches.
6. TaN, Ta, and the Cu seed are deposited.
7. Copper is deposited using electroplating.
8. The copper is planarized using CMP.

9. A silicon carbon nitride cap is deposited to serve as a diffusion barrier between the copper and dielectric, an etch stop for subsequent processing steps, and as protection for the copper from corrosion, oxidation, and humidity. An SiO_2 layer is deposited on top of the SiCN, most likely to improve adhesion between the cap and subsequently deposited low- k dielectric.
10. Although lower layers of interconnect, called the local interconnect, use low- k dielectric and SiCN caps, the upper layers, called the intermediate and global interconnect, simply use an SiO_2 dielectric and SiN/SiON hardmasks.

The interconnect processes described above are necessarily complex, due to the myriad of mechanical and chemical problems process engineers have faced integrating copper, tungsten, and low- k dielectrics. The complexity of the interconnect stack, as well as the mechanical and chemical problems associated with it, will become clearer as waveguides are explored within this material system.

THE SUBTRACTIVE PHOTONICS CONCEPT

2.1 Introduction to the Concept

Many researchers studying devices on the micro- or nanoscale attempt to justify the economic feasibility of their devices by using "CMOS compatible" materials, but the definition of CMOS compatibility is often expanded far beyond what any CMOS foundry would reasonably consider compatible. For example, in silicon photonics "CMOS compatibility is considered a gold standard in SiP because its basis lies in leveraging of the CMOS fabrication infrastructure and process. However, notion [*sic*] of compatibility is often vague to the point where any material not explicitly named incompatible with CMOS such as gold is phrased to be CMOS compatible. For a [*sic*] straightforward adoption by the CMOS industry, we adopt definition [*sic*] of compatibility as consisting exclusively of materials already in use in commercial CMOS process [*sic*] ..." [68]

Even when the definition of CMOS compatibility is restricted to those materials already being used in commercial CMOS processes, such a definition still encompasses a huge variety of devices, some of which may require substantial process modification for proper integration into CMOS foundries. Despite the large number of papers claiming CMOS compatibility, CMOS foundries devoted to mass manufacturing have deviated little from their primary mission of fabricating electronics. Silicon photonics foundries are a prominent exception, but so far rely on SOI substrates, which form but a small fraction of the total market. In contrast, bulk CMOS processes "comprise 92% of CMOS logic production on 300 mm wafers." [47].

Considering the general failure of exotic "CMOS compatible" devices in the literature to influence commercial foundries, it is reasonable to instead start with an existing CMOS process and ask what can be done to create devices with the manufacturing processes that are already operated at very large scales. In other words, instead of expecting "CMOS compatibility" to act as a middle ground that CMOS foundries will naturally meet in the middle at (Fig. 2.1), devices can be formed in actual CMOS processes using post-processing. In section 1.2 it was seen that both zero-change FEOL integration and BEOL integration were attempts at integrating photonics using existing CMOS processes. However, the zero-change FEOL method resulted

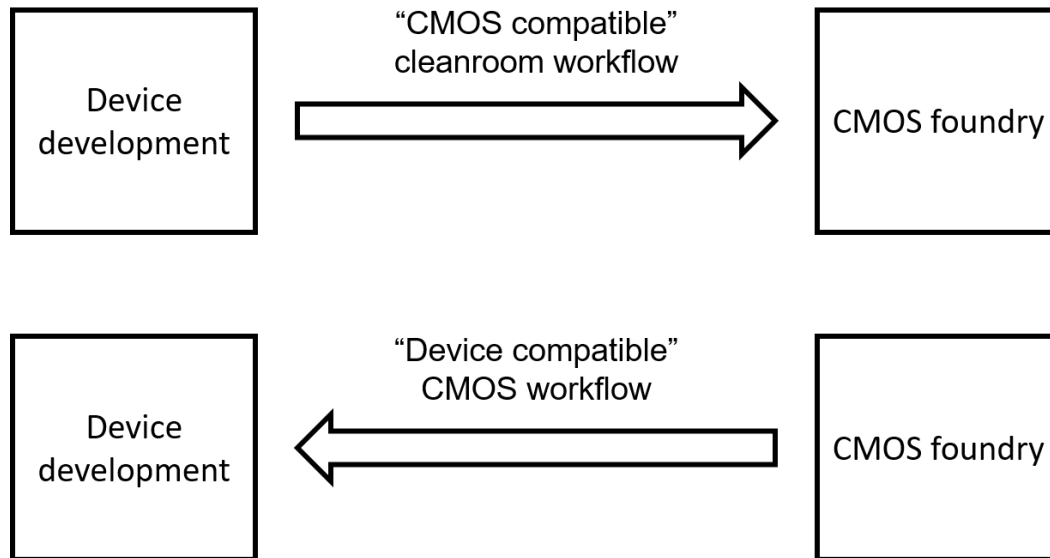


Figure 2.1: Opposing workflows for developing devices that can be mass manufactured in CMOS foundries.

in waveguides with prohibitive losses, and the deposition of waveguide materials requires either process modification, which is costly and unappealing to bulk CMOS foundries, or a dedicated service (similar to that which outsourced assembly and test (OSAT) vendors provide), which would also be costly.

The method of subtractive photonics is intended to capture almost all of the benefits of back end integration in a very low cost, accessible manner. The essence of the method is the use of the back end dielectrics to form suspended waveguides, as shown in Fig. 2.2. The interconnect metal is designed to wrap around the dielectric waveguide, so that after the chip is manufactured at a commercial bulk CMOS foundry, it can be submerged in etchant to remove the metal and reveal the waveguide. We typically use Aluminum Etch Type A at 80°C for etching the bulk of the aluminum or copper interconnect [78], [79], while a 1:1 solution of EDTA and hydrogen peroxide (H_2O_2) is used to strip any remaining barrier layers [80], [81], typically between 60°C and 80°C.

The advantages of this method over the back end deposition of materials are clear:

- Subtractive photonics can be used *now* to create low loss optical waveguides in unmodified bulk CMOS.
- The cost of the post-processing consists in the etching equipment and the etchant itself (which is low cost), creating a very low barrier to entry.

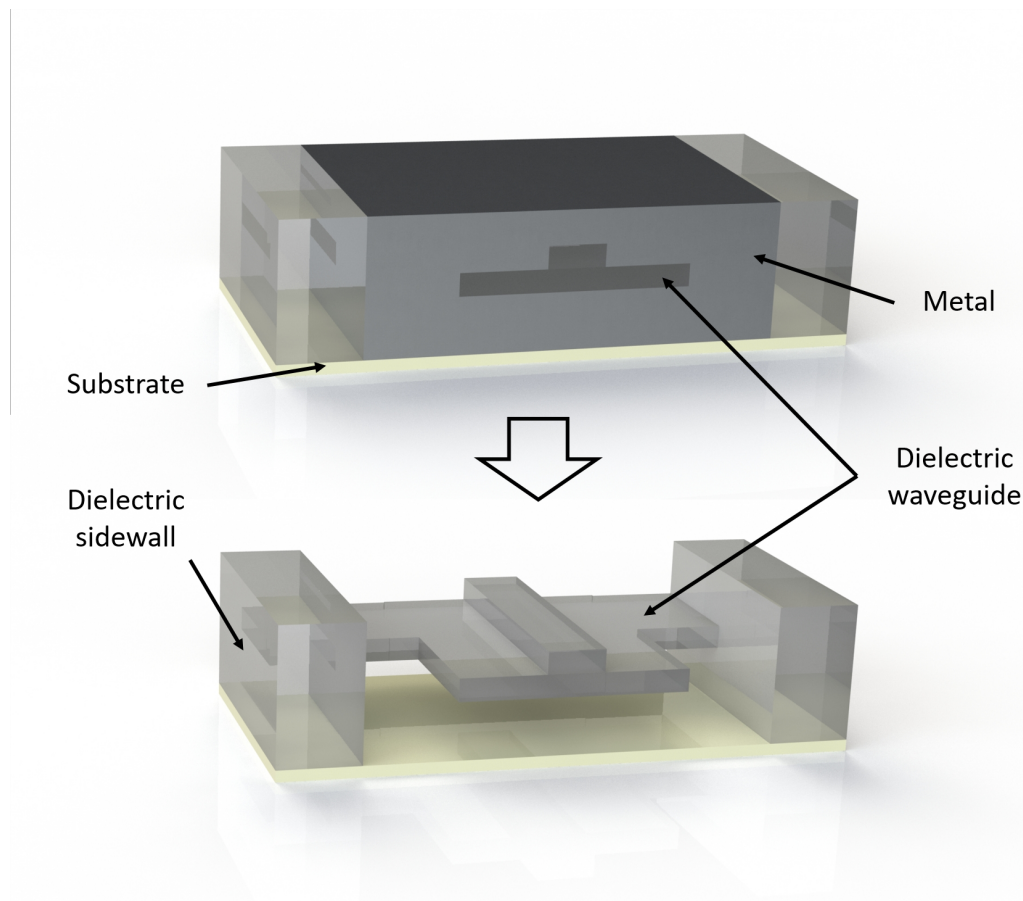


Figure 2.2: The method of subtractive photonics. Metal surrounding the waveguide is etched in post-processing, leaving suspended dielectric waveguides.

The advantages of this method are economic in nature. The method of subtractive photonics provides, in essence, a technical solution to the economic problems of back end deposition methods.

The primary disadvantages compared to back end deposition are:

- There is a relatively low index contrast between glass and air.
- There are mechanical constraints on what kind of structures can be fabricated due to the material properties of the back end interconnect, as well as the dynamics of the etching process.

The disadvantages of this method are technical in nature, which raises the possibility of addressing or circumventing these disadvantages using technical solutions. For example, the index contrast, although low, can be maximized using a channel

waveguide geometry (Fig. 2.3). The mechanical constraints are difficult to address *a priori* because the back end dielectric stackup is not exactly known. Without exact material properties for accurate simulations, experimental testing becomes necessary to determine the mechanical limits of the materials. Such testing has taken place over a number of tapeouts, as shown in Table 2.1.

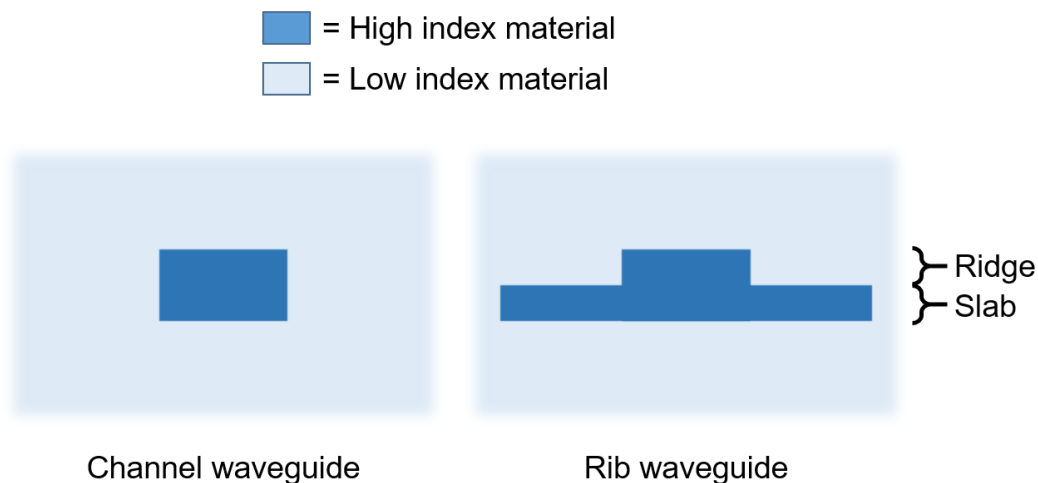


Figure 2.3: Terminology for the waveguides used in this thesis.

Name	Month	Node (nm)	RF process
TO0	May 2019	65	Yes
TO1	June 2021	180	Yes
TO2	September 2021	180	Yes
TO3	March 2022	180	Yes
TO4	June 2022	180	Yes
TO5	November 2022	180	No
TO6	March 2023	180	No

Table 2.1: List of tapeouts.

Since it is possible to design a large number of different waveguide geometries using the method of subtractive photonics, it is useful to define a nomenclature. Channel waveguides will be denoted by a list of the layers that comprise the waveguide, as in C1C2C3. Rib waveguides will be denoted by a list of the layers that comprise the ridge, followed by a list of layers that comprise the slab, as in R1R2R3-S1S2S3.

2.2 TO0

Note: The work described in this section was performed in collaboration with Reza Fatemi and Aroutin Khachaturian.

The first tapeout using the subtractive photonic method, termed Tapeout 0 or TO0, was intended to address both disadvantages of subtractive photonics by using channel waveguides, and by experimenting with the mechanical properties of the back end materials using a variety of structures. TO0 was designed in a 65 nm bulk CMOS process optimized for RF operation, including an ultra-thick top metal layer (Fig. 2.4). Note that the layer names are often abbreviated, as in M9 for the metal 9 layer, V8 for the via 8 layer, etc. This process uses a dual damascene copper interconnect, which means it most likely uses a TaN/Ta barrier and organosilicate low-k dielectric layers.

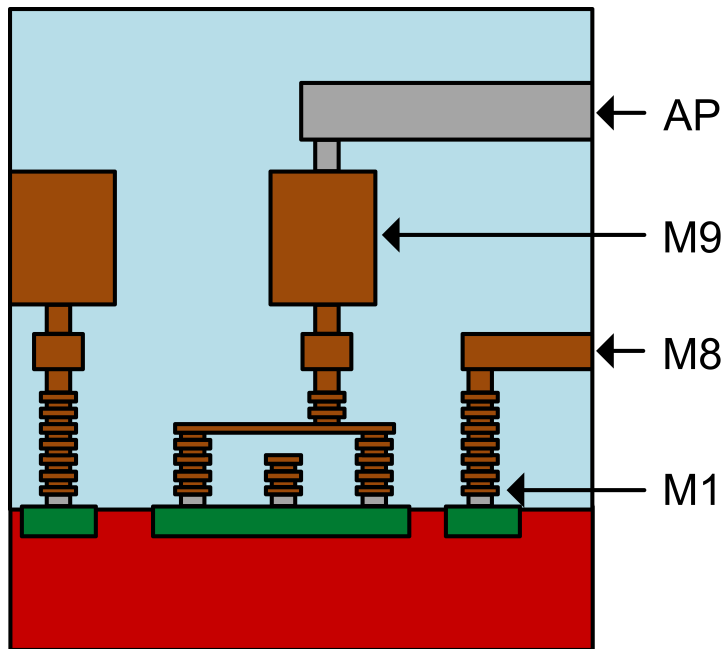


Figure 2.4: Depiction of the interconnect layers of the 65 nm process.

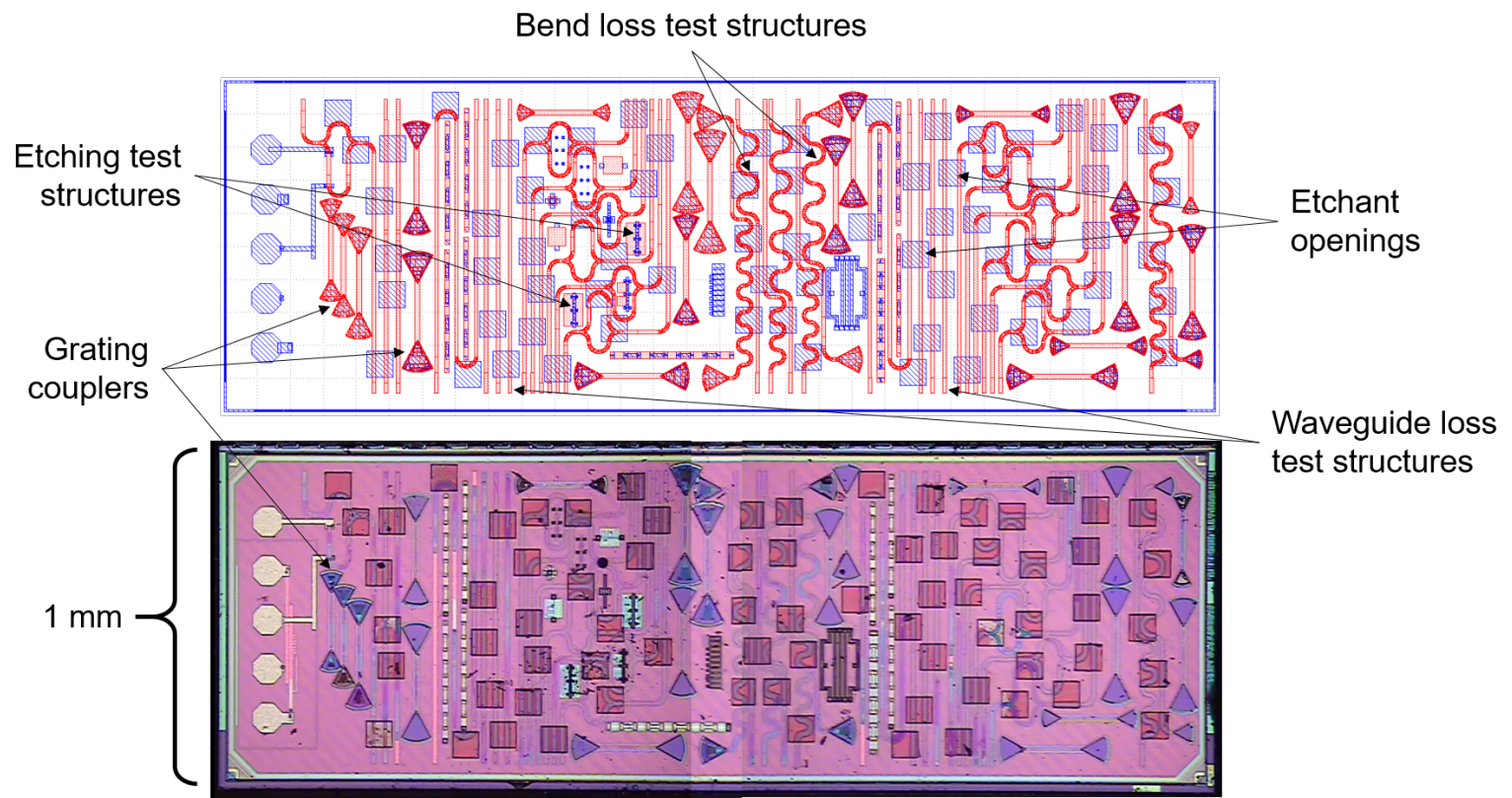


Figure 2.5: Layout of the 65nm test chip.

The layout of the chip is shown in Fig. 2.5. It features a variety of structures for testing fluid flow, mechanical integrity, and optical transmission. The primary question this tapeout was intended to answer was, could any of these structures be fabricated at all? The number of DRC violations is enormous for these types of photonic structures fabricated in electronic interconnect layers meant for Manhattan geometries, so it was unknown whether the foundry would permit such designs to be fabricated at all, or without modification. Other questions included: How big do the cavities need to be to ensure quick etching of the chip? Will there be residue after etching that would significantly affect optical loss? How much support do the waveguides need?

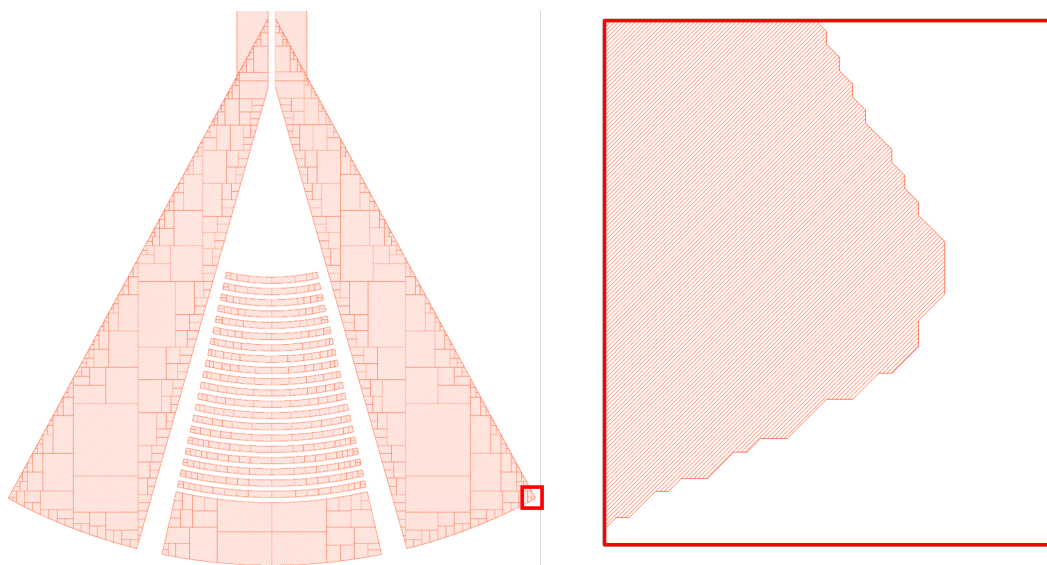


Figure 2.6: Example of fitting a photonic structure with curves to the minimum Manhattan grid.

In order to mitigate any problems with the lack of Manhattan geometry, all curves in the layout were fit to the minimum grid specified by the foundry, using only horizontal, vertical, or 45° edges (Fig. 2.6). Since these Manhattan edges were on the scale of nanometers, which the lithography at the 65 nm node cannot resolve, curved structures would be fabricated. Despite this effort, there were still over 32 million DRC errors.

Initial etching began with Aluminum Etch Type A (AETA) [78], which etches copper well [79]. After a day in room temperature AETA, the barrier layer was exposed and prevented further etching (Fig. 2.7). A FIB cut was performed to verify that the buried waveguides were indeed fabricated correctly (Fig. 2.8). The grating coupler waveguide that was opened by FIB was on the V7 layer, so this structure will be

referred to as V7-FIB.

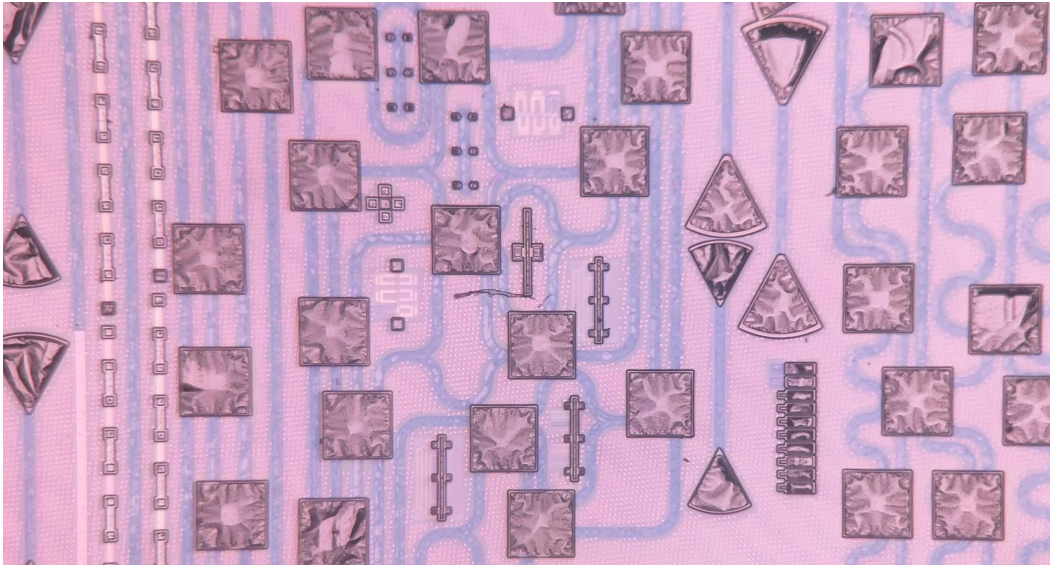


Figure 2.7: Initial etching tests using Aluminum Etch Type A.

Additional etching in AETA eventually exposed grating couplers (Fig. 2.10, top left). Since AETA does not etch TaN/Ta effectively [82], an alternate etchant is needed for these barrier layers. An initial test of chrome and titanium etchants caused little change. Next, the chip was etched overnight in a 1:1 solution of hydrogen peroxide and EDTA at room temperature [80], [81]¹. The results clearly show pieces of barrier layer cleanly separated from their original cavities (Fig. 2.9).

Next, an $\text{H}_2\text{O}_2/\text{EDTA}$ etch at 60°C over a duration of one hour resulted in a very clean etch of V7-FIB. Fig. 2.10 displays the progression of several grating coupler structures over the course of this testing, clearly showing that V7-FIB etched much cleaner than the two grating couplers surrounding it, since all the metal layers were exposed to etchant simultaneously through the FIB cut. Fig. 2.11 further shows that the entire cavity of V7-FIB, from grating coupler to grating coupler, etched very cleanly, validating the ability of the $\text{H}_2\text{O}_2/\text{EDTA}$ etchant to remove the barrier layers.

With replacement of the $\text{H}_2\text{O}_2/\text{EDTA}$ etchant and another hour long etch at 60°C , the surrounding grating couplers began approaching the cleanliness of V7-FIB (Fig. 2.10, bottom, and Fig. 2.12). In the middle images of Fig. 2.12, the waveguide of V7-FIB can be seen on the left. The waveguides on the right side of these images clearly need further etching. After the additional hour in $\text{H}_2\text{O}_2/\text{EDTA}$, the waveguides on the

¹This recipe came from the posts by Ian Yee and Jon Quirt in the cited forum threads.

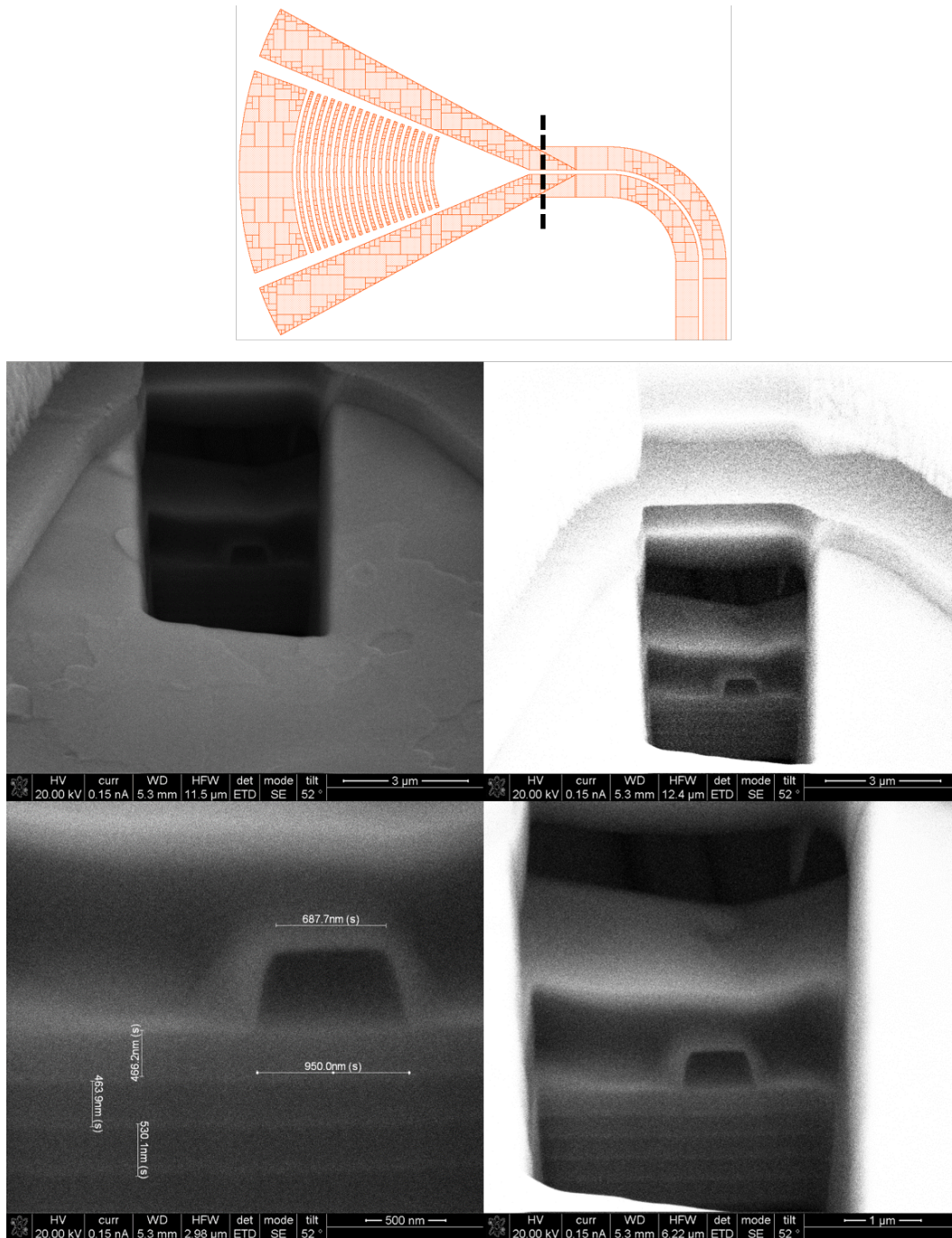


Figure 2.8: A waveguide fib cut. Top: Diagram depicting the location of the FIB cut relative to the V7 grating coupler. Middle: SEM images of a grating coupler cavity with a FIB opening for exposing the buried waveguide. These are images of the same location but with varying contrast. Bottom: Close-up images of the buried waveguide. SEM images courtesy A. Khachaturian.

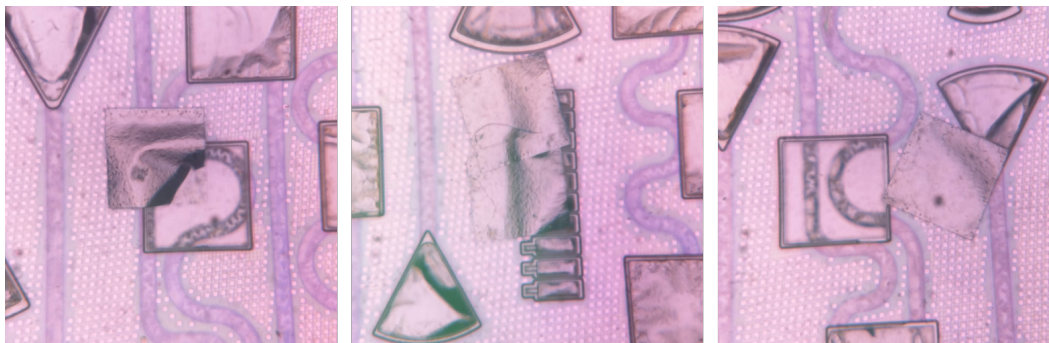


Figure 2.9: Images demonstrating the clean removal of the barrier layer.

right side of these images are approaching the cleanliness of V7-FIB in the area of the etchant openings. The reduced etching outside the etchant openings indicates the need for openings all along the length of the waveguide for uniform etching. More V7 gratings are shown in the bottom images of Fig. 2.12. Similar to the grating couplers around V7-FIB, the additional hour in $\text{H}_2\text{O}_2/\text{EDTA}$ has resulted in grating couplers that are close to fully etched.

So far the AETA was used at room temperature. Fig. 2.13 displays similar results from an expedited etching process on a different chip. The top images show that the aluminum has been etched down to the barrier layer, and in some grating coupler structures, the barrier layer has started peeling back to unveil the gratings. The lower left image displays the cleanliness of some grating couplers after an $\text{H}_2\text{O}_2/\text{EDTA}$ etch. The grating couplers do not all etch at the same rate, possibly because of the accumulation of debris in some cavities. The lower right image shows progress for the partially etched grating coupler after further etching in AETA, with the barrier layer appearing as a crumpled transparent layer. Despite some non-uniformity in etch rates, this etching recipe clearly can yield very clean photonic structures.

Figure 2.14 displays the results of etching another chip. This etching sequence consisted of an initial etch of 2 hours of AETA at 80°C , followed by an additional hour in AETA at 80°C , and finished with 2 hours of $\text{H}_2\text{O}_2/\text{EDTA}$ at 40°C . The right images of Fig. 2.14, as well as the top images of Fig. 2.13, show that the AETA can break open the barrier layer and even cause its removal, although this removal is not consistent across structures, and takes much longer than with a $\text{H}_2\text{O}_2/\text{EDTA}$ etch.

It is interesting to observe the barrier layer through SEM imaging as well. The chip of Fig. 2.14 was etched again in $\text{H}_2\text{O}_2/\text{EDTA}$ at 40°C for an hour. Fig. 2.15 shows a partially etched V7 grating coupler through both an optical microscope and an

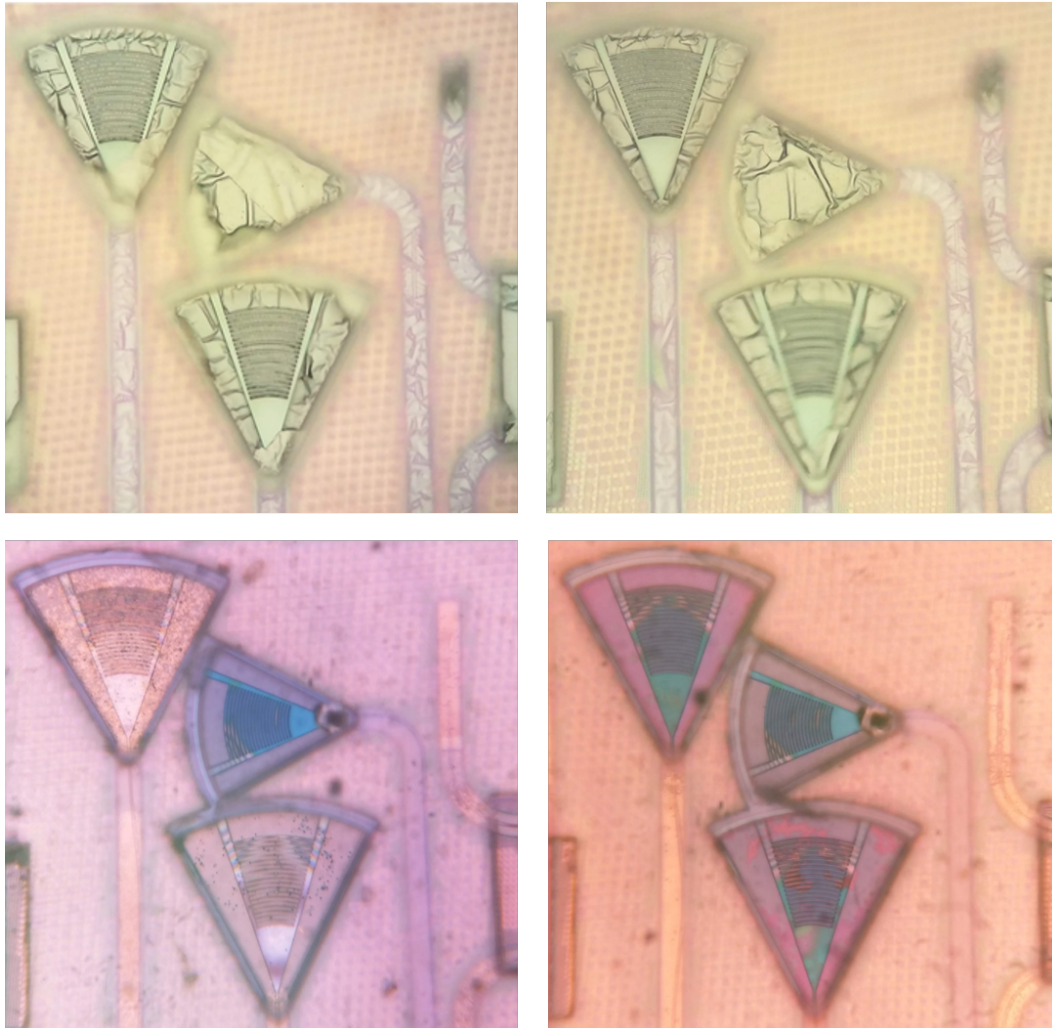


Figure 2.10: Initial etching tests. Top left: Grating couplers after initial etching in AETA. Top right: Grating couplers after tests with chromium and titanium etchants. Bottom left: Grating couplers after 1 hour in a 1:1 H₂O₂/EDTA solution at 60°C. Bottom right: Grating couplers after 1 hour in a refreshed 1:1 H₂O₂/EDTA solution at 60°C.

electron microscope, after this final etch step. The SEM image starkly shows how thin and wispy the barrier layers are, and is reinforced by Fig. 2.16, which displays two grating coupler cavities with partially removed barrier layers.

The V7 grating couplers of Figures 2.12 and 2.13 display discoloration near the supports. Upon closer examination, as shown in Fig. 2.17, it can be seen that the grating couplers have collapsed onto the bottom of the cavity. Clearly this indicates the need for many more supports than these gratings have. Gratings made from thinner layers, such as gratings that use the M6 layer for a slab, and the M7 layer for

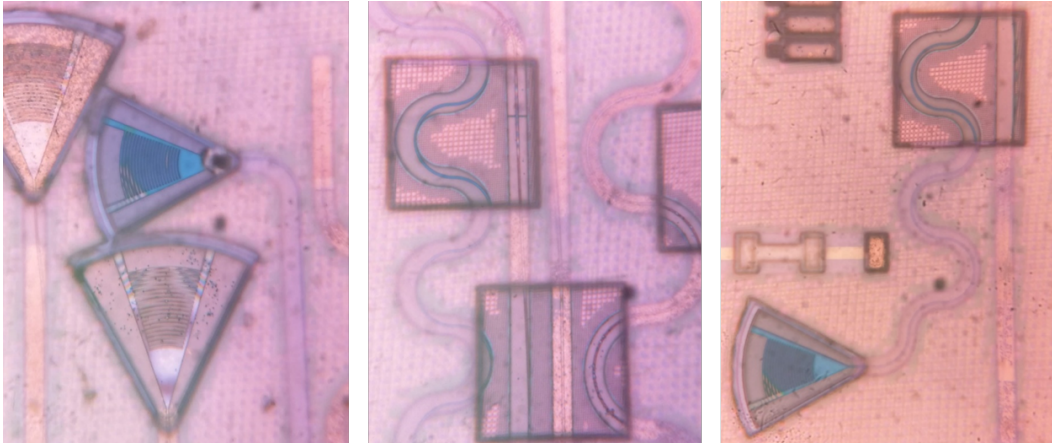


Figure 2.11: Images of the top, middle, and bottom portions of the V7-FIB structure.

a ridge, were entirely destroyed. Fig. 2.18 depicts both a partially etched and a fully etched cavity that contained M7-V6 gratings. The remnants of the supports can be seen in both images, indicating that these layers, on the order of 100 nm thick, are simply too thin to support photonic structures of this size.

The waveguides attached to the V7 grating couplers are also collapsed since the blue of the V7 grating couplers extends the length of their attached waveguides (Fig. 2.11). However, Figure 2.19 displays a supported waveguide that survived the etching process, indicating that simply by adding more supports a suspended photonic system is possible. The trade-offs associated with these supports will be studied below.

2.3 Risk Mitigation

The issues in TO0 indicate the need for not only a change in waveguide geometry, as discussed in the following chapter, but also the adoption of a general strategy for risk mitigation. In typical photonics or electronics research, the need for risk mitigation is either minimized or already addressed. Figure 2.20 displays the situation.

Research groups focused on devices typically fabricate those devices in research cleanrooms. The devices are individually characterized and then optimized through many fabrication trials. The cost of fabricating another device is low, so risk is low. Research groups focused on systems design their systems for production in established foundries. The foundries supply process development kits (PDKs) with fully characterized and parameterized cells, which are used in conjunction with industry standard tools to design, lay out, and verify the systems. The quality of the

device modeling ensures that simulations provide an accurate depiction of fabricated system performance, greatly reducing risk. For both types of groups, design for manufacturing is not necessary.

In contrast to both device research and system research, subtractive photonics falls into a uniquely risky research gap, the "chasm of manufacturability" depicted in Figure 2.20. The devices are manufactured using a standard CMOS process, which makes every fabrication run expensive compared to cleanroom fabrication. But the devices are not characterized or parameterized by the foundry, and so are susceptible to all of the risk that cleanroom devices are exposed to. In other words, subtractive photonics suffers from the worst of both worlds. The cost of fabrication is high, raising the cost of errors, and there are no experimentally verified device models to help avoid errors.

This compounded risk indicates the need for a risk management strategy. The risk can generally be divided into three categories, as shown in Figure 2.21: design risk, manufacturing risk, and measurement risk. Manufacturing risk, which is the possibility that a device will not function as expected due to unforeseen incompatibility with the manufacturing process, is out of the control of the designer. When manufacturing using CMOS foundries, the manufacturing process is set and cannot be changed. The only way the designer can mitigate any of this risk is to minimize DRC errors and attempt to meet all density requirements, the latter of which strongly affects wafer processing. Still, the manufacturing risk is very high due to the novelty of the devices and the limited information on the BEOL supplied by the foundry. This high manufacturing risk (the unknown unknowns) implies that the designer should minimize design risk and measurement risk (the known unknowns) in order to minimize overall risk. Once manufacturable devices have been fabricated and characterized, manufacturing risk becomes low, which allows the designer to expand design risk in the form of system design. Measurement risk should generally be minimized since there is usually no benefit in increasing it.

One way in which design risk is mitigated is through the use of scripting. Almost all of the structures in the chips subsequently discussed in this thesis were fully scripted in Python. The scripting provides several benefits. First, it mitigates design risk by minimizing repetition, which occurs frequently with manual polygon layout, and by easing the precise placement of polygon coordinates, which is time-consuming in manual layout. Second, it enables devices, systems, and entire layouts to be fully parameterized. This is very useful since the design space for devices in subtractive

photonics is quite large due to the large number of layers available. Third, scripting enables complex functions to be used in simulation and polygon manipulation that could be difficult to implement in a GUI.

Other types of design and measurement risk reduction will be highlighted in the later tapeouts covered by this thesis.

2.4 Waveguide Optimization

The problems that arose in TO0, described above, indicate the need for a change in waveguide design. The primary change to be made is to include many more supports for mechanical strength. However, for a suspended channel waveguide, the reflections from the supports are not negligible, and can cause significant loss over moderate distances. The immediate solution to this problem is to switch to a rib waveguide geometry (Fig. 2.3), which enables the supports to sit farther from the optical mode. The support distance becomes a parameter in the waveguide design which can be used to trade reflection loss with overall waveguide width. There is another advantage to using a rib geometry; due to the lower index contrast, it makes large area singlemode waveguides possible, which is a significant aid when designing waveguides for mechanical strength. Large area singlemode waveguides here refers to waveguides with cross-sections that are significantly larger than singlemode channel waveguides. The main disadvantage of a rib geometry is that the same reduced index contrast increases bend loss. Techniques to minimize bend loss will be discussed in section 2.4 below.

The secondary change is to focus on edge coupling. The TO0 chip included both grating couplers and edge couplers, but the grating couplers consume a relatively large amount of area and are usable over a relatively narrow bandwidth. In contrast, edge couplers offer a number of advantages:

- The ability to form large area waveguides by combining a number of interconnect layers enables high fiber coupling efficiency, since the waveguide facet size can be made comparable to the size of the core of a visible wavelength fiber, or the spot size of a lensed infrared wavelength fiber.
- Edge couplers are inherently wideband devices, and so can be used across the full useful optical spectrum, from visible to infrared. The only constraints on bandwidth would be requirements for singlemode behavior, as well as the transparency of the material itself.

- Edge couplers enable in-plane fiber coupling, which makes fiber packaging easier than the out-of-plane coupling required for grating couplers.

A coupling method with even more advantages than edge coupling is adiabatic coupling. In this method, the photonic chip is flipped on top of a substrate with a waveguide, and the waveguides of the chip and the substrate are aligned. This method can result in coupling losses of less than 1.5 dB for the TE polarization and less than 1 dB for the TM polarization. However, this method requires co-design with the substrate, so for initial characterization of the on-chip waveguides, edge coupling is sufficient.

The suspended rib waveguide introduces some nuance into the concepts of confined modes and radiation modes. In order to understand these properties, we must first study the propagation of light in a slab waveguide, in which the refractive index varies in only one dimension. Figure 2.22 shows such a slab waveguide, with \hat{y} pointing out of the page. The discussion below follows Yariv [83].

Symmetric Slab

Assuming a monochromatic electromagnetic field with a time dependence given by $e^{j\omega t}$, Maxwell's equations in a dielectric medium are

$$\nabla \times \vec{E} = -j\omega\mu\vec{H} \quad (2.1)$$

$$\nabla \times \vec{H} = j\omega\varepsilon\vec{E}. \quad (2.2)$$

The electric field of the transverse electric (TE) polarization propagating along \hat{z} can be written as

$$E_y(x, t) = E_m(x)e^{j(\omega t - \beta z)}. \quad (2.3)$$

For TE modes, Maxwell's equations are then reduced to

$$\frac{\partial^2}{\partial x^2} E_y + (k_o^2 n^2 - \beta^2) E_y = 0 \quad (2.4)$$

$$H_z = -\frac{1}{j\omega\mu} \frac{\partial}{\partial x} E_y. \quad (2.5)$$

There are several possible solutions to Equations 2.4 and 2.5. The first case is a growing or decaying exponential that crosses all three regions of the waveguide, with $\beta > k_o n_1$, but this is unrealizable and so not a physical solution. For a symmetric

slab, this leaves two solutions. One solution is sinusoidal in the waveguide core and exponentially decreasing in the cladding, a solution that is referred to as a confined mode. Another solution is sinusoidal in both the core and the cladding, a solution that is referred to as a radiation mode.

The solution for the confined mode is

$$E_m(x) = \begin{cases} A \sin(hx) + B \cos(hx), & -\frac{d}{2} \leq x \leq \frac{d}{2} \\ C e^{-qx}, & x > \frac{d}{2} \\ D e^{qx}, & x < -\frac{d}{2} \end{cases} \quad (2.6)$$

where

$$h = \sqrt{k_o^2 n_1^2 - \beta^2} \quad (2.7)$$

$$q = \sqrt{\beta^2 - k_o^2 n_2^2}. \quad (2.8)$$

To satisfy boundary conditions, the fields tangential to the boundary, E_y and H_z , must be continuous across the boundary. Since μ is taken to be constant across the entire waveguide, Equation 2.5 can be simplified to:

$$\frac{\partial}{\partial x} E_m(x) = \begin{cases} hA \cos(hx) - hB \sin(hx), & -\frac{d}{2} \leq x \leq \frac{d}{2} \\ -qC e^{-qx}, & x > \frac{d}{2} \\ qD e^{qx}, & x < -\frac{d}{2}. \end{cases} \quad (2.9)$$

Matching the boundaries results in

$$A \sin\left(\frac{1}{2}hd\right) + B \cos\left(\frac{1}{2}hd\right) = C e^{-\frac{1}{2}qd} \quad (2.10)$$

$$hA \cos\left(\frac{1}{2}hd\right) - hB \sin\left(\frac{1}{2}hd\right) = -qC e^{-\frac{1}{2}qd} \quad (2.11)$$

$$-A \sin\left(\frac{1}{2}hd\right) + B \cos\left(\frac{1}{2}hd\right) = D e^{-\frac{1}{2}qd} \quad (2.12)$$

$$hA \cos\left(\frac{1}{2}hd\right) + hB \sin\left(\frac{1}{2}hd\right) = qD e^{-\frac{1}{2}qd} \quad (2.13)$$

which can be consolidated into

$$2B \cos\left(\frac{1}{2}hd\right) = e^{-\frac{1}{2}qd} (C + D) \quad (2.14)$$

$$2A \sin\left(\frac{1}{2}hd\right) = e^{-\frac{1}{2}qd} (C - D) \quad (2.15)$$

$$2hA \cos\left(\frac{1}{2}hd\right) = q e^{-\frac{1}{2}qd} (D - C) \quad (2.16)$$

$$2hB \sin\left(\frac{1}{2}hd\right) = q e^{-\frac{1}{2}qd} (D + C). \quad (2.17)$$

This leaves two options:

$$C = D \rightarrow A = 0 \quad \text{symmetric and even modes} \quad (2.18)$$

$$C = -D \rightarrow B = 0 \quad \text{antisymmetric and odd modes.} \quad (2.19)$$

For even modes we find

$$\begin{aligned} 2hB \sin\left(\frac{1}{2}hd\right) &= 2Cq e^{-\frac{1}{2}qd} \\ 2B \cos\left(\frac{1}{2}hd\right) &= 2C e^{-\frac{1}{2}qd} \\ \hline h \tan\left(\frac{1}{2}hd\right) &= q. \end{aligned} \quad (2.20)$$

With $u = \frac{1}{2}hd$ and $v = \frac{1}{2}qd$, the solution becomes

$$u \tan(u) = v. \quad (2.21)$$

For odd modes we find

$$\begin{aligned} 2hA \cos\left(\frac{1}{2}hd\right) &= 2Dq e^{-\frac{1}{2}qd} \\ 2A \sin\left(\frac{1}{2}hd\right) &= -2D e^{-\frac{1}{2}qd} \\ \hline h \cot\left(\frac{1}{2}hd\right) &= -q \end{aligned} \quad (2.22)$$

which can be rewritten as

$$u \cot(u) = -v. \quad (2.23)$$

For modes of TM polarization, propagation along \hat{z} can be written as:

$$H_y(x, t) = H_m(x) e^{j(\omega t - \beta z)}. \quad (2.24)$$

After application of Maxwell's equations, we find

$$\frac{\partial^2}{\partial x^2} H_y + (k_o^2 n^2 - \beta^2) H_y = 0 \quad (2.25)$$

$$E_z = \frac{1}{j\omega\epsilon} \frac{\partial}{\partial x} H_y. \quad (2.26)$$

Similar to Equation 2.6, the solution is

$$H_m(x) = \begin{cases} A \sin(hx) + B \cos(hx), & -\frac{d}{2} \leq x \leq \frac{d}{2} \\ C e^{-qx}, & x > \frac{d}{2} \\ D e^{qx}, & x < -\frac{d}{2} \end{cases} \quad (2.27)$$

where

$$h = \sqrt{k_o^2 n_1^2 - \beta^2} \quad (2.28)$$

$$q = \sqrt{\beta^2 - k_o^2 n_2^2}. \quad (2.29)$$

H_y and E_z must be continuous across the boundary. In Equation 2.26, $\varepsilon = \varepsilon_o \varepsilon_r = \varepsilon_o n^2$, so the change in refractive index across the waveguide boundaries must be taken into account. The equations for E_z can be simplified to:

$$\frac{1}{n^2} \frac{\partial}{\partial x} H_m(x) = \begin{cases} [hA \cos(hx) - hB \sin(hx)] (1/n_2^2), & -\frac{d}{2} \leq x \leq \frac{d}{2} \\ -qC e^{-qx} (1/n_1^2), & x > \frac{d}{2} \\ qD e^{qx} (1/n_1^2), & x < -\frac{d}{2}. \end{cases} \quad (2.30)$$

Matching the boundaries results in

$$A \sin\left(\frac{1}{2}hd\right) + B \cos\left(\frac{1}{2}hd\right) = C e^{-\frac{1}{2}qd} \quad (2.31)$$

$$[hA \cos\left(\frac{1}{2}hd\right) - hB \sin\left(\frac{1}{2}hd\right)] (1/n_2^2) = -qC e^{-\frac{1}{2}qd} (1/n_1^2) \quad (2.32)$$

$$-A \sin\left(\frac{1}{2}hd\right) + B \cos\left(\frac{1}{2}hd\right) = D e^{-\frac{1}{2}qd} \quad (2.33)$$

$$[hA \cos\left(\frac{1}{2}hd\right) + hB \sin\left(\frac{1}{2}hd\right)] (1/n_2^2) = qD e^{-\frac{1}{2}qd} (1/n_1^2) \quad (2.34)$$

which can be consolidated into

$$2B \cos\left(\frac{1}{2}hd\right) = e^{-\frac{1}{2}qd} (C + D) \quad (2.35)$$

$$2A \sin\left(\frac{1}{2}hd\right) = e^{-\frac{1}{2}qd} (C - D) \quad (2.36)$$

$$[2hA \cos\left(\frac{1}{2}hd\right)] (1/n_2^2) = q e^{-\frac{1}{2}qd} (1/n_1^2) (D - C) \quad (2.37)$$

$$[2hB \sin\left(\frac{1}{2}hd\right)] (1/n_2^2) = q e^{-\frac{1}{2}qd} (1/n_1^2) (D + C). \quad (2.38)$$

This leaves two options:

$$C = D \rightarrow A = 0 \quad \text{symmetric and even modes} \quad (2.39)$$

$$C = -D \rightarrow B = 0 \quad \text{antisymmetric and odd modes} \quad (2.40)$$

For even modes we find

$$\begin{aligned} [2hB \sin\left(\frac{1}{2}hd\right)] (1/n_2^2) &= 2C q e^{-\frac{1}{2}qd} (1/n_1^2) \\ 2B \cos\left(\frac{1}{2}hd\right) &= 2C e^{-\frac{1}{2}qd} \\ \hline h \tan\left(\frac{1}{2}hd\right) &= (n_2^2/n_1^2) q \end{aligned} \quad (2.41)$$

which can be rewritten as

$$u \tan(u) = \bar{v} \quad (2.42)$$

where $\bar{v} = \frac{1}{2}(n_2^2/n_1^2)qd$.

For odd modes we find

$$\begin{aligned} [2hA \cos(\frac{1}{2}hd)](1/n_2^2) &= 2Dqe^{-\frac{1}{2}qd}(1/n_1^2) \\ 2A \sin(\frac{1}{2}hd) &= -2De^{-\frac{1}{2}qd} \\ \hline h \cot(\frac{1}{2}hd) &= -(n_2^2/n_1^2) q \end{aligned} \quad (2.43)$$

which can be rewritten as

$$u \cot(u) = -\bar{v}. \quad (2.44)$$

Eigenvalue equations 2.21, 2.23, 2.42, and 2.44 may be solved graphically or numerically for the propagation constant β of the confined modes. Substituting β into Equations 2.6, 2.8, 2.27, and 2.29, we can plot the resulting mode profiles. The fundamental TE mode of a 1D slab waveguide is shown in Figure 2.23.

In Equations 2.6 and 2.27, we adopted solutions for which the field decays exponentially into the cladding. For this to be the case, we must have that $\beta^2 > k_o^2 n_2^2$, or $n_{eff} > n_2$, which results in a real q in Equations 2.8 and 2.29. These modes are commonly referred as confined modes, since the power is confined almost entirely within the core of the waveguide. If we instead take $\beta < k_o^2 n_2^2$, or $n_{eff} < n_2$, then q becomes imaginary and the solutions in the cladding become oscillatory to infinity, which are the radiation modes that were discussed earlier.

The situation becomes difficult when we try to move from a waveguide with index variation in one dimension to a waveguide with index variation in two dimensions, as there are no analytic solutions for the modes of a rib or channel waveguide [83]. However, the one dimensional solution makes it intuitive that the effective index of a confined mode must be greater than the index of the cladding. For a rib waveguide, the effective index of a confined mode must be greater than the effective index of the surrounding slab, where the effective index of the slab is the effective index of the slab's fundamental mode. This criterion separates the confined modes of a rib waveguide from the radiation modes, and leads to the definition of what may be termed the normalized index contrast,

$$\frac{n_{eff} - n_{slab}}{n_{glass} - n_{slab}}. \quad (2.45)$$

The normalized index contrast is used to characterize the relative confinement of modes in later chapters.

It is here that the slab adds nuance, since it acts as part of the cladding of a rib waveguide. If we take the slab to be infinite, then the radiation modes have either $n_{slab} > n_{eff} > n_{air}$ or $n_{air} > n_{eff}$. In the first case, the modes are confined in one dimension by the slab, but extend to infinity in the other two dimensions, so in the context of subtractive photonics they are referred to as slab modes. In the second case, the modes are not confined in any sense, and so in the context of subtractive photonics they are referred to as radiation modes.

For a realistic rib waveguide, the slab is divided into two categories, as shown in Figure 2.24. In the portions of a suspended rib waveguide where the slab functions as a mechanical support, the slab is connected to the cavity sidewall and is effectively infinite. In those portions where the slab is not connected to the cavity sidewall, the slab is finite and supports confined modes. Figure 2.24 shows a rib waveguide that is single mode with an infinite slab, but is multimode with a finite slab.

The effect of the confined slab modes is illustrated in Figure 2.25, in which a rib waveguide is imaged by scanning a lensed fiber across the waveguide facet and recording the current from a connected CMOS photodiode. It can be clearly seen that light can propagate through the entire slab of the waveguide, even though this is a single mode (i.e., single confined mode) geometry when the slab is connected to the sidewall. As a result, the slab acts as a wide channel waveguide, and so suffers from loss due to reflections from supports.

Supports

Reflections from supports can be modeled using mode overlap calculations in a simulation using the eigenmode expansion (EME) method. By sweeping the extension of the slab (Fig. 2.26) we can model the effect of slab width on reflection from supports. Fig. 2.27 shows an example of a slab extension sweep using a V5M5-V4M4 waveguide. At zero extension, which corresponds to a channel waveguide, the fundamental mode experiences essentially zero transmission over a distance of a centimeter due to reflections from supports. In contrast, at slab extensions over about $3 \mu\text{m}$, the fundamental mode experiences essentially full transmission over a centimeter. This justifies the switch from a channel geometry to a rib geometry.

As mentioned above, the finite slab of a suspended rib waveguide supports confined modes, and functions as a wide channel waveguide. Just as Figure 2.27 shows the

loss of a V5M5-V4M4 channel waveguide due to reflections from supports, it also shows how much loss confined slab modes will experience due to supports. The lack of confinement the slab modes experience when traveling through bent support sections also contributes to the loss. As a result, it is reasonable for designers to consider the slab as infinite for the purposes of analysis.

Bends

In addition to slab extension, a parameter important for rib waveguide design is the bend loss. Bend loss is most intuitively understood using an index transformation, which is illustrated in Figure 2.28. Consider the bent waveguide shown in the left part of Figure 2.28. It has a refractive index profile given in the plot at the bottom, and the effective index sits partway between the index of the core and the index of the cladding. The mode has a phase front at ϕ_1 and propagates around the bend to ϕ_2 . Since the wavefront must arrive entirely in phase at ϕ_2 , the wavefront must experience a radially increasing phase velocity. The phase velocity c_p can be written as [83]

$$c_p = \frac{c_o}{n_{eff}} \left(1 + \frac{r'}{R} \right) \quad (2.46)$$

where R is the center of the waveguide core, r' is the distance beyond R , and n_{eff} is the effective index of the mode. At some radius the phase velocity of the wavefront will exceed the speed of light in the medium c_o/n and the light will break off from the wavefront, resulting in loss. This loss can be modeled using an equivalent straight waveguide formed using an index transformation.

To form the equivalent straight waveguide, suppose that the bent waveguide is stretched into a straight waveguide, with ϕ_1 and ϕ_2 moved accordingly, as in right sight of Figure 2.28, but the wavefront retains the laterally increasing phase velocity, as denoted by the arrow thickness. The only way a wavefront starting at ϕ_1 can arrive at ϕ_2 entirely in phase is if a laterally increasing refractive index profile is added to the waveguide to compensate for the laterally increasing phase velocity. The refractive index profile can be written as [83]

$$n_{eq}(r') = n(r') \left(1 + \frac{r'}{R} \right), \quad (2.47)$$

where $n(r')$ is the physical refractive index profile, and is shown in the right side of Figure 2.28.

From the equivalent index $n_{eq}(r')$ in Fig. 2.28 we can immediately draw two conclusions. First, the modes of this waveguide will be different from the modes of a

straight waveguide, so there will be losses due to mode mismatch. Second, since the equivalent index is continually increasing towards the outside edge of the bend, at some distance from the core of the waveguide the equivalent index will be greater than the effective index of the confined mode. At this point the mode can couple into the higher index region, similar to coupling into a high index substrate. This models the loss due to a portion of the wavefront coupling into radiation modes due to excessive phase velocity.

The mode mismatch due to a change in curvature is similar to the mode mismatch between two waveguides of different widths. A common solution to the latter problem is an adiabatic taper. Similarly, the mode mismatch between two waveguides of different curvature can be solved using an adiabatic taper of curvature. Bends with a linear increase in curvature are called Euler bends [84], although generalizations are possible [85].

Although the laterally increasing equivalent index cannot be overcome since it is intrinsic to the bend, it can be mitigated by reducing the refractive index of the waveguide cladding on the outside edge of the bend [84]. For the rib waveguides considered in the rest of this thesis, this implies eliminating the slab on the outside edge of the bend.

A comparison of these techniques to mitigate bend loss are shown in Figure 2.29 for a V5M5-V4M4 waveguide. The bend radius that is swept in Figure 2.29 refers to the radius of the standard waveguide, which has constant curvature throughout the bend. An Euler bend that connects the same two points must necessarily have a maximum curvature greater than the curvature of the standard bend, since the Euler starts with a curvature of zero. This distinction is shown on the right side of Figure 2.29. Since the Euler bend has higher curvature, it suffers from greater loss for the same nominal bend radius.

The single-sided slab geometry is compared to the geometry of the standard bend in the upper right of Figure 2.29. For the waveguide geometries considered in this thesis, the single-sided slab geometry is multimode due to the increased index contrast on one side of the ridge. The oscillations in transmission of the single-sided slab in Figure 2.29 are due to coupling between the fundamental mode and higher order modes in the bend. The transmission is high when the bend path length is an integer multiple of the beat length between the fundamental and higher order mode, and can be used to design a so-called matched bend [84].

The Euler bend with a single-sided slab clearly provides the best solution. The single-sided slab ensures transmission is higher than the standard bend, and the smooth change in curvature ensures there is minimal coupling to higher order modes. As a result, low loss bends with radii less than $100\ \mu\text{m}$ can be used to build complex systems on-chip.

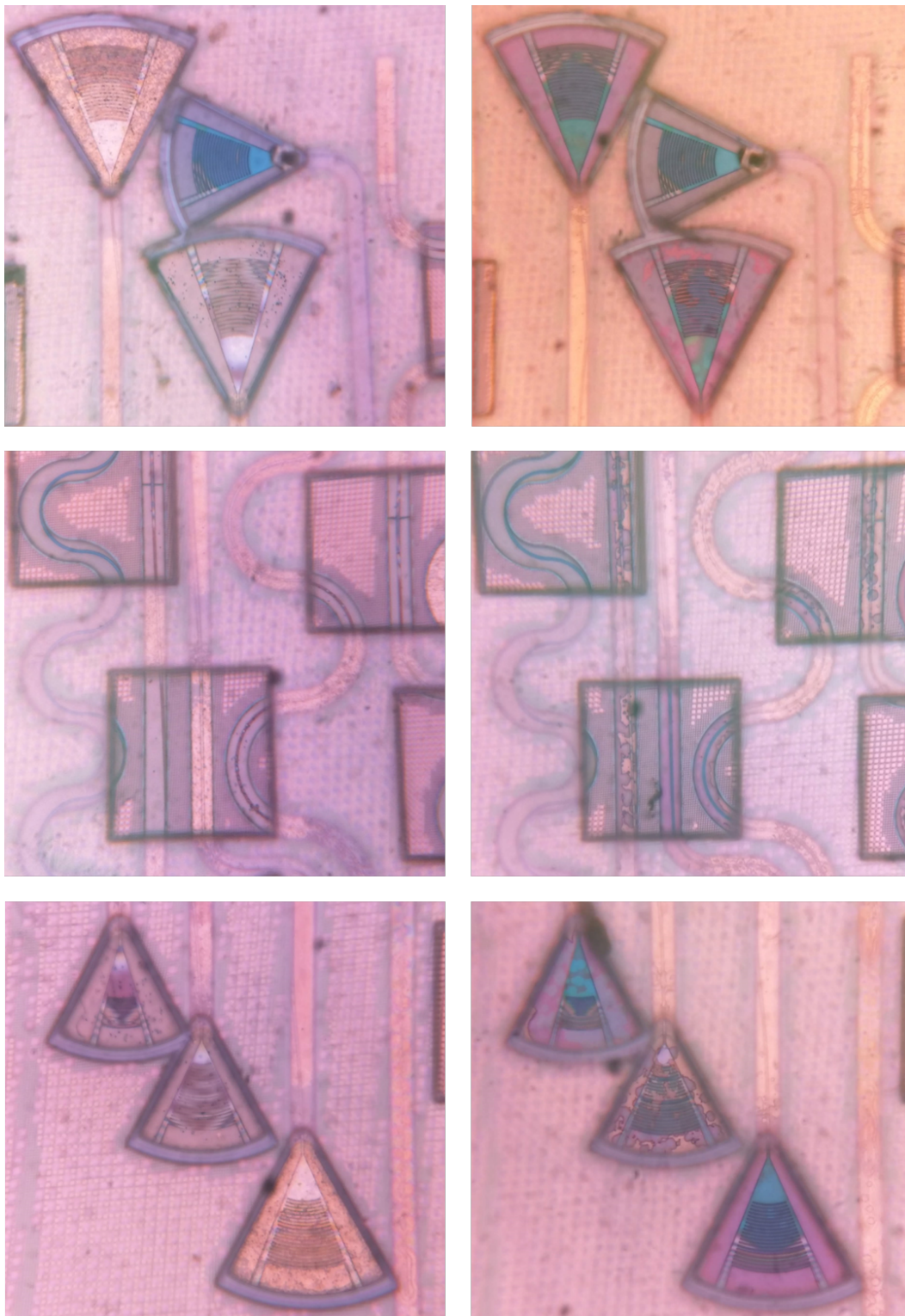


Figure 2.12: Illustration of etching progress. Left: Test structures after 1 hour in a 1:1 H₂O₂/EDTA solution at 60°C. Right: Test structures after 1 hour in a refreshed 1:1 H₂O₂/EDTA solution at 60°C.

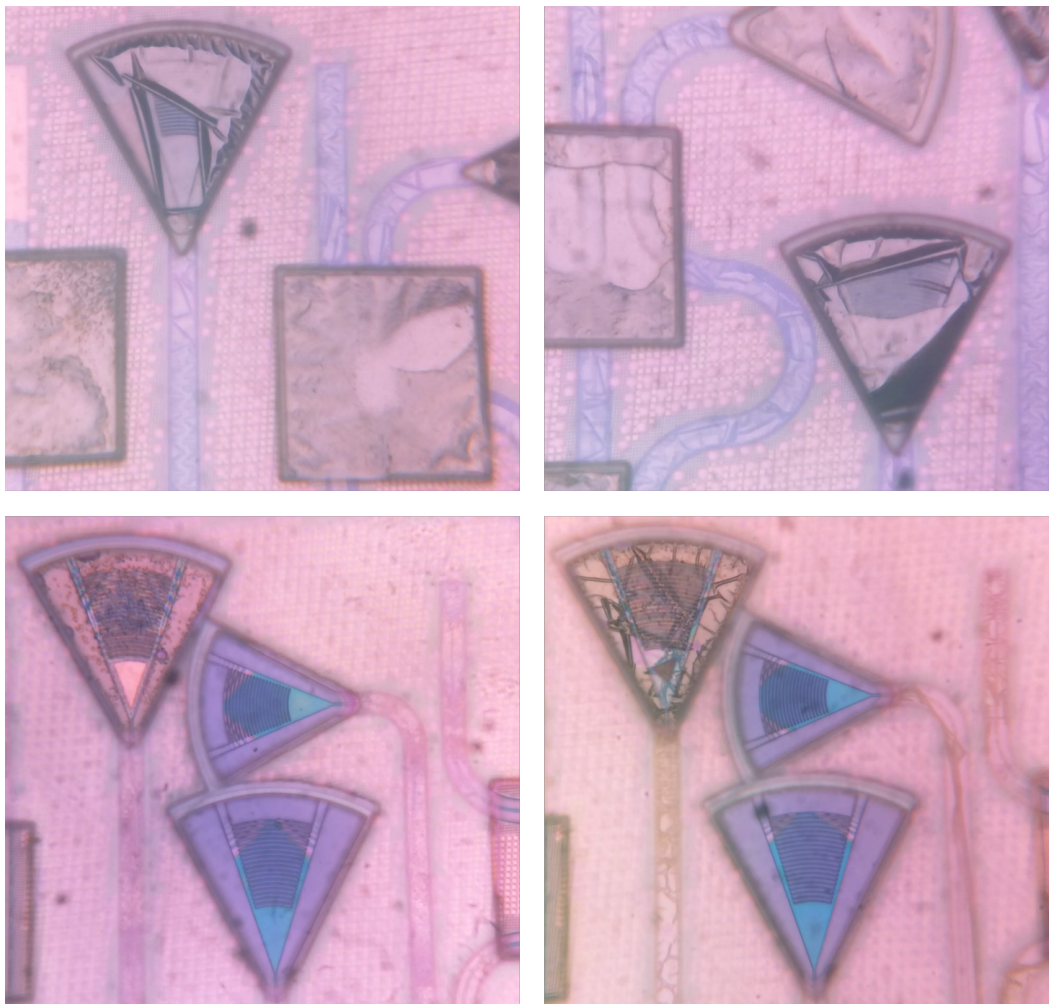


Figure 2.13: Results of expedited etching. Top: Test structures after 1 hour in AETA at 80°C. Bottom left: Test structures after 1 hour in 1:1 H₂O₂/EDTA at 80°C. Bottom right: Test structures after another hour in AETA at 80°C.

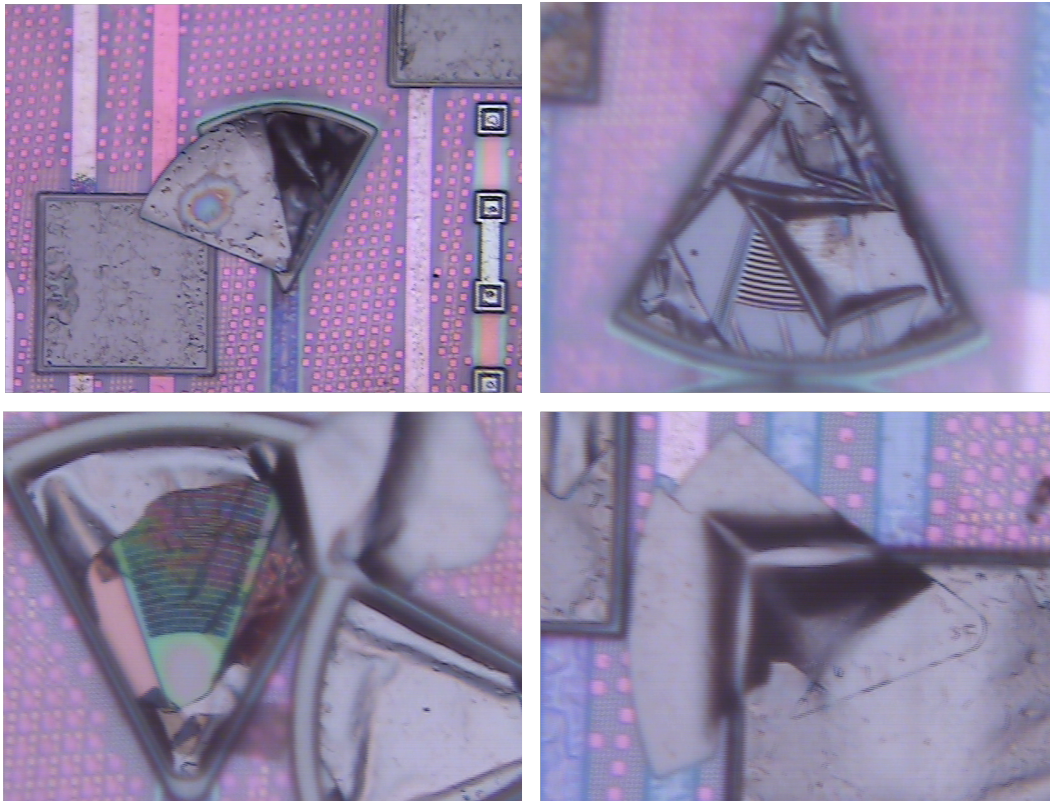


Figure 2.14: Some results from a three part etching sequence. Top left: Aluminum pad peeling off after 2 hours in AETA at 80°C. Right: Test structures after an additional hour in AETA at 80°C. Bottom left: An exposed V7 grating coupler after 2 hours in H₂O₂/EDTA at 40°C.

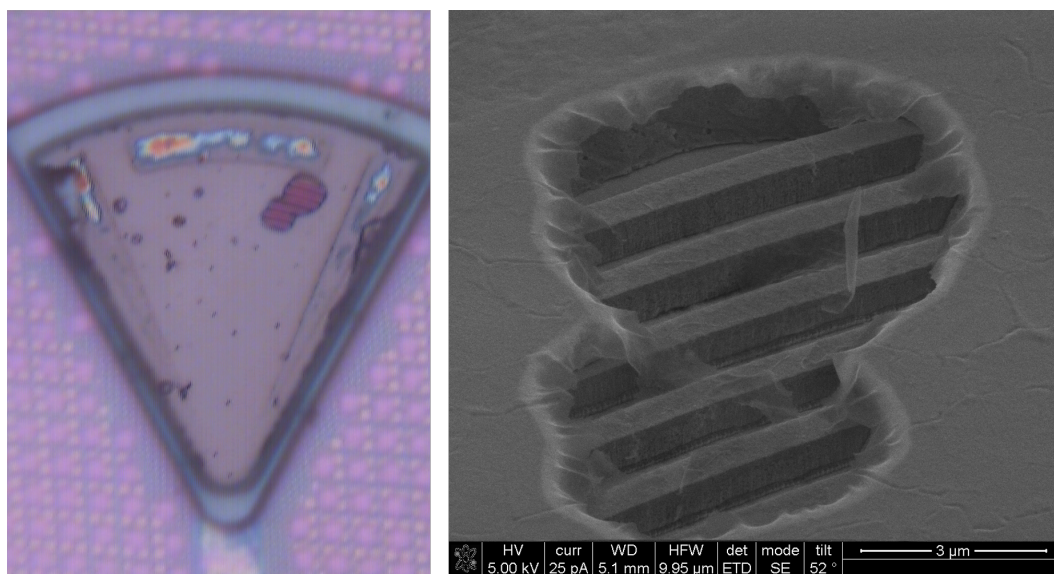


Figure 2.15: A partially exposed grating coupler demonstrating the nature of the barrier layer. SEM image courtesy A. Khachaturian.

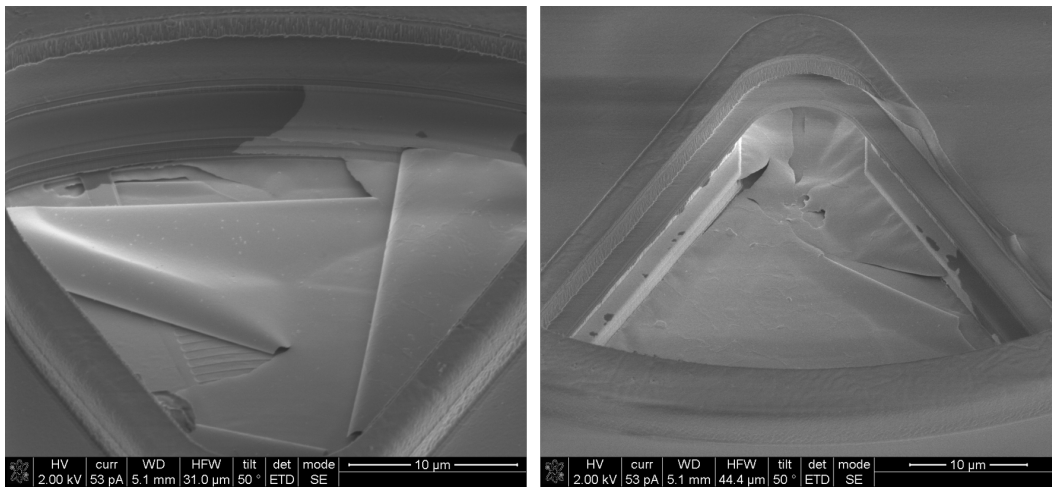


Figure 2.16: Partially etched grating couplers demonstrating the thinness of the barrier layer. SEM images courtesy A. Khachaturian.

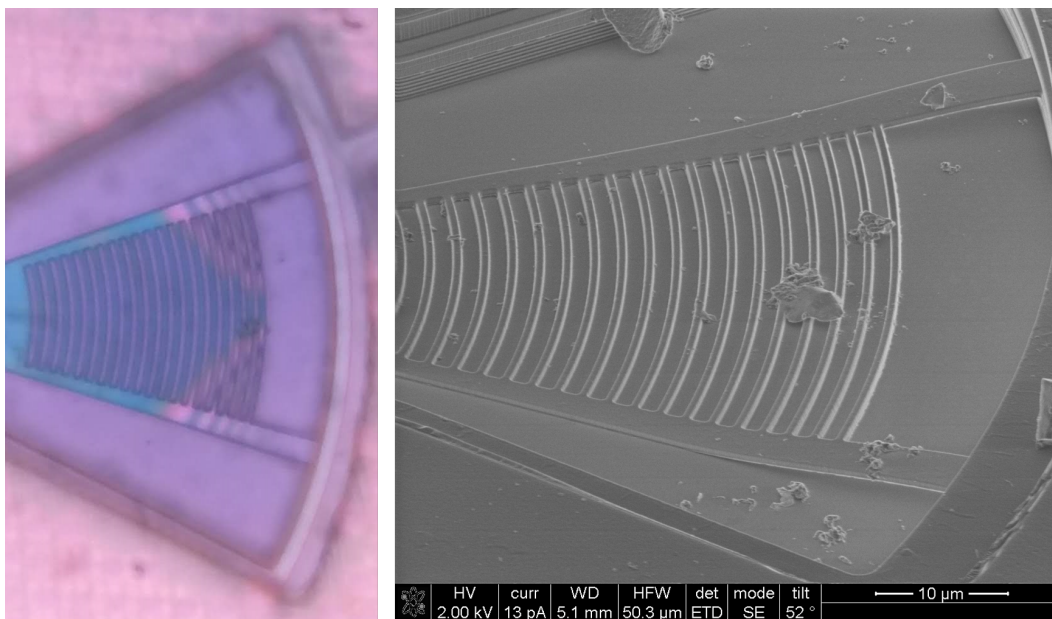


Figure 2.17: A collapsed grating coupler formed from the V7 layer. SEM image courtesy A. Khachaturian.

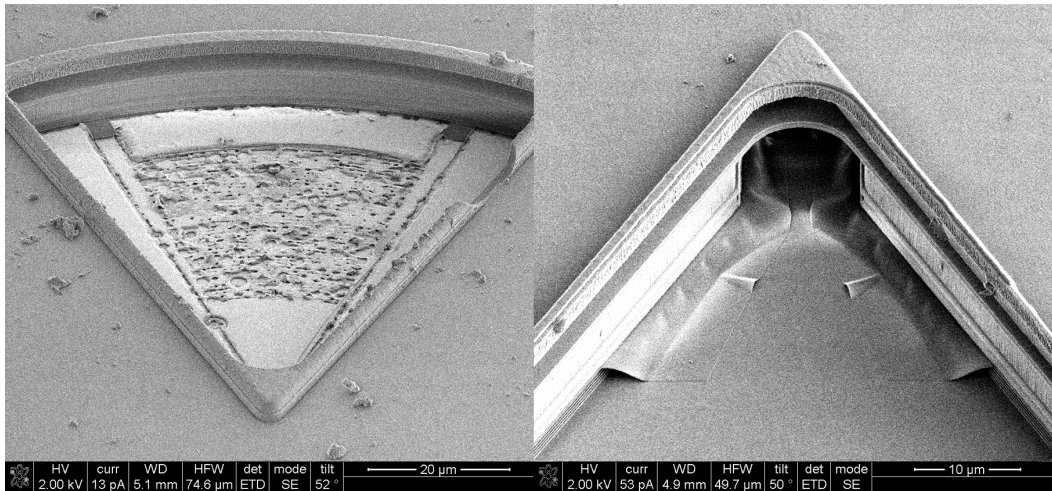


Figure 2.18: Remnants of grating couplers formed using the M7 layer as a ridge and the V6 layer as a slab. Left: A partially etched cavity with the support remnants visible against the back wall of the cavity. Right: A fully etched cavity with the V6 slab remnants visibly collapsed onto the bottom of the cavity. SEM images courtesy A. Khachaturian.

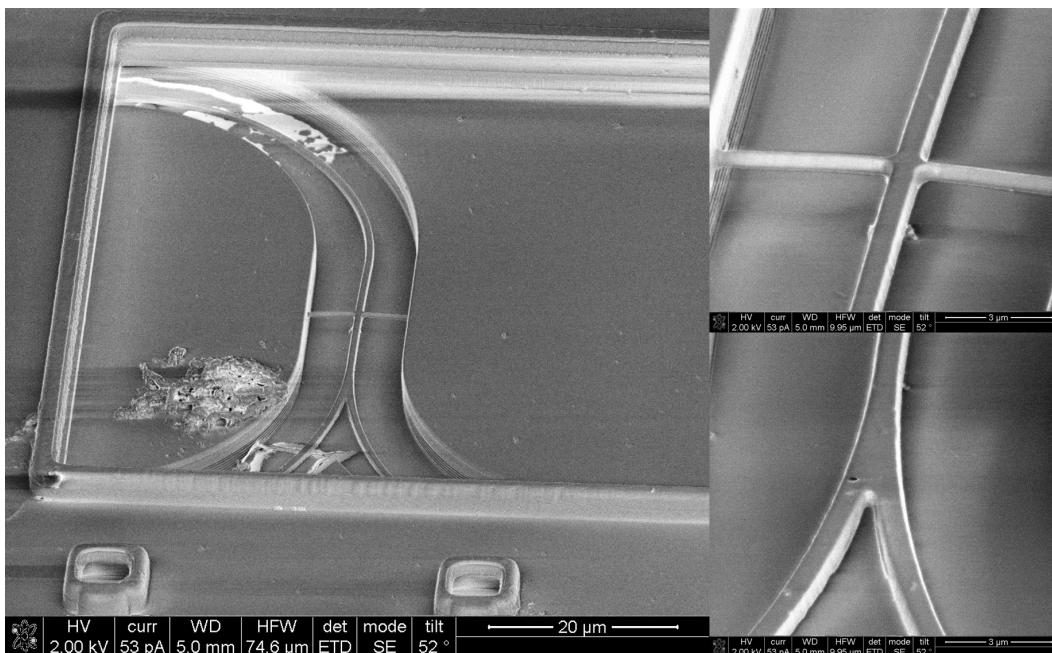


Figure 2.19: A suspended channel waveguide. SEM images courtesy A. Khachaturian.

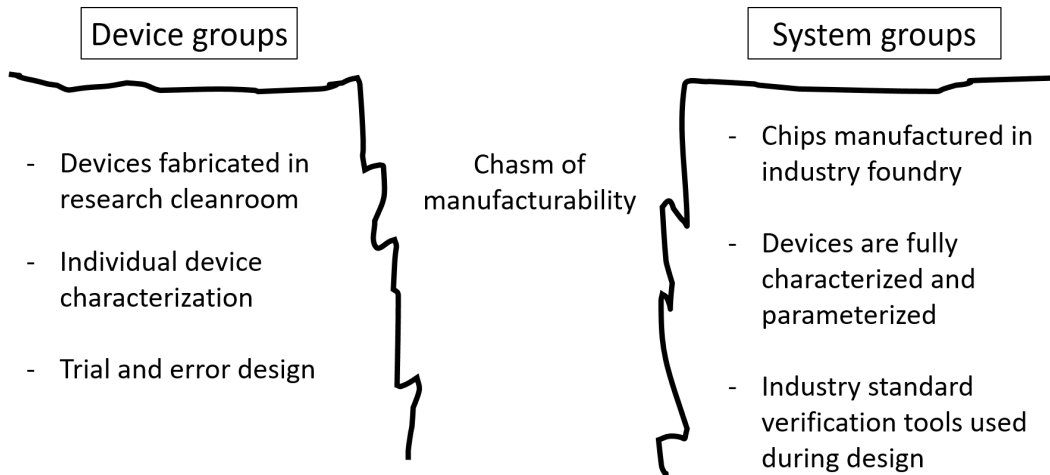


Figure 2.20: The gap between device and system research groups.

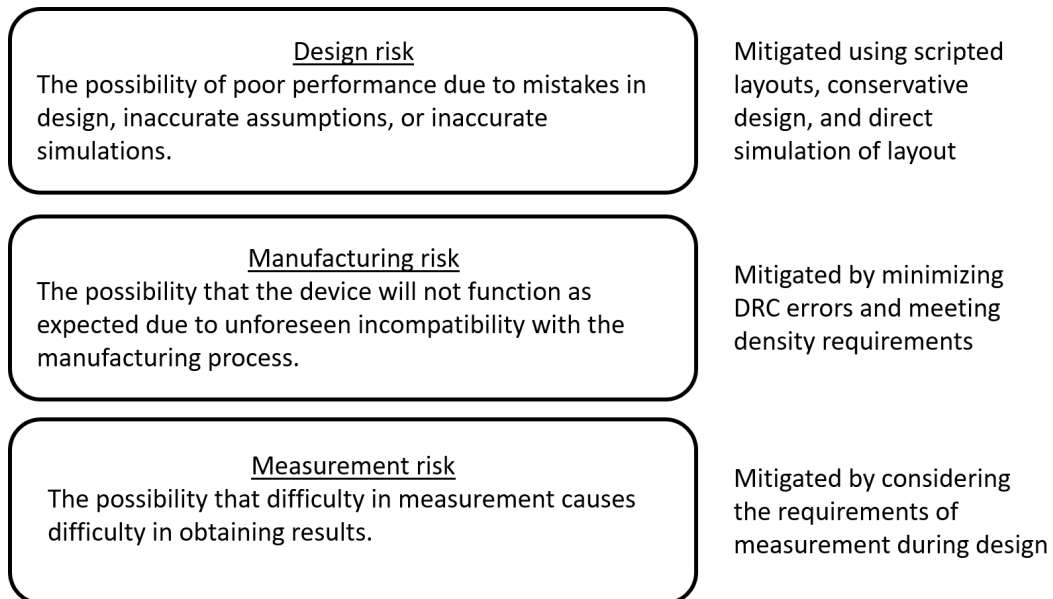


Figure 2.21: The risks that the subtractive photonics project is exposed to, and means of mitigating those risks.

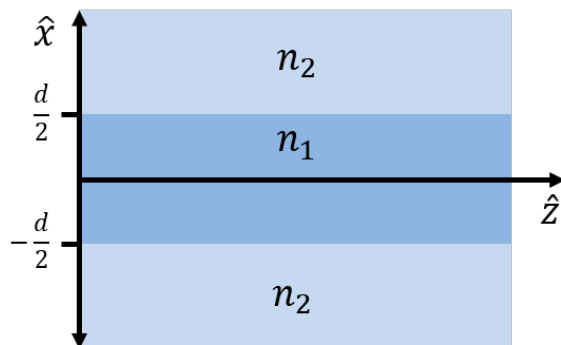


Figure 2.22: A symmetric slab waveguide, with refractive index variation only in \hat{x} .

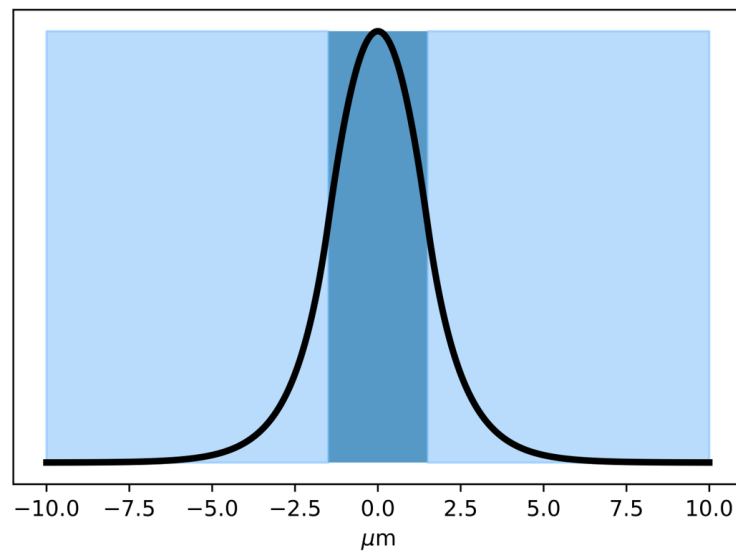


Figure 2.23: The fundamental mode of a 1D waveguide with a $3 \mu\text{m}$ wide slab at a wavelength of 1550 nm . The index of the core is 1.495 and the index of the cladding is 1.466 .

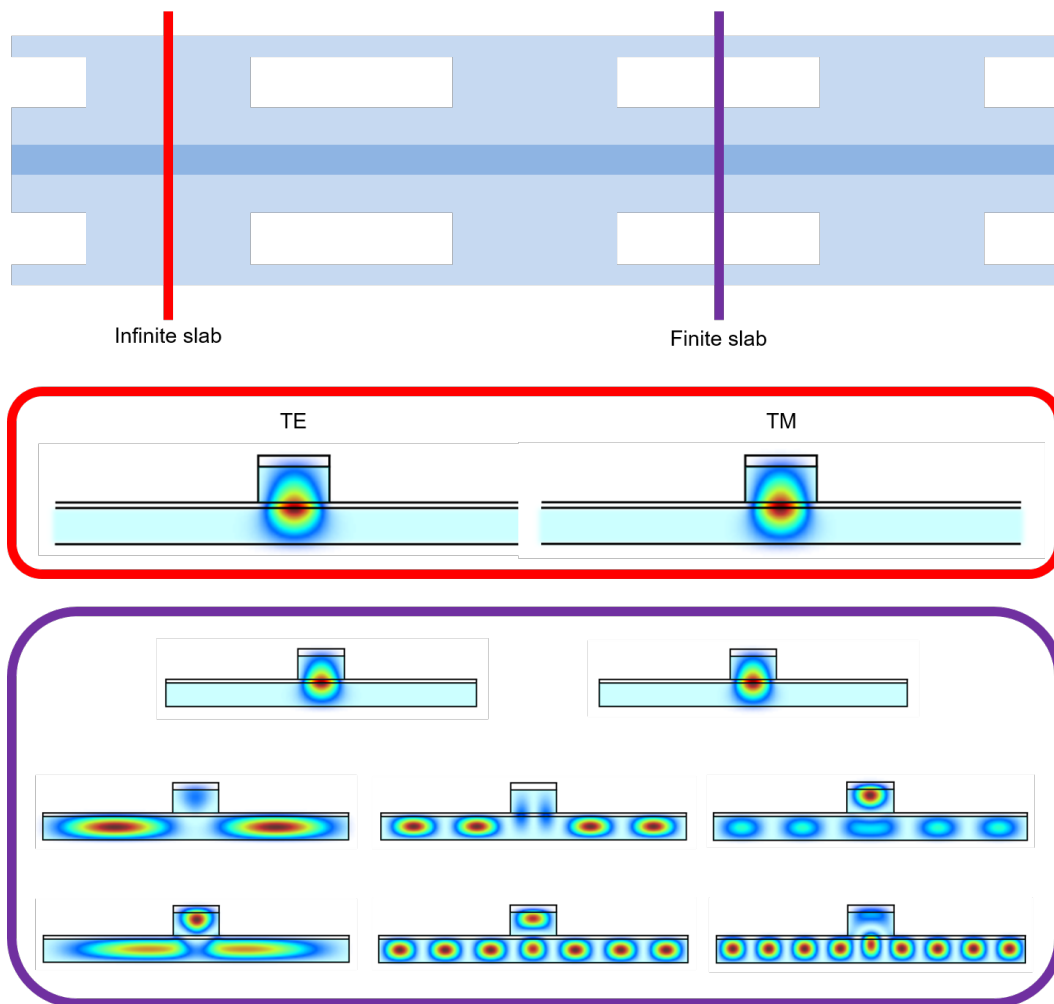


Figure 2.24: Example of a rib waveguide that is single mode with an infinite slab, but multimode with a finite slab. Note that for the case of the finite slab, many more modes are supported than are shown.

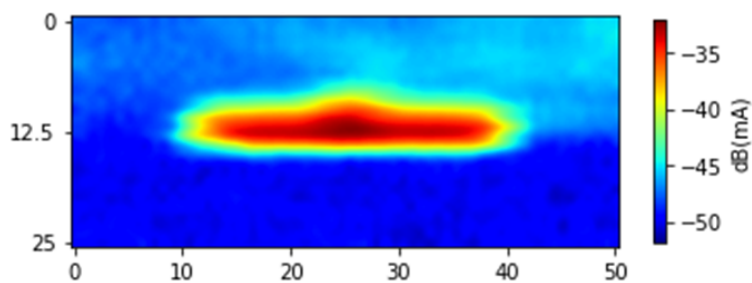


Figure 2.25: Imaging a waveguide facet by scanning a fiber across it.

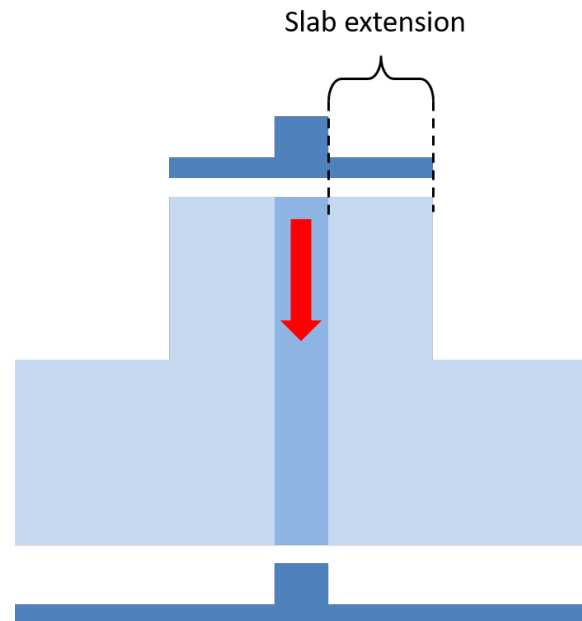


Figure 2.26: Definition of slab extension and the two waveguide widths used in mode overlap calculations.

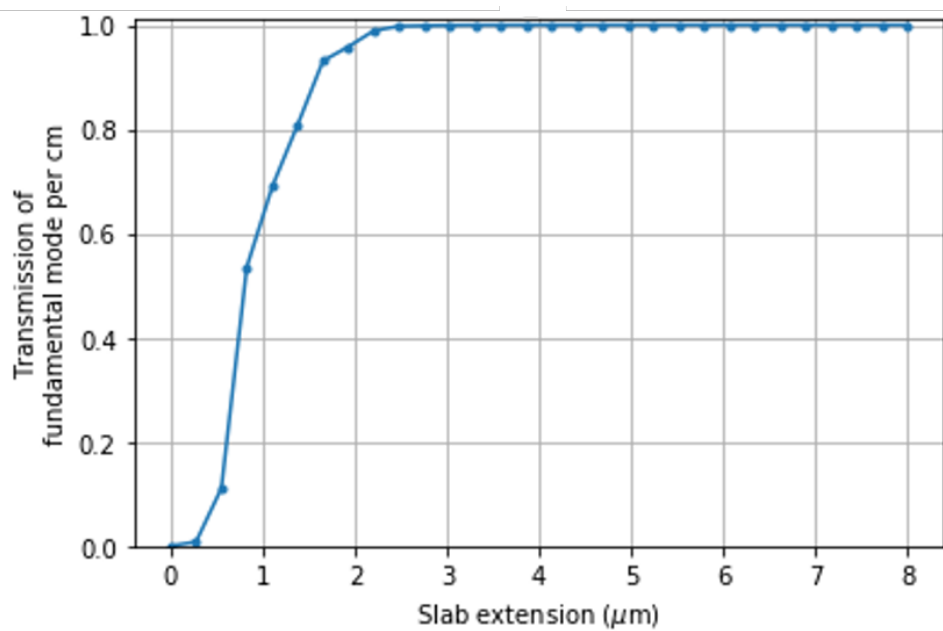


Figure 2.27: Optimization of slab extension for a V5M5-V4M4 waveguide.

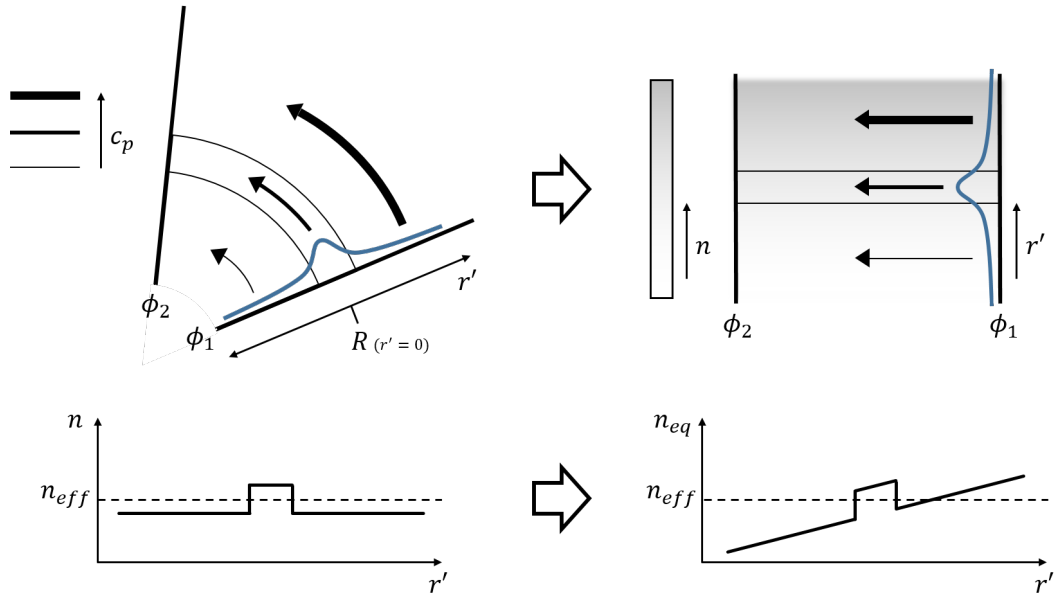


Figure 2.28: Modeling a bend using an index transformation. Left: A bent waveguide has an index profile given by n_{eq} . Phase velocity increases radially in the bend in order for the wavefront at ϕ_1 to arrive in phase at ϕ_2 . Right: The bend is modeled using a straight waveguide with a refractive index gradient to compensate for the increasing phase velocity.

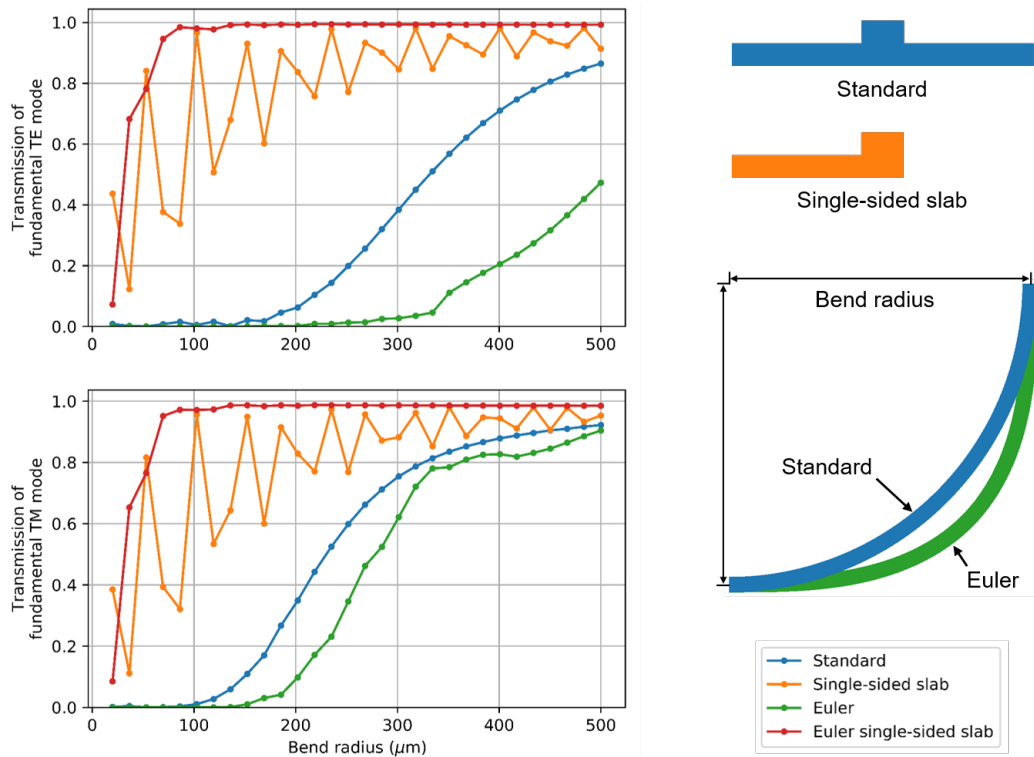


Figure 2.29: Bend loss of a V5M5-V4M4 waveguide.

Chapter 3

DESIGNS AND EXPERIMENTAL RESULTS

The rib waveguides introduced in section 2.4 have some parameters requiring experimental validation, such as the minimum slab thickness and the minimum ridge thickness. In this chapter, we will explore the variety of experimental waveguides we fabricated and the lessons learned from them.

All subsequent tapeouts used a 180 nm bulk CMOS process. There were a number of reasons for switching from 65 nm to 180 nm:

- There is a greater tolerance for DRC violations.
- The area is five times cheaper.
- There are a greater number of MPW shuttles available.
- The fabrication turnaround is faster.
- Higher supply voltages are better for modulators and photodetectors.
- 180 nm is already commonly used in non-standard processes such as MEMS and CMOS image sensors.

The greater tolerance for DRC violations, lower cost, and fast turnaround all made experimenting much easier than in 65 nm.

The 180 nm process used from this point forward possesses six metal layers and five via layers, as depicted in Fig. 3.1. This process uses an aluminum-copper alloy interconnect with tungsten vias, which means it most likely uses a Ti/TiN barrier bilayer and fluorosilicate low- k dielectric layers. The fluorosilicate glass has a refractive index of 1.43 [86], [87]¹, which is not far from the refractive index of glass, 1.46 [89]. Nevertheless, a waveguide mode modeled using the entire dielectric stackup, including low- k layers, will look distorted compared to a waveguide mode modeled using only glass, as shown in Fig. 3.2. Consequently, simulations requiring the most accuracy should model the entire dielectric stackup.

¹Based on similar papers, such as [88], this refractive index was most likely measured at 633 nm.

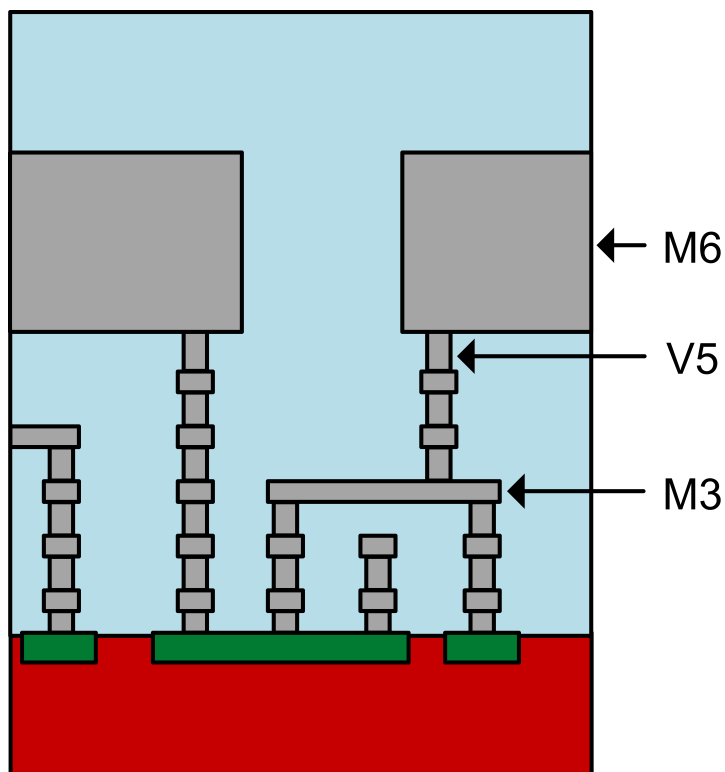


Figure 3.1: Depiction of the interconnect layers of the 180 nm process.

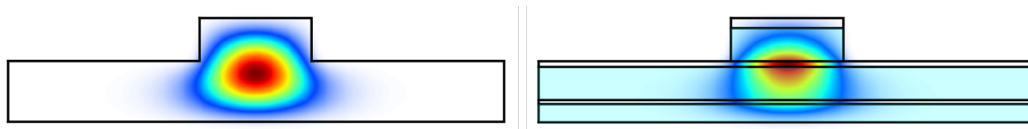


Figure 3.2: Example of the distortion caused by modeling the full dielectric stackup. The light blue regions denote low-k dielectric.

3.1 Risk Mitigation in 180 nm

As described in Figure 2.21, the project risks can be divided into design risk, manufacturing risk, and measurement risk. The impacts of each of these will be considered in turn.

Design Risk

The three main design risks under consideration for TO1 are shown in Table 3.1. The first, that of high loss loss due to supports on a channel waveguide, can be mitigated by switching to a rib waveguide geometry, as described previously. The second design risk is poor fiber coupling and the third risk is that the waveguides might collapse due to insufficient mechanical support.

Risk	Mitigation
High loss due to supports	Rib waveguide
High loss due to fiber coupling	Large cross section waveguide
Waveguides collapse	Thick slab

The 180 nm process used for this tapeout is optimized for RF systems, so it has an ultra thick top metal (UTM) layer that is several micron thick. This thick metal is ideal for fabricating relatively large multimode waveguides, which have two advantages:

- The fundamental mode will experience low loss relative to other waveguide geometries in this platform due to the greater confinement.
- Low loss edge coupling is possible when using a visible wavelength fiber, which have core diameters around $5\ \mu\text{m}$, or a lensed infrared wavelength fiber, which have spot diameters around $5\ \mu\text{m}$.

The second advantage, low loss edge coupling, is not only an advantage over singlemode waveguides in this process, but also an advantage over other photonics processes in general. In silicon photonics, for example, the waveguides are typically 220 nm thick and 400-500 nm wide, presenting an enormous mode mismatch between the waveguide and the core of a fiber. Although waveguide tapers are often used to improve the mode matching, these tapers consume chip area that the multimode waveguides in this process do not need at all. The low index contrast of the glass waveguides further reduces reflections at the edge coupling interface.

With these considerations in mind, a waveguide was designed that uses the UTM layer, which is also the M6 layer, as the ridge, and the V5, M5, and V4 layers as the slab, referred to as a M6-V5M5V4 waveguide according to the convention of this thesis (Fig. 3.3). A three layer slab was chosen in consideration of the fragility of the slab layers in TO0. The only surviving support in TO0 (Fig. 2.19) used a layer more than half a micron thick, which provided a very rough estimate that the slab of the waveguide should be at least $1\ \mu\text{m}$ thick. Given the size of the M6 layer, however, confinement is not significantly impacted by increasing the slab thickness further to $2\ \mu\text{m}$, an even more conservative value.

The $5\ \mu\text{m}$ core diameter and spot diameter mentioned above suggest that a good width for the ridge of this multimode waveguide would be $5\ \mu\text{m}$, since the UTM layer is about $5\ \mu\text{m}$ thick. At this width, the M6-V5M5V4 waveguide supports six confined

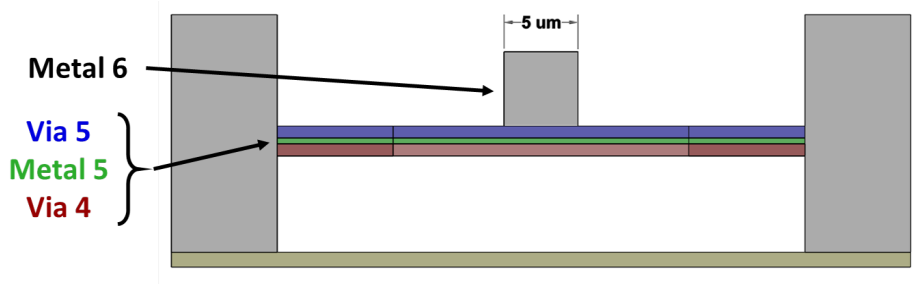


Figure 3.3: Schematic cross section of the M6-V5M5V4 waveguide geometry.

modes at a wavelength of 1550 nm, shown in Figure 3.4. The normalized index contrast of the confined modes is plotted against ridge width in Fig. 3.5. Although the M6-V5M5V4 waveguide is singlemode at a ridge width of $2 \mu\text{m}$, the high aspect ratio of the resulting ridge would force the mode to be confined mainly in the slab, reducing overall mode confinement and increasing bend loss accordingly.

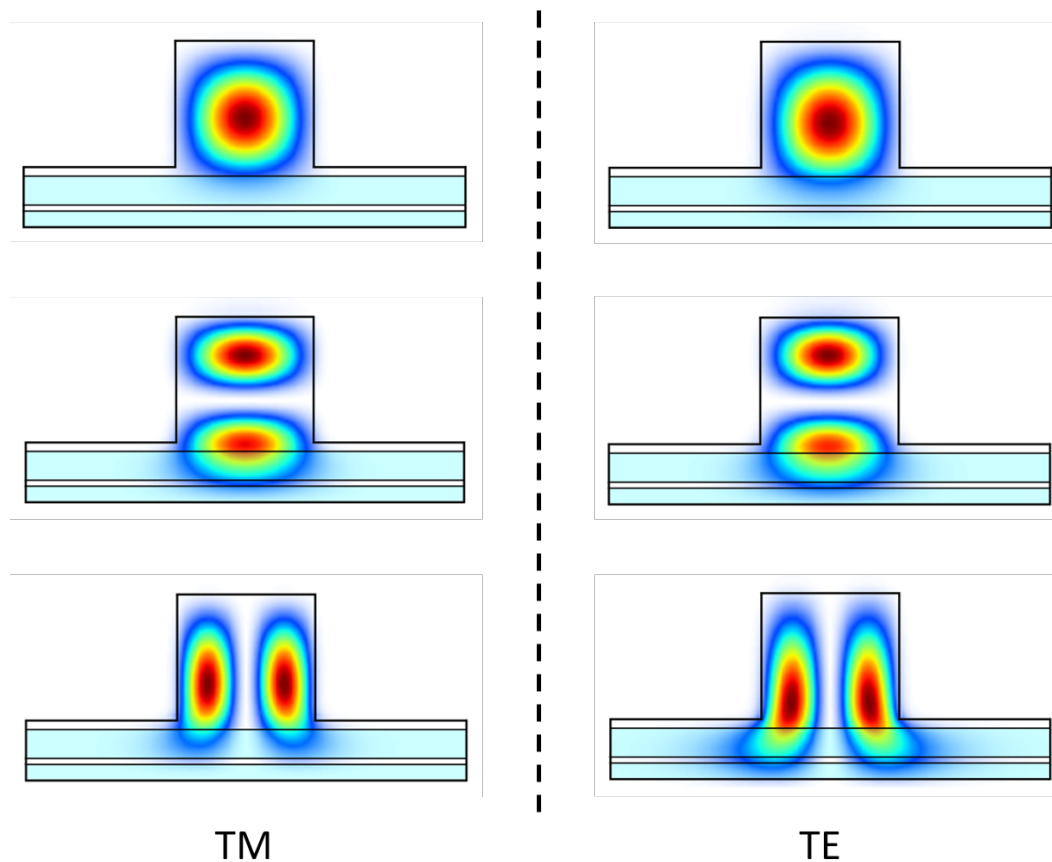


Figure 3.4: Confined modes of the M6-V5M5V4 waveguide at a wavelength of 1550 nm and a ridge width of $5 \mu\text{m}$.

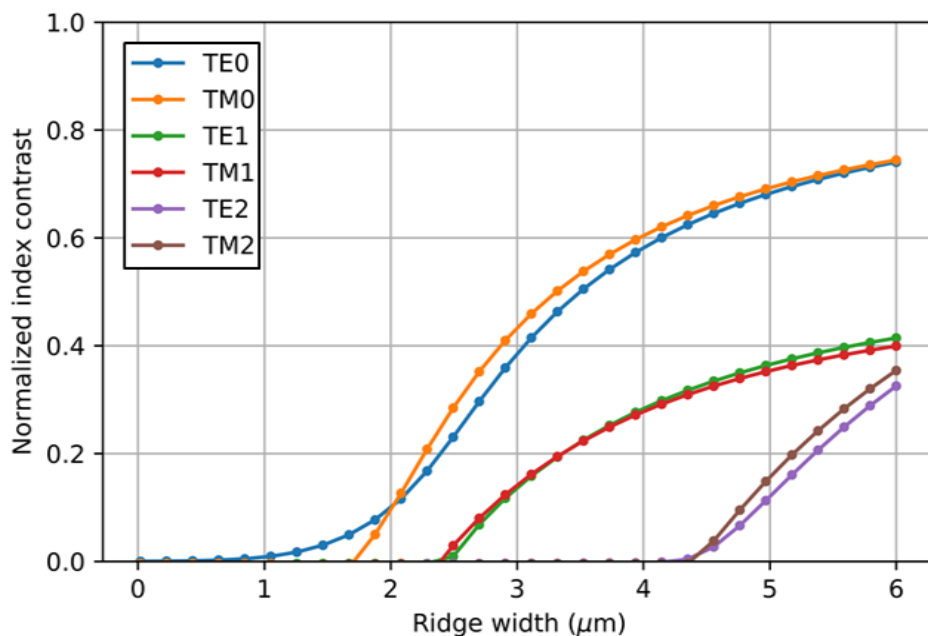


Figure 3.5: Normalized effective index of the confined modes of the M6-V5M5V4 waveguide across ridge widths at a wavelength of 1550 nm.

Manufacturing Risk

The manufacturing risks considered in TO1 are shown in Table 3.1.

Risk	Mitigation
Waveguides fail due to layer interface problems	Multiple waveguide geometries
Foundry rejects chip because of metal-via overlap violations	Sidewalls with metal-via overlap greatly reduce DRC violations
Foundry rejects chip because of off-grid coordinates	Gridding algorithm fits all polygons to a 5nm grid
Foundry rejects chip because of polygons with angles other than 90° and 45°	Gridding algorithm forces all angles to 90° and 45°
Foundry rejects chip because of acute angles	All corners on slab layers are rounded until no acute angles remain
Foundry rejects chip because of one-big via violation	All V5 metal is converted into waffling

In considering manufacturing risk, it is worthwhile at this point to return to TO0. The total number of design rule check (DRC) violations in TO0 was 32,620,405 - an impressive number considering almost all electronic chips have zero DRC

violations. For electronic chips the DRC rules provide a very useful check on the manufacturability of the system, so electronics designers have strong incentive to eliminate all DRC violations. The manufacturers are not concerned with the functionality of the systems being fabricated though, and so in practice many DRC rules can be violated with no consequence to the physical manufacturability of the system.

For one subset of rules, however, the manufacturer must enforce compliance. The density rules require that each metal layer have a certain minimum density to ensure adequate local and global planarization [31]. Global planarization in particular must be kept within tolerance since many different customers may have designs on one wafer. For all the subtractive photonics tapeouts described in this thesis, density requirements were met or nearly met, ensuring that other designs on the wafer were not affected.

The DRC errors of TO0 mainly consisted of two categories:

- Design rules for optical proximity correction that require edges smaller than the minimum width of the layer cannot be adjacent to another edge smaller than the minimum width of the layer.
- Enclosure rules that require metal layers contacting vias to extend beyond the via (Fig. 3.6).

A breakdown of the DRC errors for TO0 is shown on the left side of Figure 3.7. The violations related to optical proximity correction rules constitute the majority of the errors and are due to the output of the gridding algorithm, which is described below. These errors cannot be mitigated without eliminating the gridding algorithm entirely, which adds to the overall risk of rejection (see Table 3.1). The violations due to enclosure rules, however, can be mitigated by ensuring that all non-optical vertical surfaces in the chip have metal layers overlapping via layers, as shown on the right side of Figure 3.7. Non-optical vertical surfaces are those surfaces that the waveguide mode does not interact with significantly, such as the slab edges or the cavity sidewalls.

Additional manufacturing risk comes from polygons with coordinates that are off the 5 nm grid, polygons with angles other than 90° and 45° , and polygons with acute angles. Part of the risk here is that some polygons may not get fabricated as intended, depending on the foundry's processing of the GDS file, and part of the risk is that the

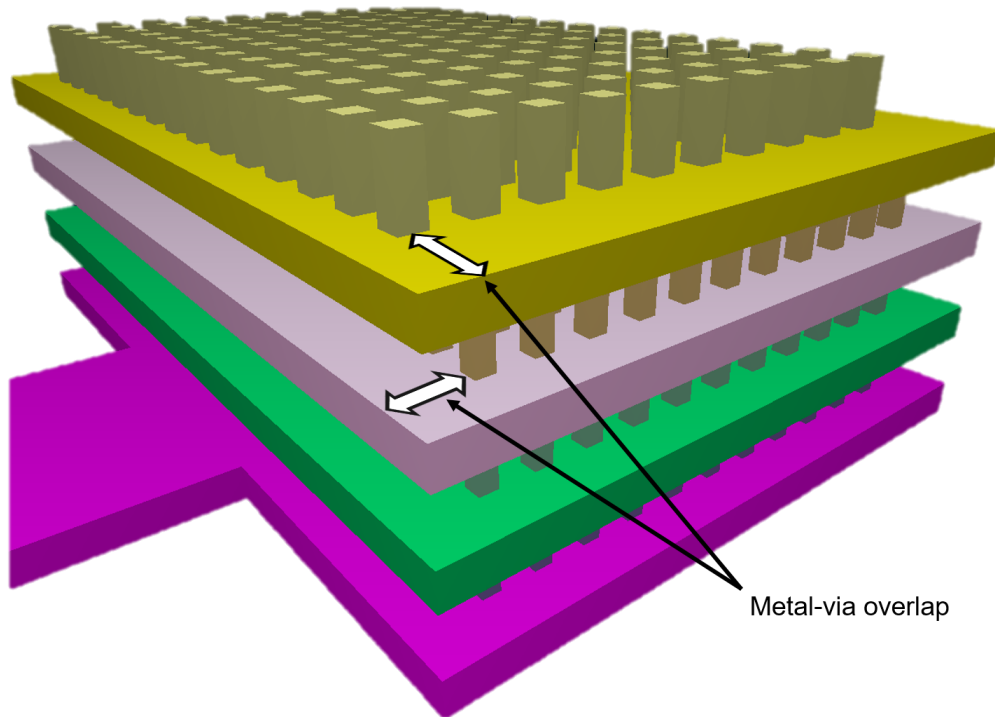


Figure 3.6: Illustration of the enclosure rules. The metal layers must extend beyond the vias.

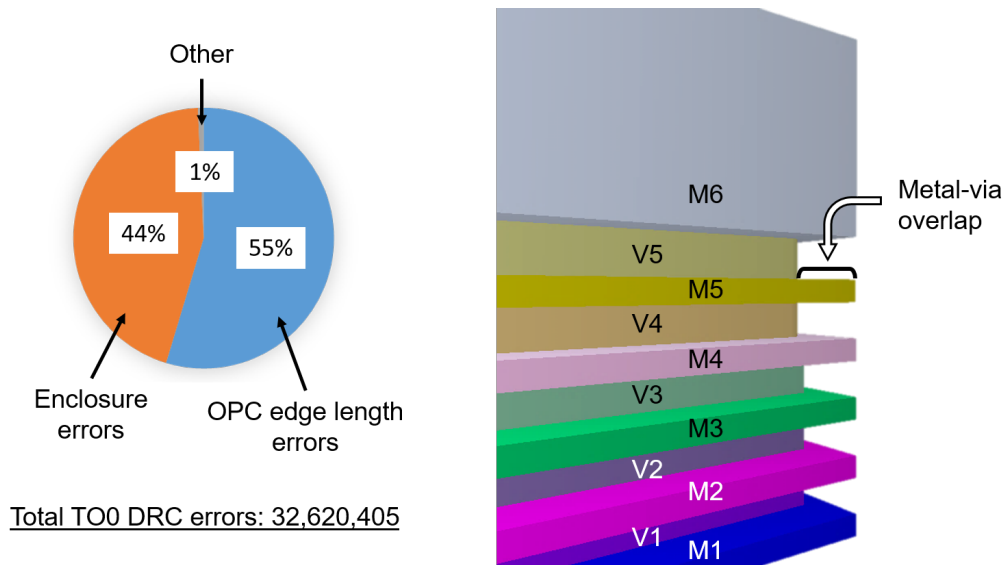


Figure 3.7: Mitigating DRC errors. Left: Breakdown of the DRC errors in TO0. Right: Illustration of the corrugation used on vertical surfaces to mitigate enclosure violations.

foundry might outright reject the chip for excessive violations, as with the metal-via overlap. The solution used in TO1-TO4 was to pass all of the layout's polygons through a gridding algorithm. First, the algorithm ensures that all polygon vertices are on a 5 nm grid, and that all angles are 90° or 45° . As shown in Figure 3.8, it does this by adding vertices to the polygon until the angle requirement is satisfied. However, acute 45° angles can still remain. To eliminate this problem, polygons on the slab layers are rounded until no acute angles remain, as shown in Figure 3.9. Only the slab layers receive this treatment because the other layers are generally smoothly curving.

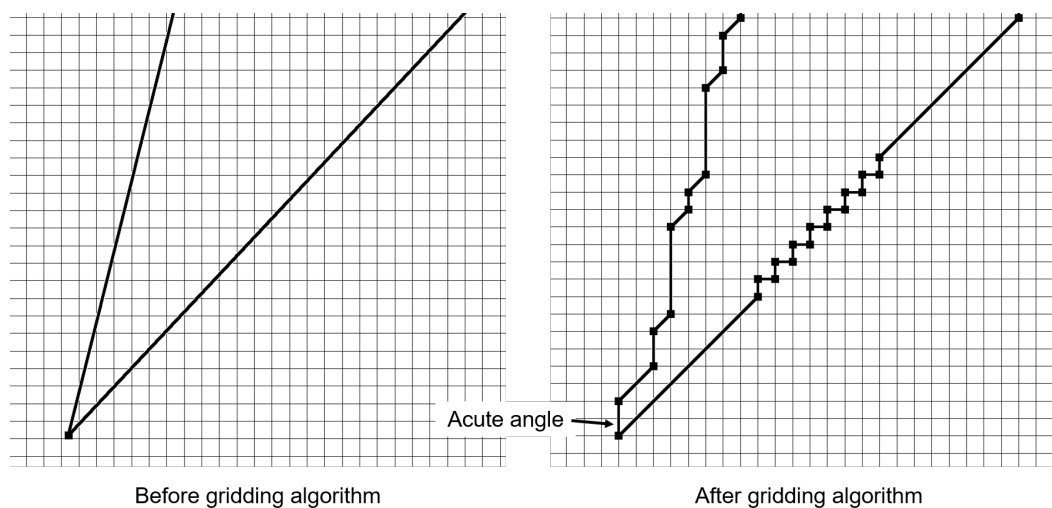


Figure 3.8: Illustration of the gridding algorithm. Left: A polygon corner before gridding. Right: The same polygon corner after gridding. Note the acute angle that must be removed by other means (Fig. 3.9).

The 180 nm process used in TO1, optimized for RF systems, caused difficulties with the top via layer (V5). The process forbids large contiguous areas of metal in this layer, obstructing normal patterning of the metal. Ideally the polygons on the V5 layer, which would normally be fully metal, would be filled using a pattern that minimized DRC violations, and thus, fabrication difficulties. Since V5 is a via layer, a via array is a natural fill pattern, and fully meets DRC rules. Another structure that meets DRC rules, and hence guarantees manufacturability, is the slot via found in the seal ring. The slot via can be patterned around the perimeter of any polygon that needs to be removed, and the interior of the polygon can be filled with vias. Since it was unclear how difficult it would be to remove these large pieces of dielectric during the etching process, the polygon interior is further split into squares about $5\ \mu\text{m}$ in length on each side, as a compromise between ease of etching and ease

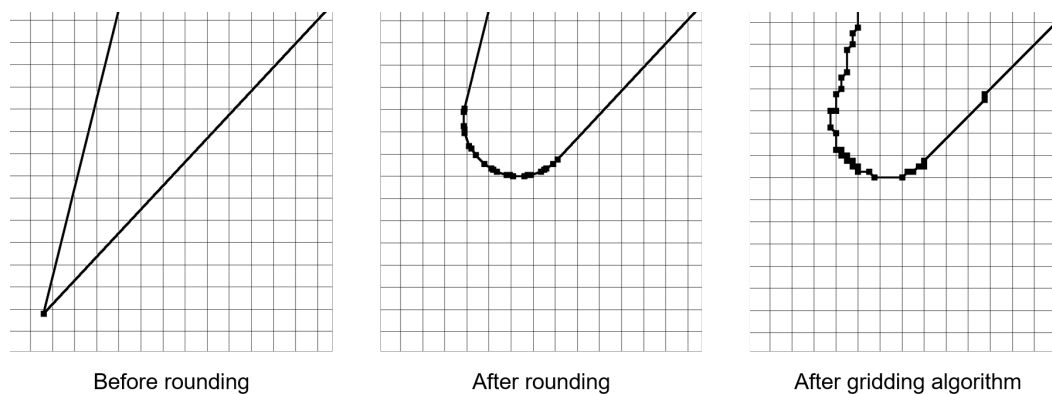


Figure 3.9: Illustration of the polygon rounding procedure. Left: A polygon corner before rounding. Center: The same polygon corner after rounding. Right: The same polygon corner after gridding. Note the absence of acute angles.

of manufacturing. The resulting polygon fill is termed waffling since the dielectric squares resemble waffles.

Measurement Risk

The measurement risks considered in TO1 are shown in Table 3.1.

Risk	Mitigation
Polishing damages waveguide facets	Waveguides are recessed from dice line for lithographically defined facet
Recess is too small so that dicing damages waveguide facets	Conservative dicing spacing of $13 \mu\text{m}$
Waveguides cannot be measured easily due to setup constraints	Test structures designed for compatibility with specialized instruments
Propagation in waveguide and waveguide cavity are difficult to distinguish	S-shaped test structures

After fabrication, the chips must be diced to provide access to waveguide edge couplers. One possibility is to dice through the waveguide and then polish the facet to remove any damage due to dicing, as shown in the upper part of Figure 3.11. However, the polishing may damage the waveguide or may not provide as smooth a waveguide facet as lithography alone. In order to obtain a lithographically defined waveguide facet, the facet is recessed from the end the waveguide cavity. Dicing within the cavity extension, as shown in the bottom of Figure 3.11, can then be followed by etching to expose the waveguide facet. This strategy raises the question

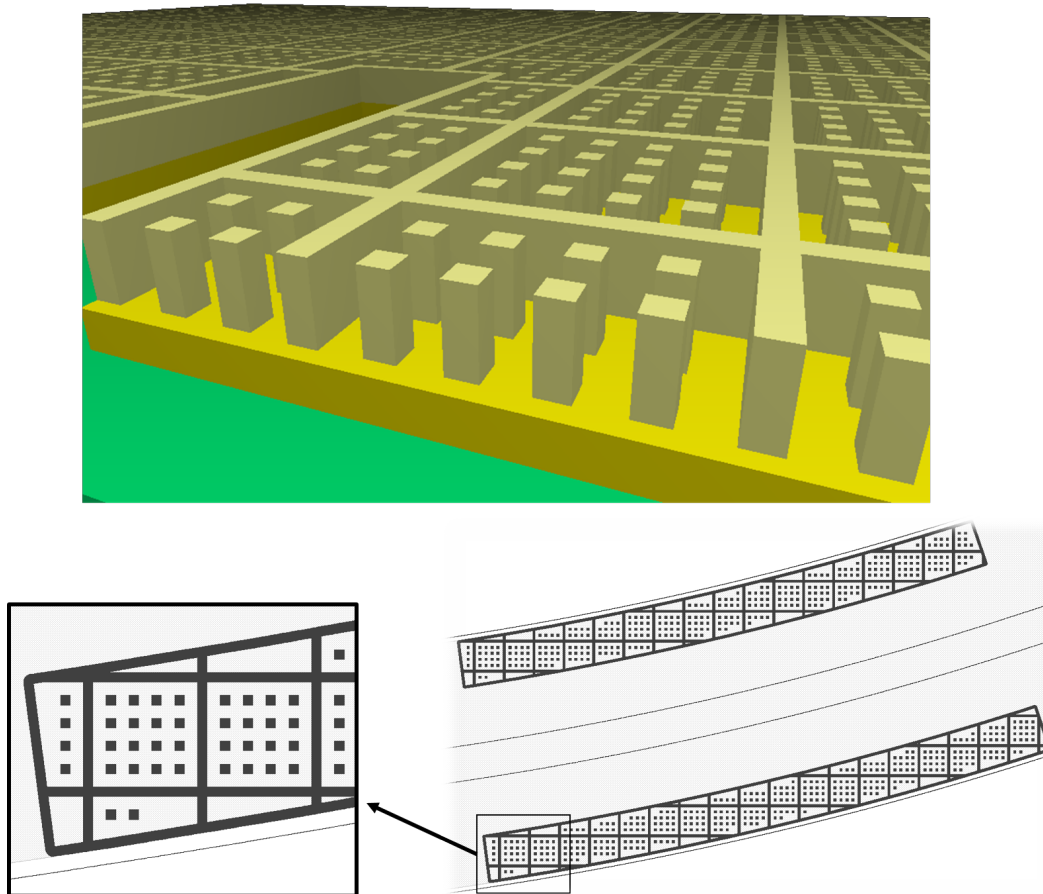


Figure 3.10: An example layout showing waffling on the V5 layer, along with a 3D model of waffling.

of where exactly the dice line should be located within the cavity extension. If the dice line is too close to the facet, as shown in the upper part of Figure 3.12, then the facet may lie within the dicing tolerance and the dicing may completely destroy the waveguide facet. Even if the dicing does not cut through the facet, however, it may come close enough to indirectly damage the glass.

The dicing tolerance is normally about $10\ \mu\text{m}$ on either side of the nominal dice line for the dicing vendor used in this thesis. Moreover, the working distance of a lensed fiber with a $5\ \mu\text{m}$ spot size at $1550\ \text{nm}$ is $26\ \mu\text{m}$. Assuming that it is equally likely for the cut to end up closer to the waveguide as farther away, the best choice for the dice line is half of the working distance of the lensed fiber, or $13\ \mu\text{m}$ from the waveguide facet. This allows enough room for the cut to be $10\ \mu\text{m}$ closer to the facet than nominal without damaging the waveguide, while also enabling the facet to remain within the working distance of the fiber if the cut is $10\ \mu\text{m}$ farther than the

nominal dice line.

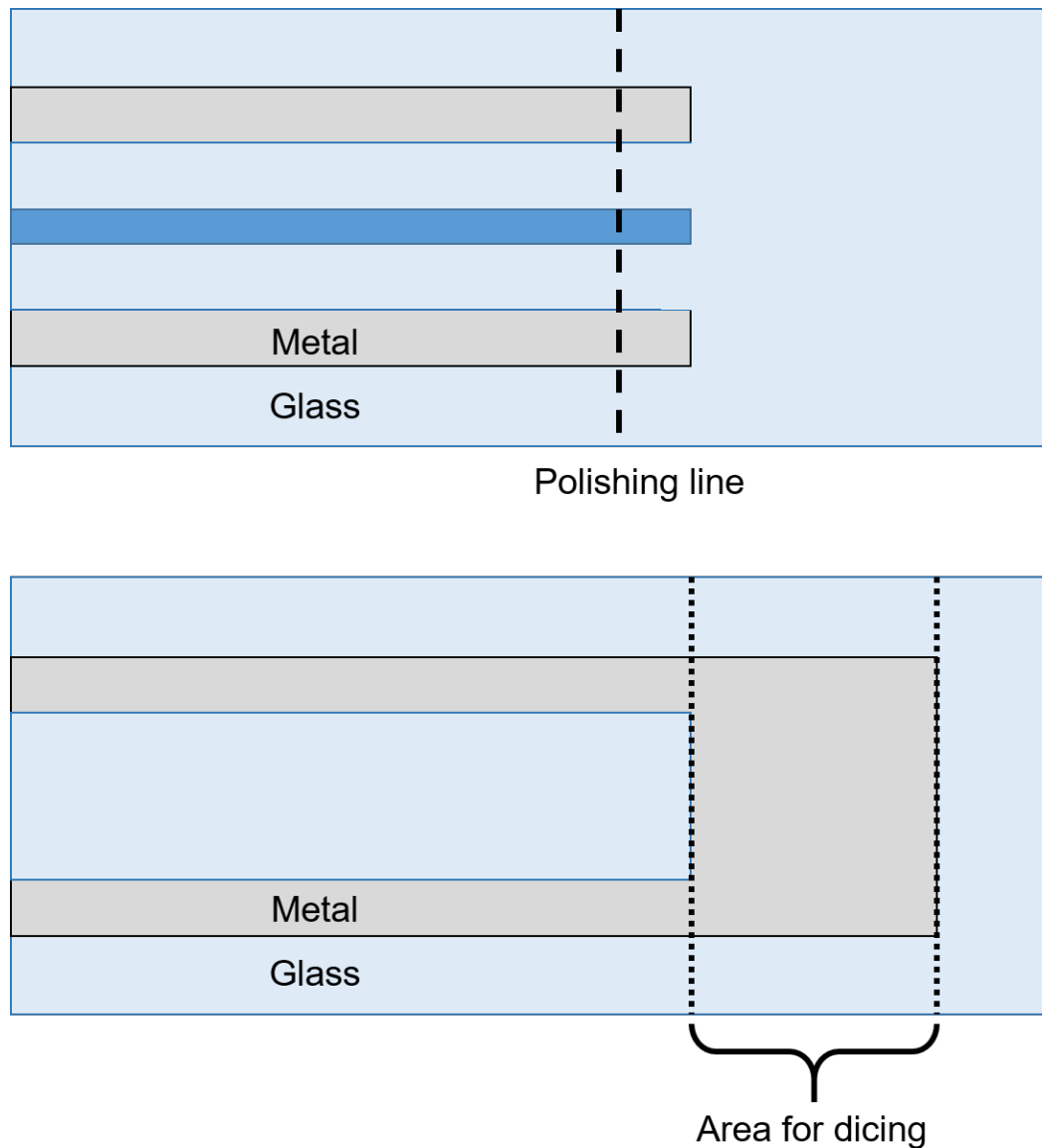


Figure 3.11: Defining waveguide facets. Top: Waveguide layout for dicing followed by polishing. Bottom: Waveguide layout for obtaining a waveguide facet defined by lithography by dicing and then etching.

The impact of the dicing tolerance on the efficacy of edge coupling may be estimated by simulating the mode overlap of the spot of a lensed fiber with the confined modes of the M6-V5M5V4 waveguide. Figure 3.13 shows the results of this simulation using a commercial mode solver and the scalar approximation to the electric field, demonstrating that the coupling of a 1550 nm lensed fiber with a spot diameter of $5\ \mu\text{m}$ into the confined modes of the M6-V5M5V4 waveguide is not significantly

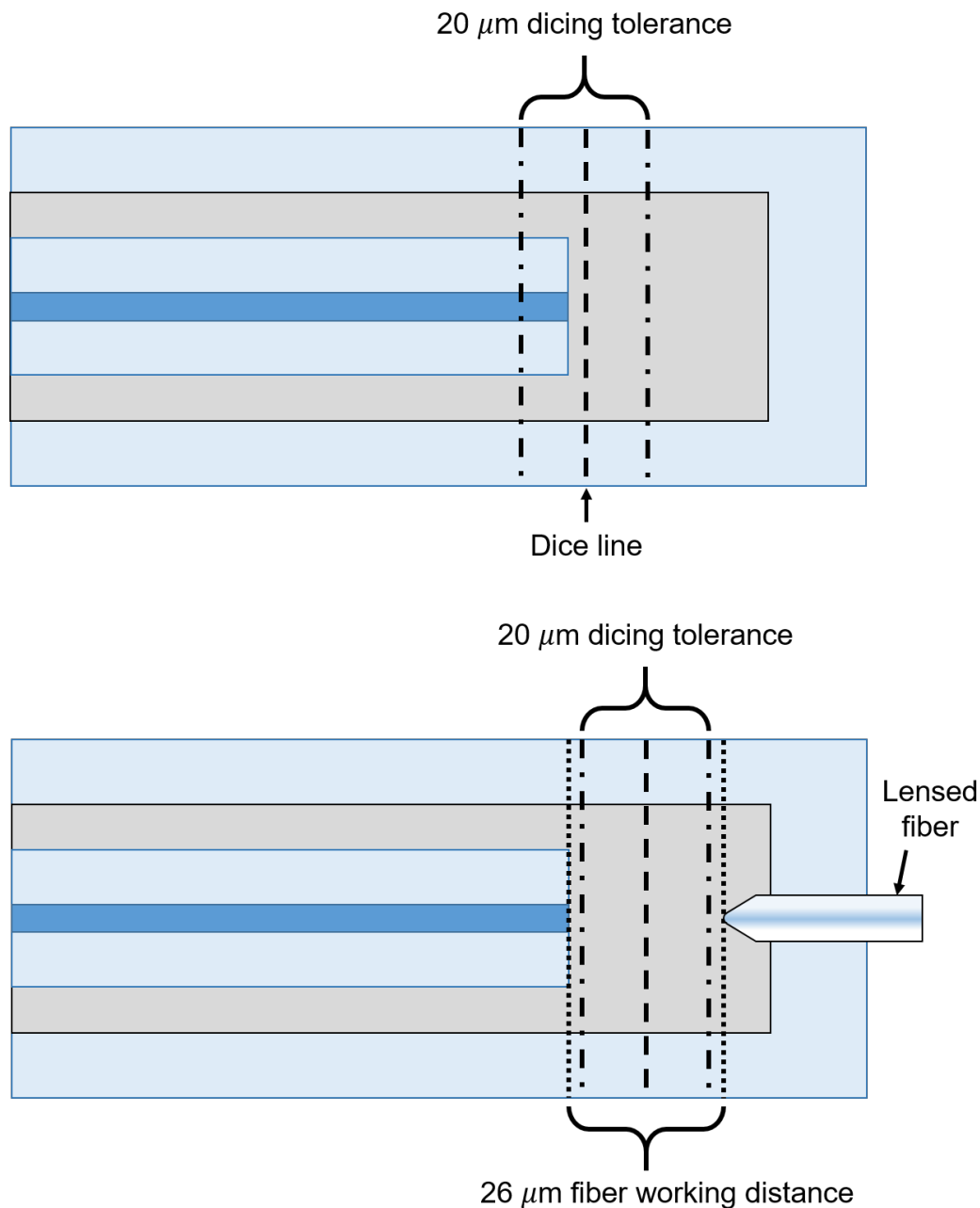


Figure 3.12: Placing the dice line. Top: A dice line that is too close to the waveguide facet. Bottom: Choosing the dice line to fit within a lensed fiber's working distance.

affected by changes in the working distance up to $10\ \mu\text{m}$. Coupling of 90% into the fundamental TE and TM modes can be achieved, dropping by only a little over 10% for up to $10\ \mu\text{m}$ away, which corroborates the fiber coupling advantage assumed at the outset.

Another concern that needs to be taken into account is the ease of measuring the test

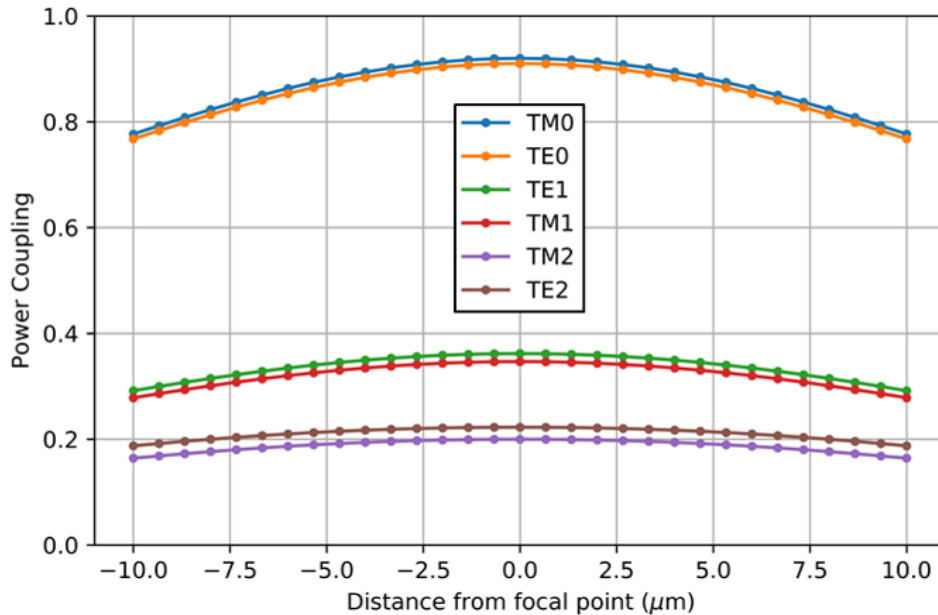


Figure 3.13: Simulated fiber coupling efficiency for all confined modes of a M6-V5M5V4 waveguide and a 1550 lensed fiber with a $5 \mu\text{m}$ spot diameter. The ridge width of the waveguide is $5 \mu\text{m}$.

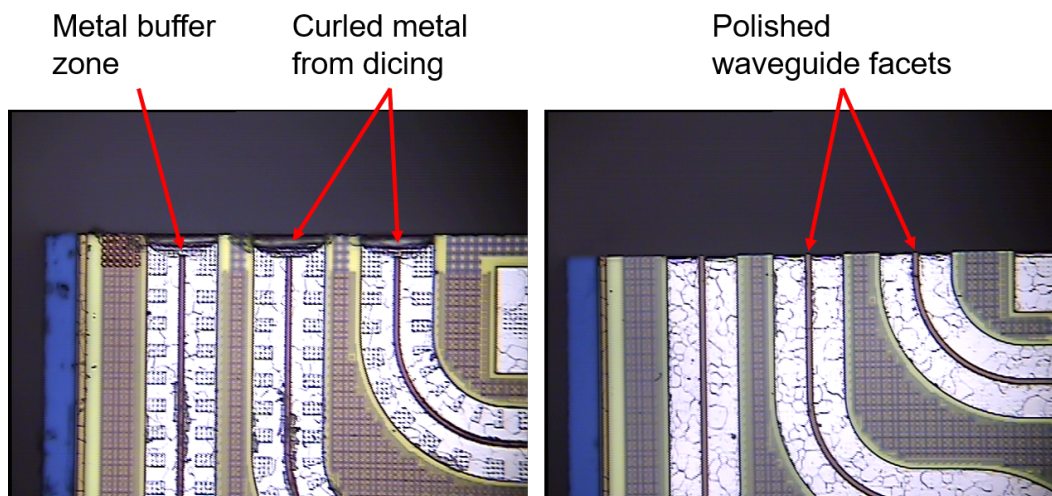


Figure 3.14: Comparison of dicing and polishing. Left: Waveguides before etching but after dicing. Right: Waveguides before etching but after dicing and polishing.

structures in an experimental setup. A setup that makes it easy to measure the test structures will enable more measurements than otherwise might be possible, and will likely make each measurement more accurate as well. For this thesis, an instrument was available with motorized stages for aligning fiber, but this instrument cannot

measure edge couplers at 90° to each other (Fig. 3.15, top), which is a structure that might otherwise be useful for measuring bend loss. Since this instrument yields much more accurate measurements than manual stages mounted on an optical table, it is well worth it to design test structures requiring the fibers to lie in the same axis (Fig. 3.15, bottom).

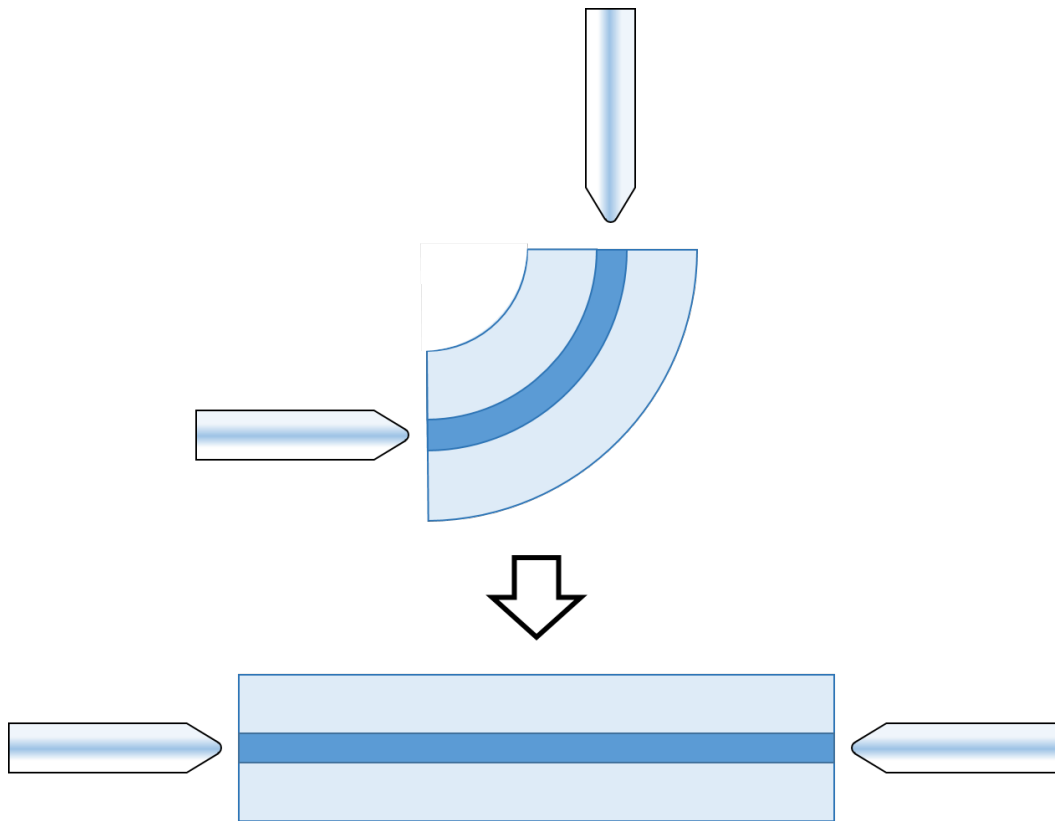


Figure 3.15: Designing a test structure for ease of measurement. Top: A test structure with edge couplers at 90° to each other. Bottom: A test structure with measurement fibers lying in the same axis.

However, if the measurement fibers are lying in-line for characterizing a straight waveguide test structure (Fig. 3.16, top), it may be difficult to distinguish light that is coupled into the waveguide from light propagating outside the waveguide. To eliminate this source of error, test structures with an S-shape were used (Fig. 3.16, bottom). The S-shape further made it easy to add path length differences in the direction parallel to the dice lines, so that the waveguide facets could all be aligned with the same dice line.

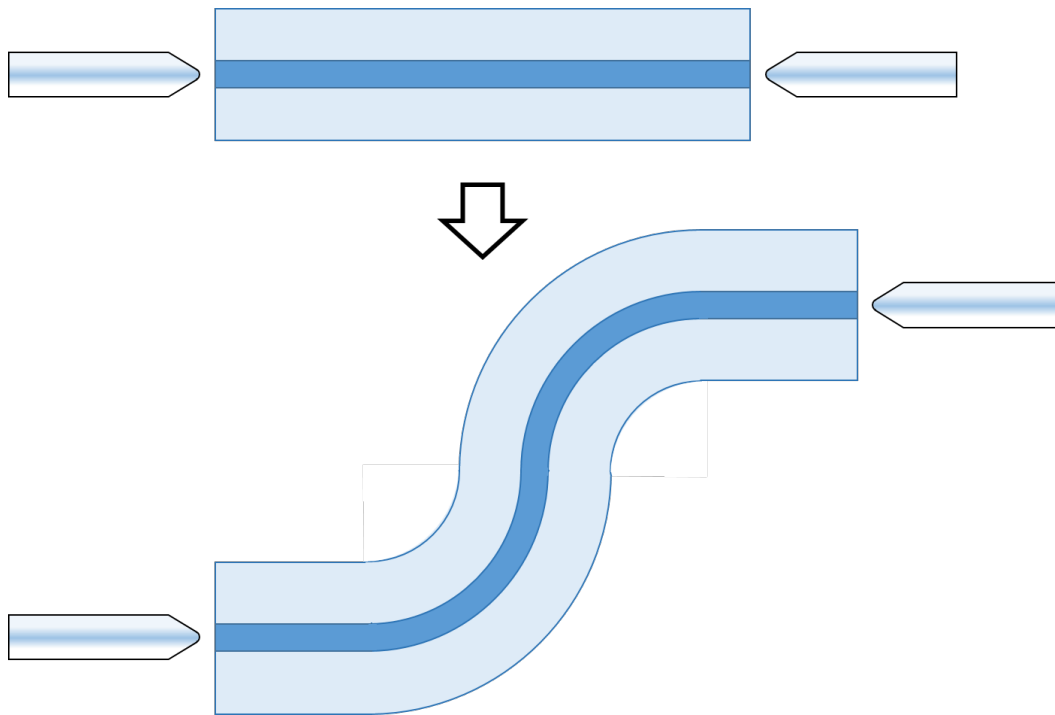


Figure 3.16: Designing a test structure to ensure waveguide coupling. Top: A straight test structure with in-line measurement fibers. Bottom: A S-shaped test structure for ensuring only confined light is measured.

3.2 Etching Procedure

Following the success of etching with Aluminum Etch Type A (AETA) and $\text{H}_2\text{O}_2/\text{EDTA}$ in TO0, the same etchants were used all following tapeouts. Although the specific etching times vary from structure to structure, the general etching recipe consists of first etching the chips in 80°C AETA for several hours in order to remove all the visible metal. Any remaining barrier layers are then removed by etching in a 1:1 solution of 80°C $\text{H}_2\text{O}_2/\text{EDTA}$ for several hours. For many structures this etching procedure is sufficient, but difficult to reach metal or stubborn barrier layers may require additional time with either etchant.

3.3 TO1

The first chip taped out in the 180 nm process has a layout shown in Figure 3.17. There are a variety of structures for testing mechanical strength, etching speed, waveguide loss, and bend loss. The two main waveguide geometries used are M6-V5M5V4 and V5M5V4-M4V3, with a $5\ \mu\text{m}$ ridge width, $10\ \mu\text{m}$ support length, and $25\ \mu\text{m}$ support spacing, unless noted otherwise. The test structures include the following:

- Three M6-V5M5V4 waveguides with 50 μm path length differences
- Three M6-V5M5V4 waveguides with 50 μm path length differences and 100 μm support spacing
- Three V5M5V4-M6 waveguides with 50 μm path length differences, for which the last minute addition of waffling caused excessively high losses
- Three V5M5V4-M4V3 waveguides with 50 μm path length differences
- Three M6-V5M5V4 waveguides with bend radii of 50, 75, and 100 μm
- Three V5M5V4-M6 waveguides with bend radii of 50, 75, and 100 μm , for which the last minute addition of waffling caused excessively high losses
- Three V5M5V4-M4V3 waveguides with bend radii of 50, 75, and 100 μm
- Mechanical test structures, with the following specs fabricated for both the M6-V5M5V4 and V5M5V4-M4V3 geometries (dimensions on metal layer only since the via layer is typically 1.2 μm larger)
 - 225 μm support spacing
 - 300 μm support spacing
 - 300 μm support spacing and 50 μm support length
 - Support widths of 5, 10, 20, 30, and 40 μm
 - Support lengths of 0.3, 0.4, 0.5, 1, 2, 5, 15, and 20 μm
 - Slab widths of 40, 50, and 60 μm
 - Support spacing of 13, 16, 60, and 75 μm .

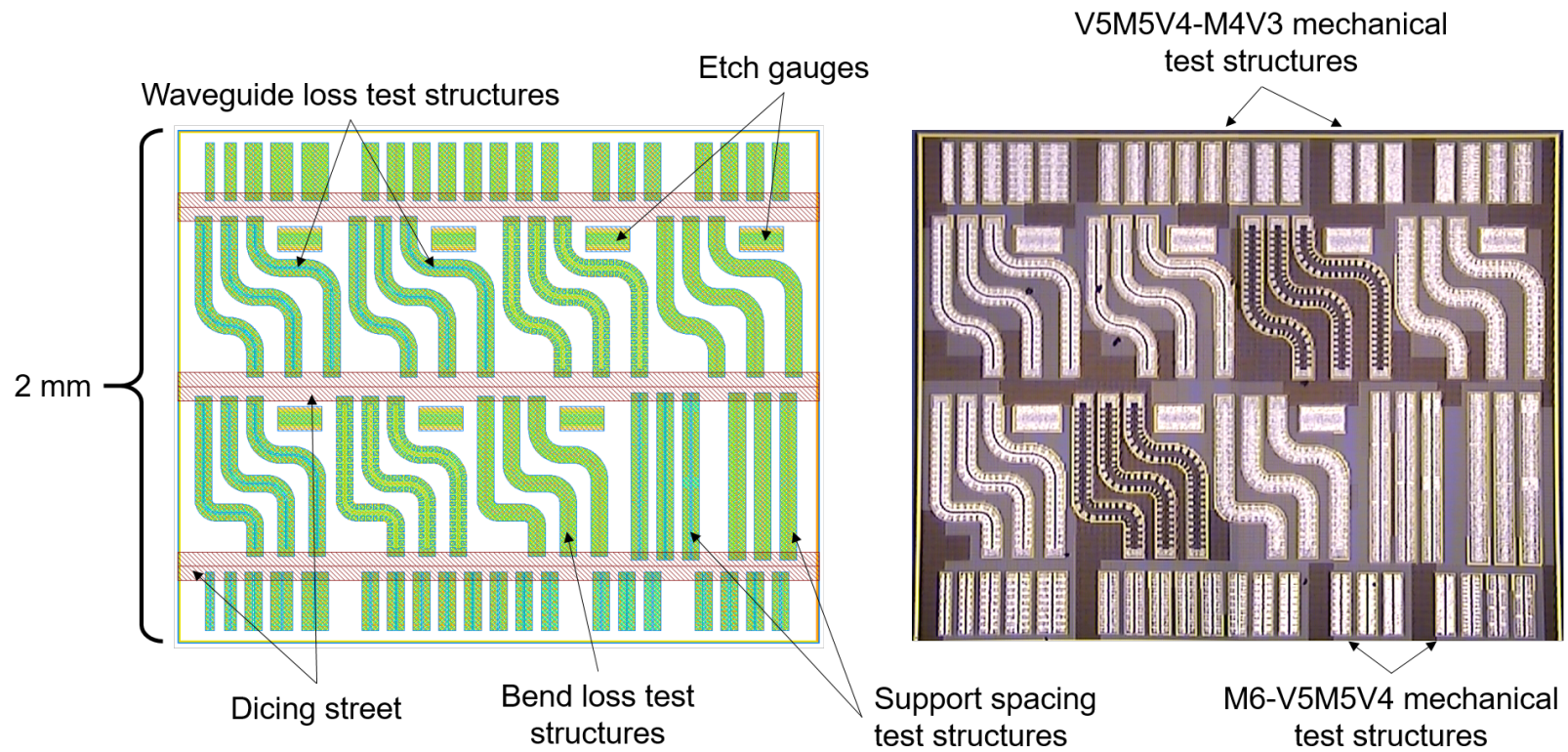


Figure 3.17: Layout of TO1.

M6-V5M5V4

The geometry and design considerations for the M6-V5M5V4 waveguide have already been discussed above in Section 3.1, so in this section we examine the results of post-processing.

Figure 3.18 shows some of the waveguide test structures in the middle of etching. The interference patterns indicate insufficient etching. The waffling falls out in sheets, most likely due to the adhesion of the barrier layer, so it poses no problem during etching. Figure 3.20 displays the M6-V5M5V4 waveguides after they are fully etched. The left image focuses on the suspended waveguide facet, and the right image focuses on the shadow reflected off the bottom of the cavity, and both demonstrate how cleanly the waveguide has been etched. Figure 3.21 shows SEM images of the fabricated waveguides. Note from the scale bars that the roughness is on the order of 10 nm, and that the roughness occurs only in the horizontal dimension, consistent with the planar lithographic processing.

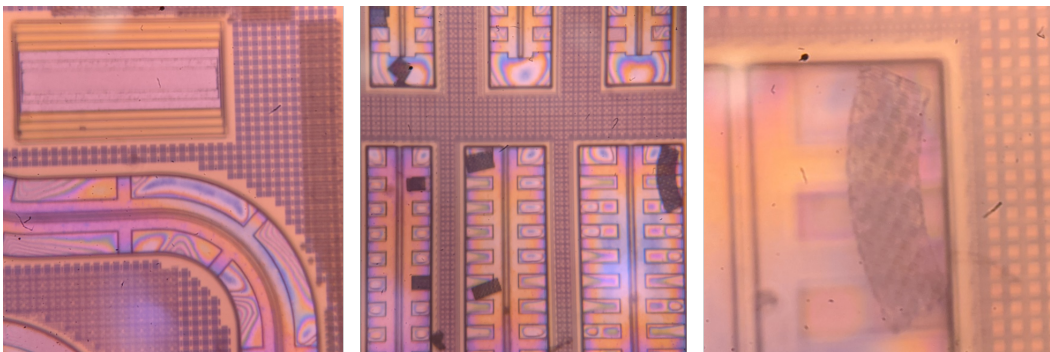


Figure 3.18: A TO1 chip in mid-etch. Left: A partially etched M6-V5M5V4 waveguide. Middle: Partially etched mechanical test structures. Right: Displaced waffling.

The left image of Figure 3.18 shows a partially etched structure above the waveguide. This structure uses terraces of metal layers to provide a visual indicator of etching progress, as shown in Figure 3.19.

The waffling shown in Figure 3.20 can be seen better in the SEM image of Figure 3.22. Note that the barrier layer is holding the pieces of dielectric together, and the via arrays can be easily seen.

For suspended waveguides fabricated using subtractive photonics, there is a trade-off between the mechanical strength of supports and the speed of etching. The shortest supports enable the fastest etching, but the longest supports are the strongest.

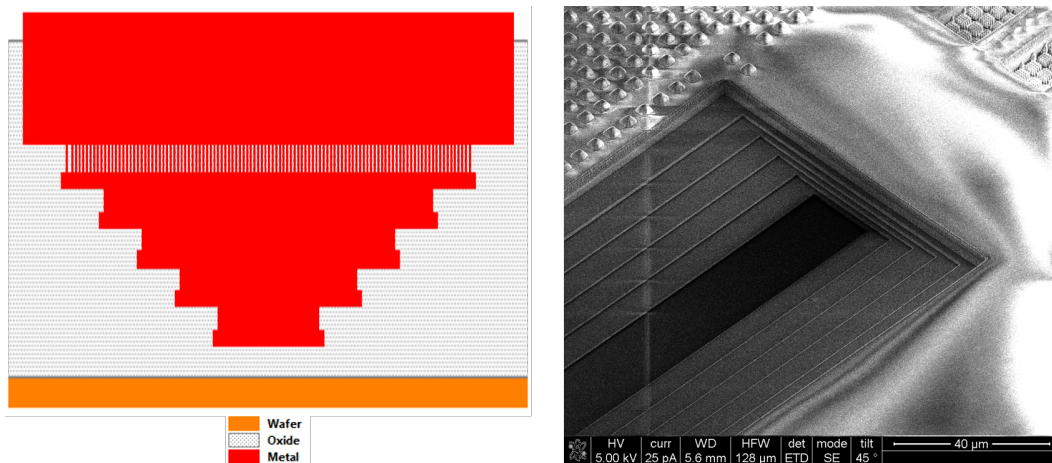


Figure 3.19: SEM image of an etch gauge. SEM image courtesy A. Khachaturian.

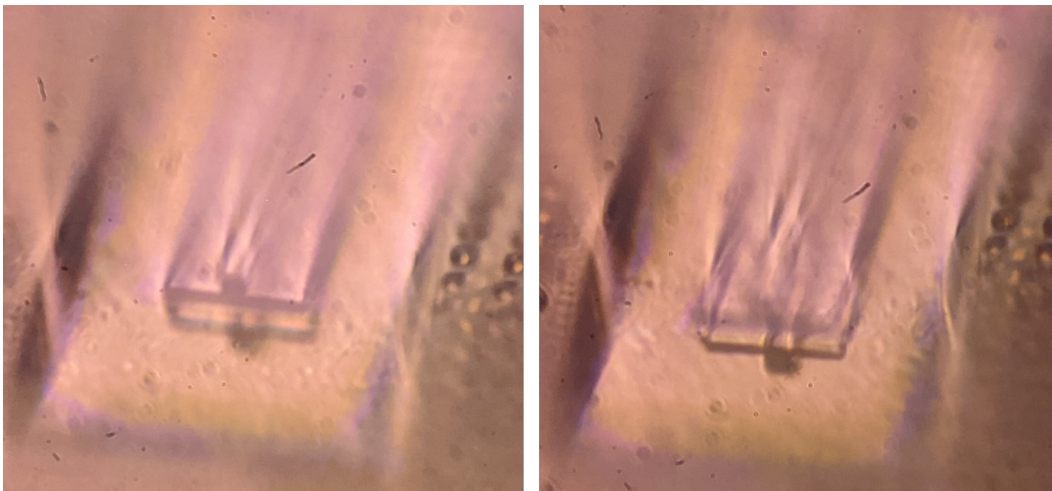


Figure 3.20: A fully etched M6-V5M5V4 waveguide. Left: Focus on the waveguide facet. Right: Focus on the reflection of the waveguide facet.

Figure 3.23 shows a series of test structures designed to take both ends of the trade-off to the extreme. With sufficient time in etchant, even the structure with the longest supports and the smallest etchant openings etched cleanly. Moreover, even the structure with the smallest support length of $1.5 \mu\text{m}$ survived etching without collapse, and there were no broken supports. This test indicates the superb strength of a three layer slab.

The waveguide test structures are designed in accordance with the considerations in Table 3.1. Each test structure consists of three waveguides, each incrementally increasing path length or bend radius. The measurement setup, and a photo of 635 nm light propagating in one of the waveguides, is shown in Figure 3.24. Variations due

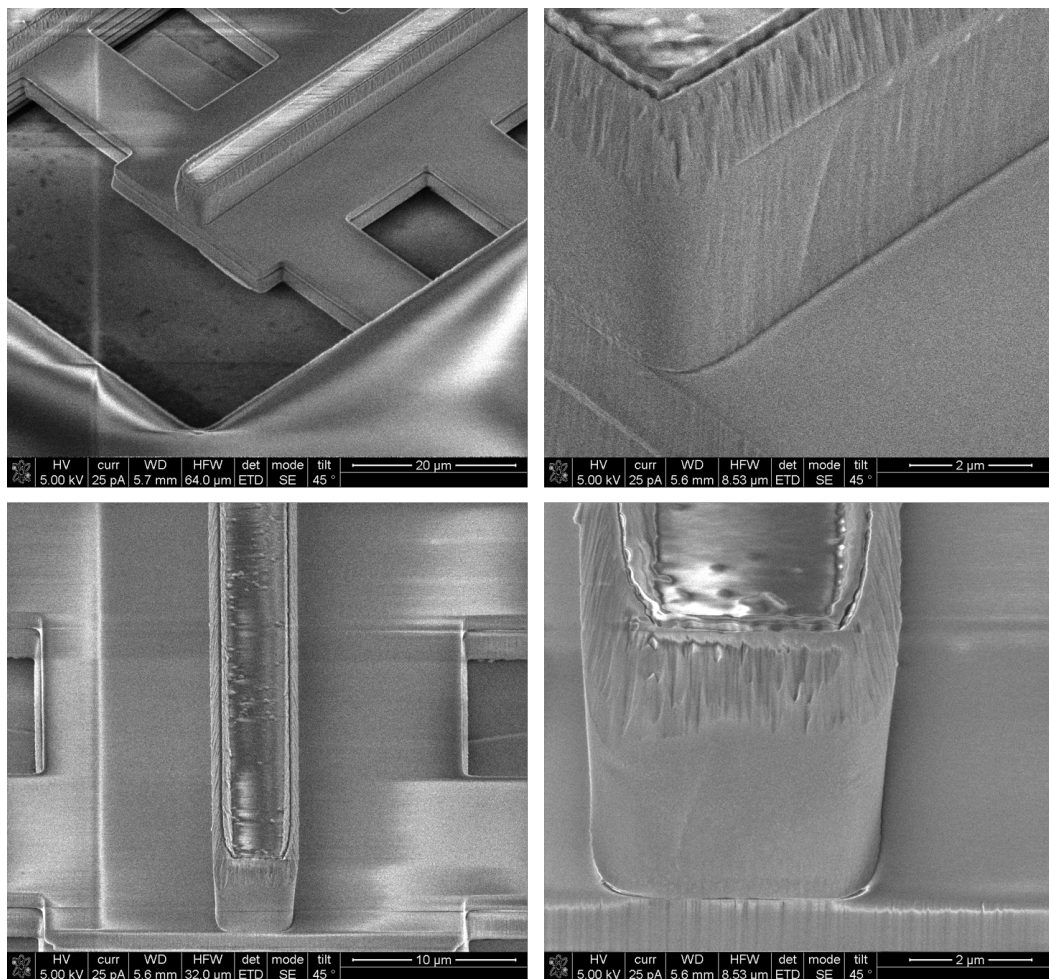


Figure 3.21: SEM images of a fabricated M6-V5M5V4 waveguide. SEM images courtesy A. Khachaturian.

to fiber movement are averaged out over 600 measurements, with the maximum and minimum insertion loss indicated by the error bars in Figure 3.25.

TO1 contains two M6-V5M5V4 waveguide loss test structures. The first uses 10 μm long supports spaced every 25 μm , and the second uses 10 μm long supports spaced every 100 μm . All of the waveguide loss test structures vary by 50 μm in path length, and both a polished and an unpolished chip were measured, with both cleaved and lensed fibers. As Figure 3.25 shows, there is no correlation between path length and insertion loss, indicating that loss is dominated by poor coupling and bend loss. The image of the illuminated waveguide in Fig. 3.24 corroborates this with brighter scattered 635 nm light at the waveguide facets, supports and first bend. Although the chip with polished facets exhibits somewhat decreasing insertion loss with increasing bend radius (Fig. 3.25, right), the trend is not entirely consistent, indicating that

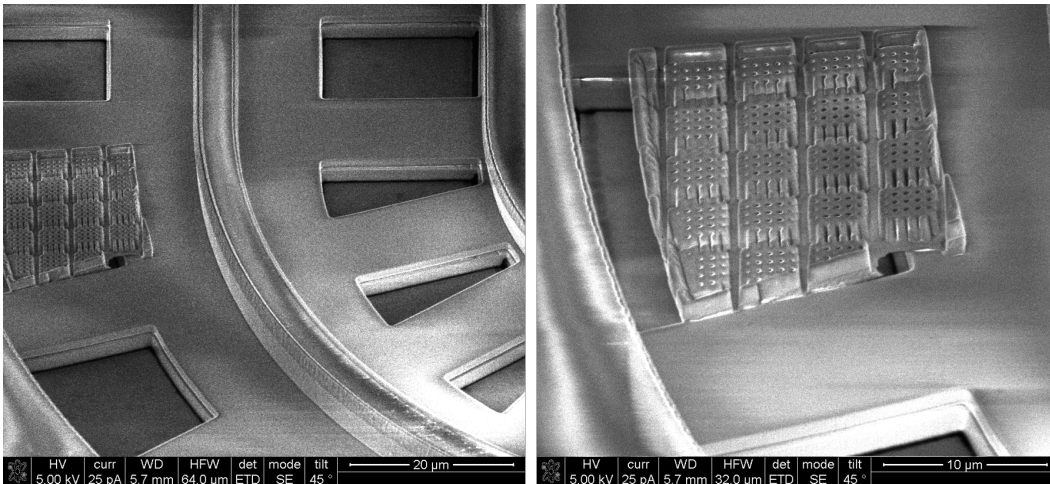


Figure 3.22: SEM images of waveguide waffling. Left: A M6-V5M5V4 waveguide with a partially removed piece of waffling. Right: Close view of the waffling. SEM images courtesy A. Khachaturian.

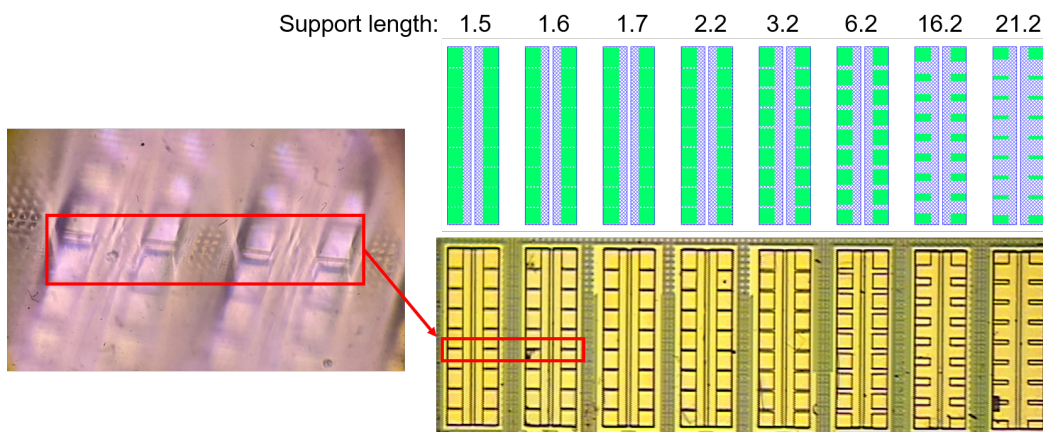


Figure 3.23: Layout and micrograph of the M6-V5M5V4 mechanical test structures.

differences in loss are dominated by variations between the individual waveguides.

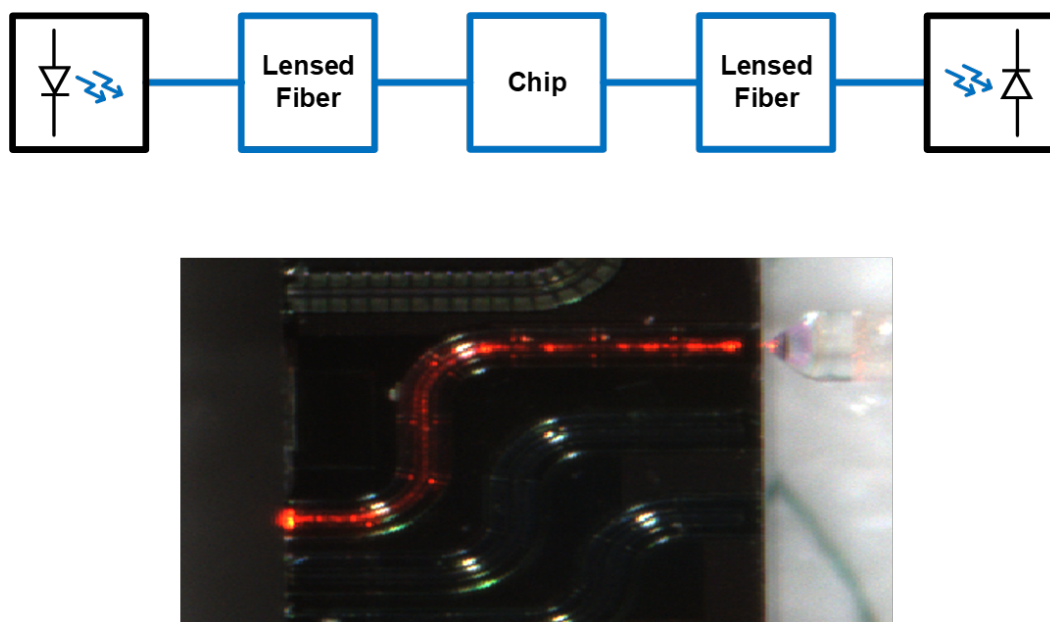


Figure 3.24: Measurement of TO1. Top: Experimental setup used for measuring insertion loss of the waveguides. Bottom: Image of 635 nm light propagating in the waveguides.

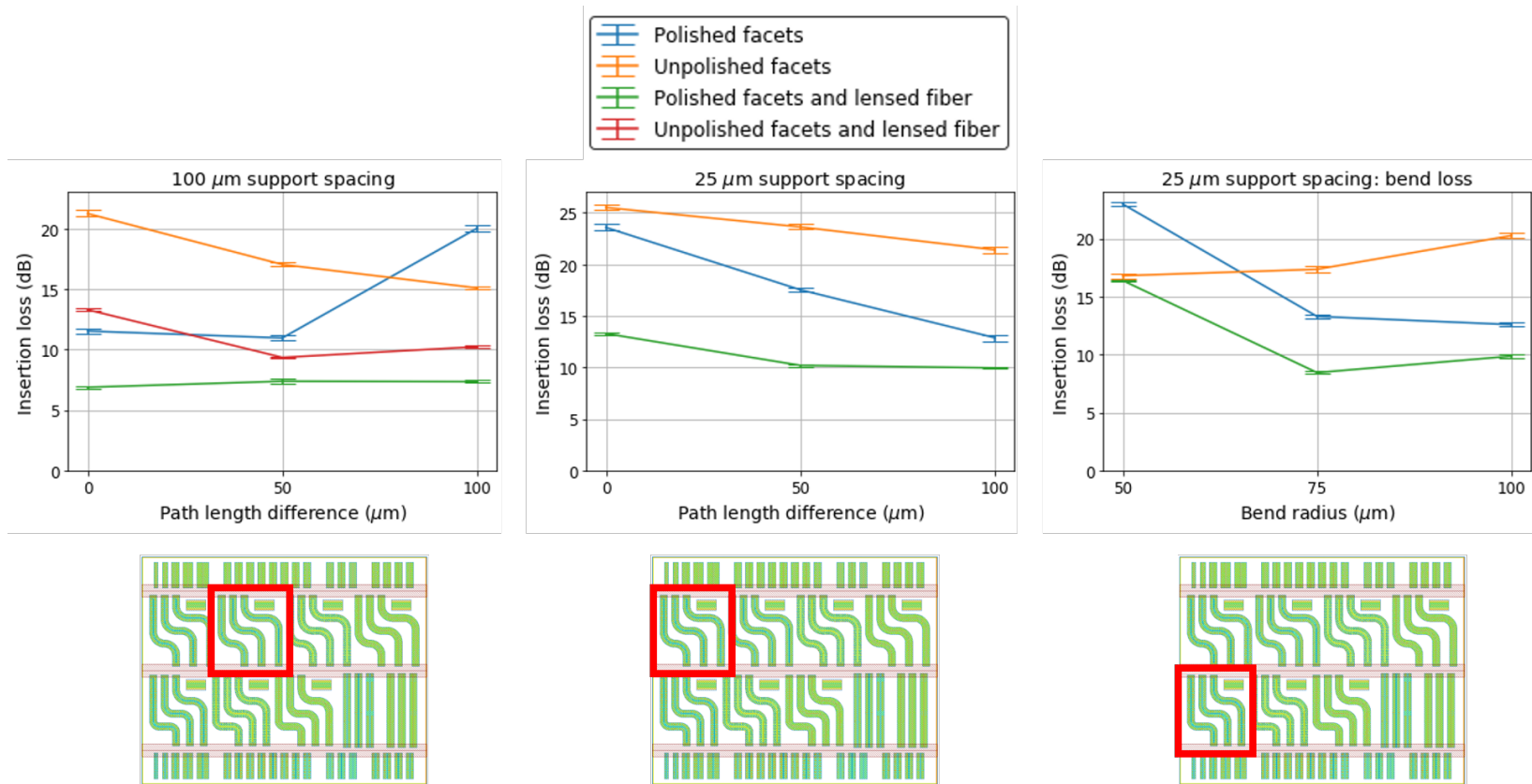


Figure 3.25: Measurements of M6-V5M5V4 waveguides with varying path lengths and bend radii. The bottom diagrams indicate the locations of the test structures in TO1.

V5M5V4-M4V3

The second waveguide geometry fabricated in TO1 uses a three layer ridge and a two layer slab (Figure 3.26). The main goal of this geometry was to determine the mechanical strength of a two layer slab, which would improve optical confinement over a three layer slab. With a $5\ \mu\text{m}$ wide ridge, again for better overlap with the spot of a lensed fiber at $1550\ \text{nm}$, the V5M5V4-M4V3 geometry has six confined modes, shown in Figure 3.27. Relatively good coupling efficiency of about 80% can be obtained for the fundamental TE and TM modes at $1550\ \text{nm}$ (Figure 3.28). A sweep of the ridge width shows that this geometry supports singlemode operation at a ridge width just below $3\ \mu\text{m}$ (Fig. 3.29).

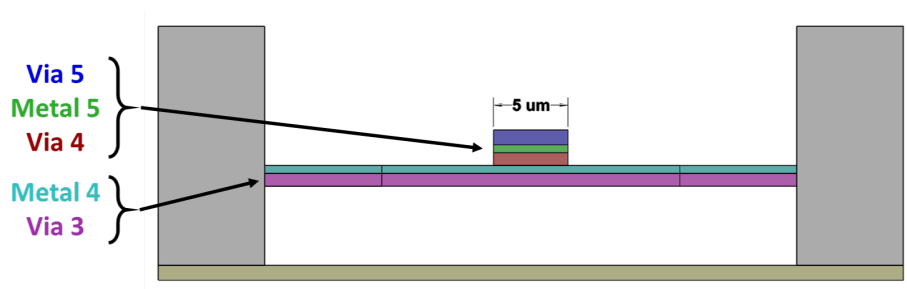


Figure 3.26: Schematic cross section of the V5M5V4-M4V3 waveguide geometry.

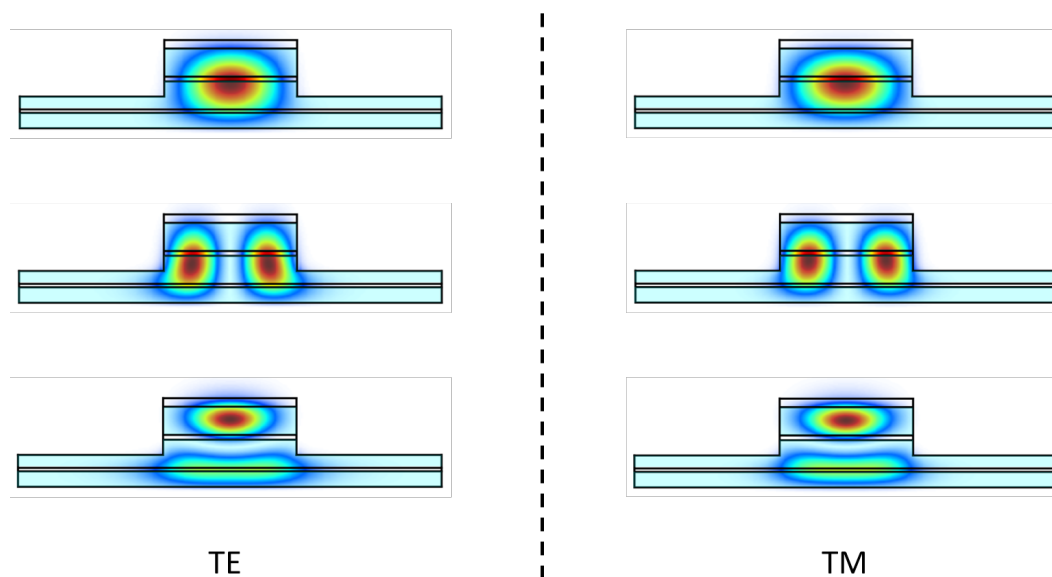


Figure 3.27: Confined modes of the V5M5V4-M4V3 waveguide at a wavelength of $1550\ \text{nm}$ and a ridge width of $5\ \mu\text{m}$.

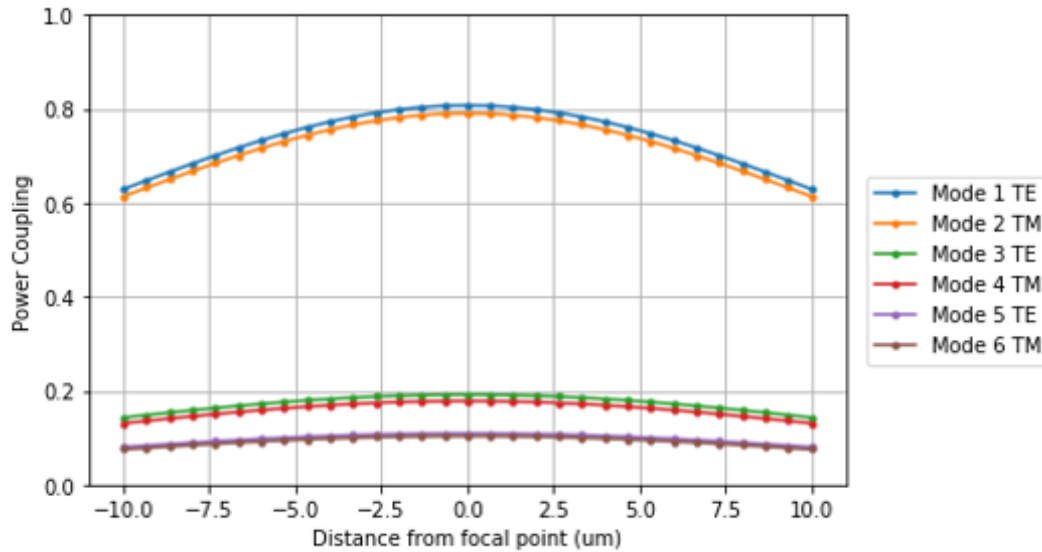


Figure 3.28: Simulated fiber coupling efficiency for all confined modes of a V5M5V4-M4V3 waveguide and a 1550 lensed fiber with a 5 μm spot diameter. The ridge width of the waveguide is 5 μm .

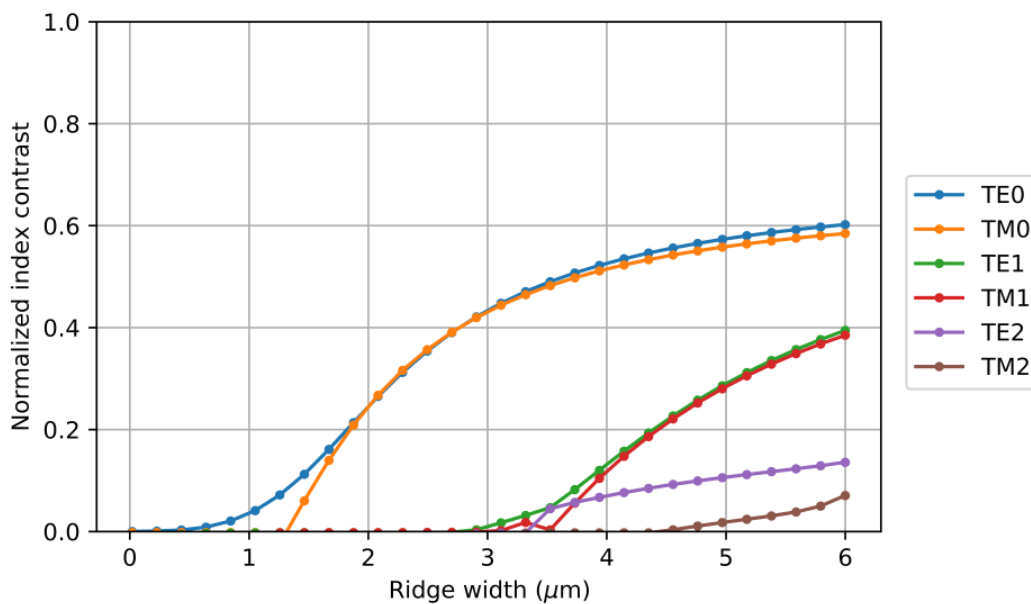


Figure 3.29: Normalized effective index of the confined modes of the V5M5V4-M4V3 waveguide across ridge widths at a wavelength of 1550 nm.

SEM images of the fabricated waveguides can be seen in Figure 3.30. Similar to the M6-V5M5V4 waveguides, roughness can be seen to be on the order of 10 nm, and occurs only in the horizontal dimension. The ridge layers of this waveguide

do not perfectly align across the entire length of the waveguide, a result of the low tolerances accepted for BEOL interconnects.

A thin layer can also be seen in the upper left SEM image of Figure 3.30. Additional views of this thin layer can be seen in the microscope and SEM images of Figure 3.31. It is likely that this thin layer is a TiN barrier layer, consistent with the observations of TO0. Note the transparency of the barrier layer in the microscope images, which will be contrasted with barrier layer observations in later chips. Note also in the SEM images that the waveguides appear to be partially collapsed, indicating excessive support width.

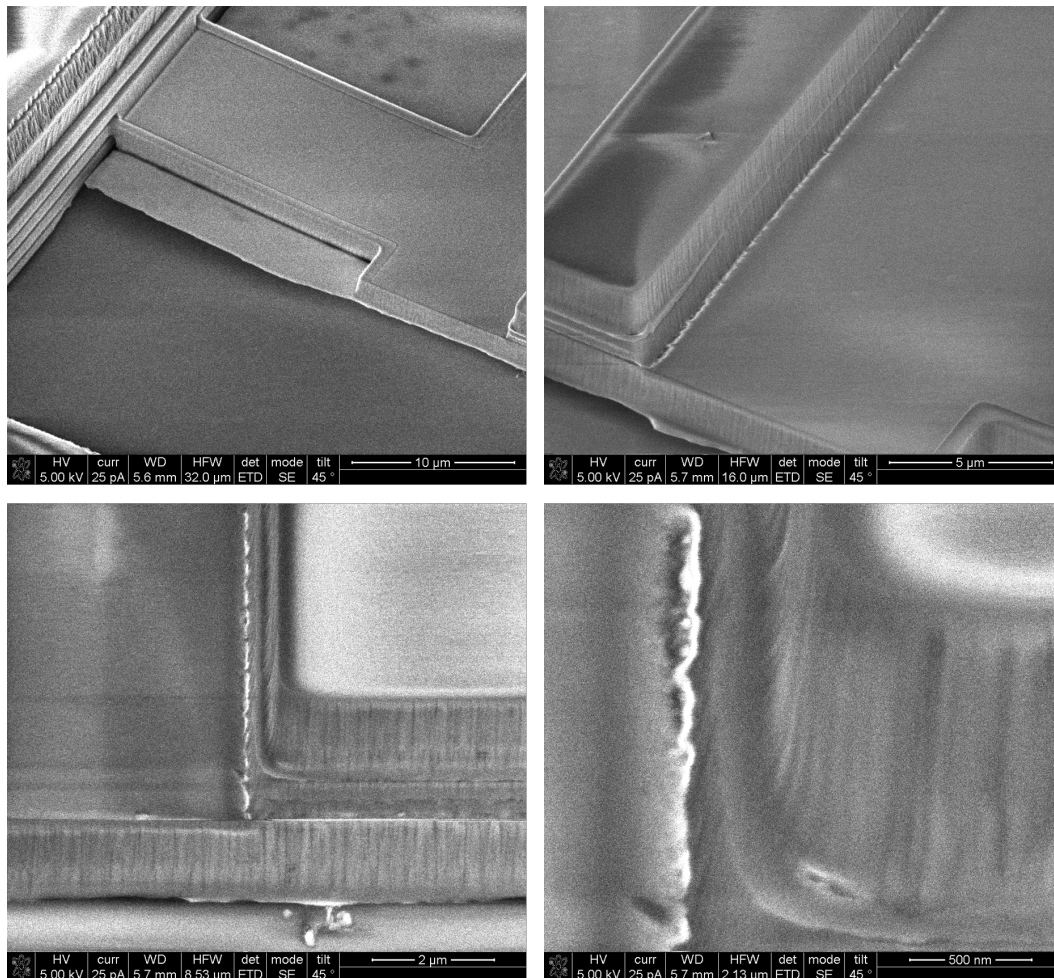


Figure 3.30: SEM images of a fabricated V5M5V4-M4V3 waveguide. SEM images courtesy A. Khachaturian.

Waveguide collapse due to excessive support width can be seen most clearly in Figure 3.32, where the support width was varied from $5\ \mu\text{m}$ to $40\ \mu\text{m}$. The microscope image of Fig. 3.32 demonstrates the insufficiency of a $40\ \mu\text{m}$ support width. For the

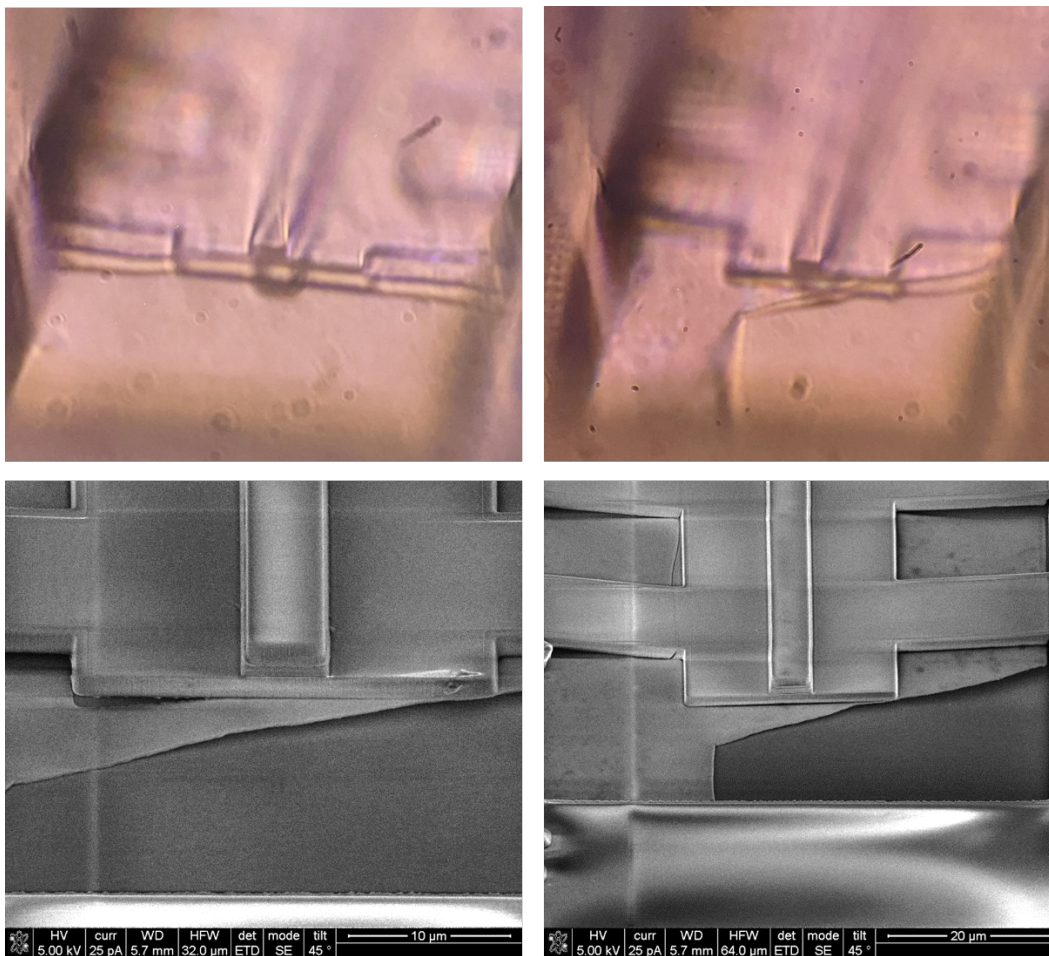


Figure 3.31: Microscope and SEM images of V5M5V4-M4V3 waveguides. SEM images courtesy A. Khachaturian.

mechanical test structures of Fig. 3.32, the support widths shorter than $40\ \mu\text{m}$ were mechanically sound, but since those test structures are also supported at both ends by connection to the surrounding dielectric, they do not fully reflect the conditions of a long waveguide. The waveguide of Fig. 3.31 demonstrates that for a long waveguide even a $20\ \mu\text{m}$ support width is insufficient for mechanical strength. At the low end, the small etchant openings of a $5\ \mu\text{m}$ support width increased etching times.

Similar mechanical test structures were fabricated for testing support length, as shown in Figure 3.33. In contrast to the M6-V5M5V4 waveguides, for which all support lengths supported the waveguides, for the V5M5V4-M4V3 geometry only support lengths of $3.2\ \mu\text{m}$ and above survived the etching process. The rightmost structures of Fig. 3.33 were intended to test the excess etching time required for large support lengths. Although the etching time depends on the etchant used and the

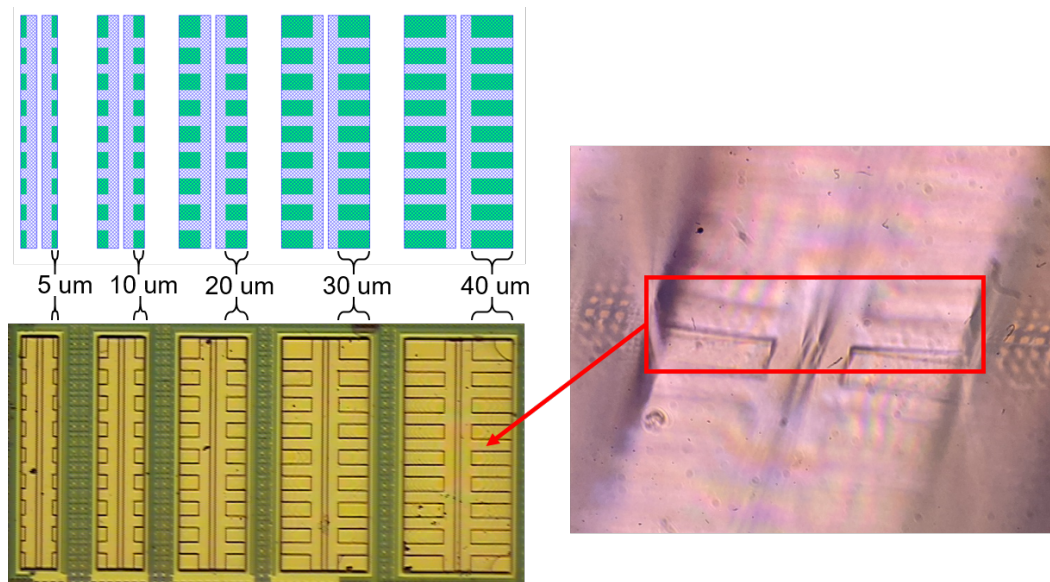


Figure 3.32: Microscope images of structures designed to test the limits of support width for the V5M5V4-M4V3 geometry.

temperature of the etch, the general conclusion is that the most reasonable etching times occur for support lengths no more than half the support period.

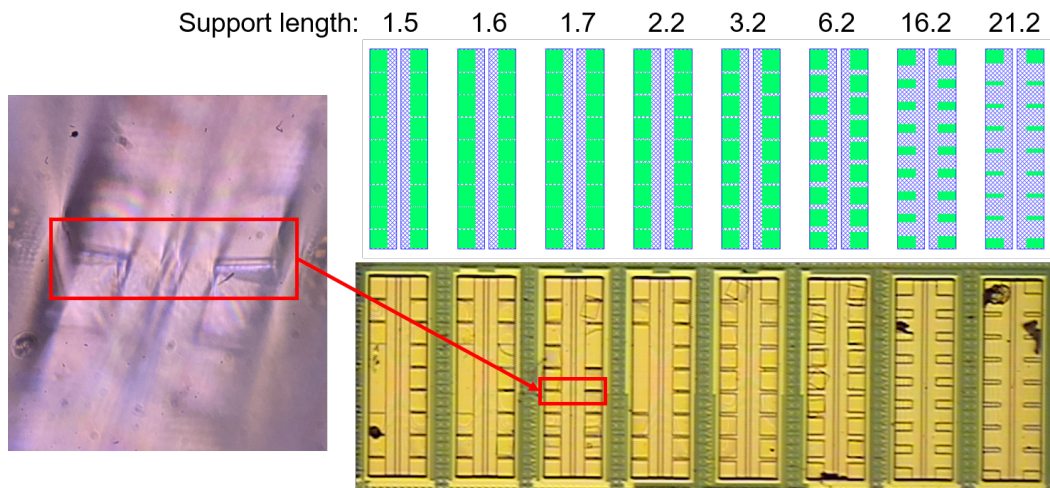


Figure 3.33: Microscope images of structures designed to test the limits of support length for the V5M5V4-M4V3 geometry.

3.4 TO2

The second chip taped out in the 180 nm process has a layout shown in Fig. 3.34. Since the results of TO1 indicated that the wavelength path length differences were not sufficient to determine waveguide loss, TO2 was intended to extend those path length

differences while still keeping chip area to the MPW minimum for risk reduction. The two main waveguide geometries used are M6-V5M5V4 and V5M5-V4M4V3. The former waveguides use a 5 μm ridge width, 10 μm support length, and 100 μm support spacing, unless noted otherwise. The latter waveguides use a 4 μm ridge width, 10 μm support length, and 100 μm support spacing, unless noted otherwise.

The chip featured the following waveguides:

- Four M6-V5M5V4 waveguides with 500 μm path length differences
- Three V5M5-V4M4V3 waveguides with 300 μm path length differences
- Three M6-V5M5V4 waveguides with 300 μm path length differences, 50 μm support length, and 300 μm support spacing
- One V5M5-V4M4V3 waveguide with 50 μm support length and 300 μm support spacing
- Three M6-V5M5V4 mechanical test structures, one each with 10 μm , 50 μm , and 100 μm support lengths, and all with increasing support spacing of 500 μm , 750 μm , and 1 mm
- One V5M5V4-M4V3 mechanical test structure with 5 μm ridge width, 10 μm support length, 300 μm support spacing, and vertical supports.

TO2 also included several CMOS photodiodes.

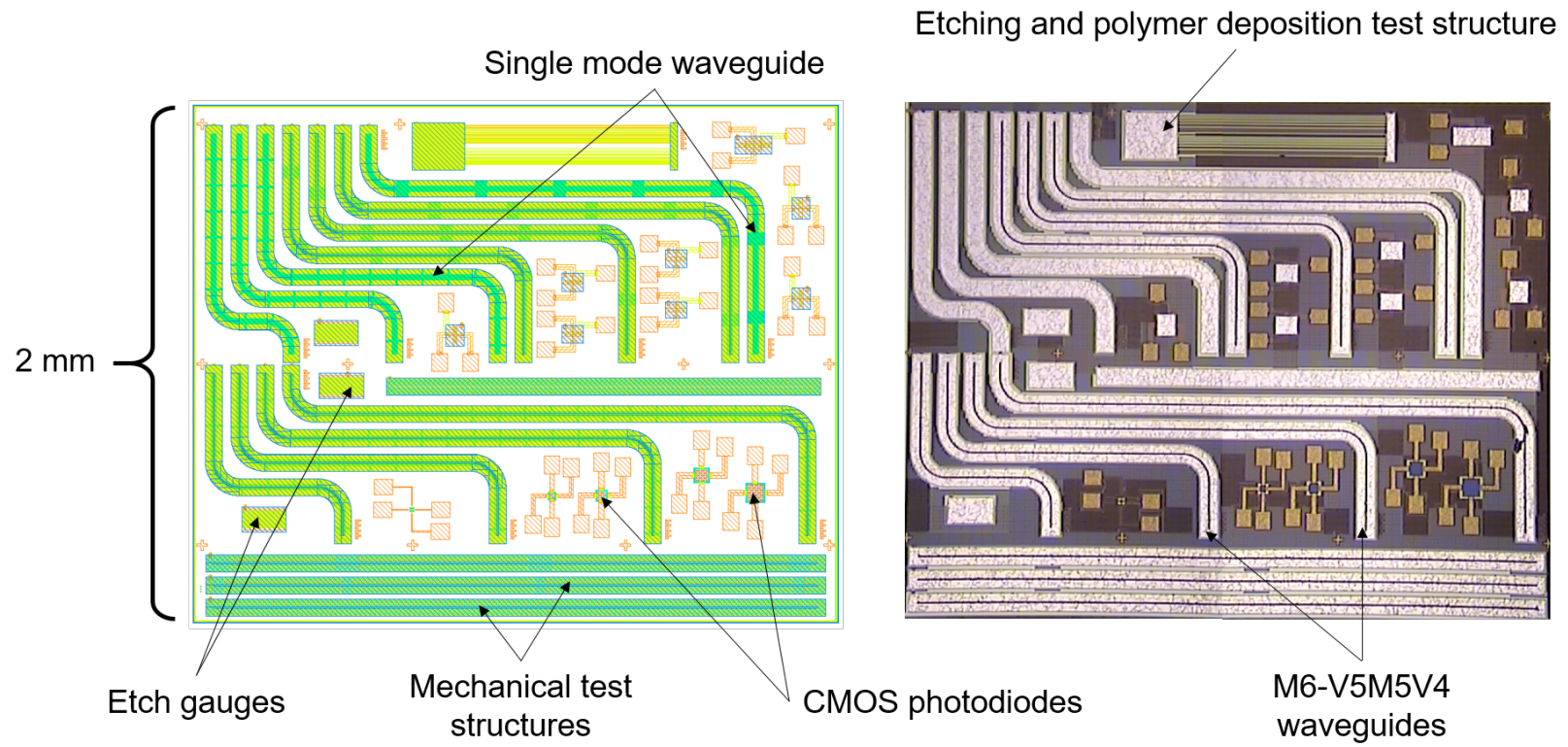


Figure 3.34: Layout of TO2.

The waveguides in this tapeout were severely compromised by a foundry requirement for waffling on all metal layers, which is illustrated in Figure 3.35. With the waffling on only the V5 layer, as in TO1, each section of waffling remained in one piece and all the waffling was cleanly removed from the chip. In TO2 significantly different results were observed. Figure 3.36 shows that the waffling in this case broke into individual waffles and was scattered all over the waveguide cavity, especially on the bottom. A variety of methods were used to try and prevent this accumulation of waffling, including etching the chips upside down, but nothing worked. As a result, many waveguides suffered from high insertion loss.

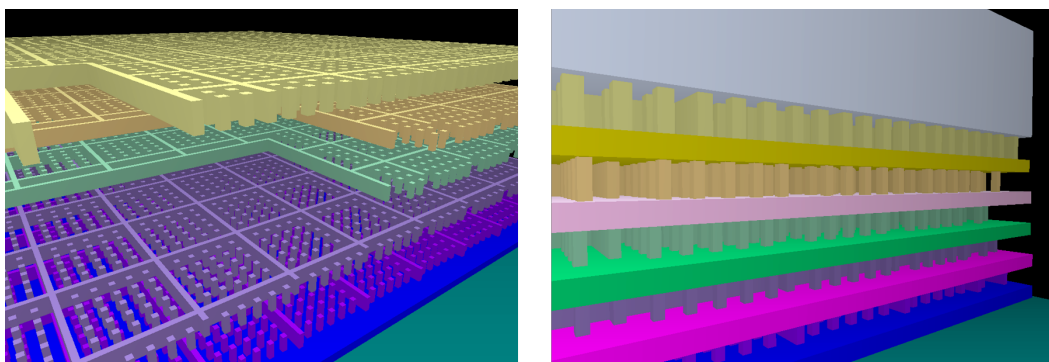


Figure 3.35: The design of waffling on all via layers. Left: Exploded view of the via layers used in waveguide construction. Right: Depiction of the metal layers of a waveguide cavity, with waffling on all via layers.

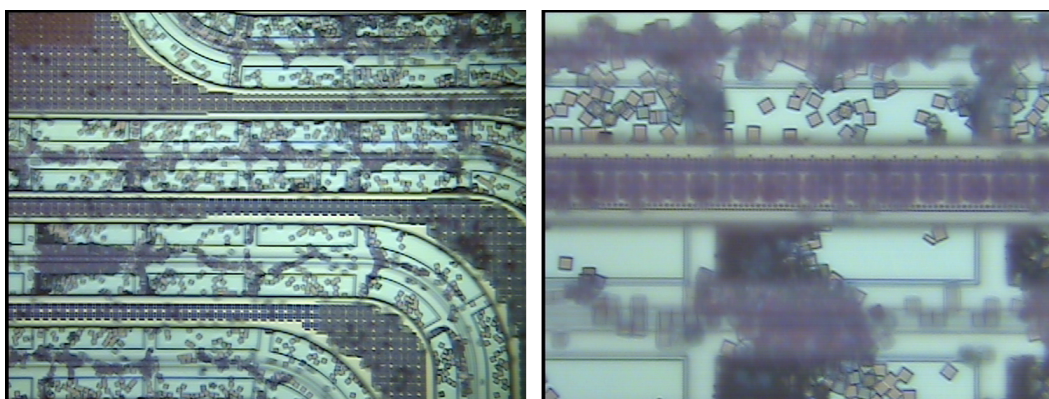


Figure 3.36: The residue of waffling on all layers.

M6-V5M5V4

Although TO1 had shown the feasibility of using a two layer slab to improve optical confinement, TO2 used M6-V5M5V4 waveguides again in order to provide continuity.

The same 100 μm bend radius was also used so that measurements of the TO1 waveguides could be combined with measurements of the TO2 waveguides.

Measurements results for three chips are shown in Figure 3.37. The insertion loss is shown with respect to the path length of each of the four M6-V5M5V4 waveguides. The longest waveguide of chip 1 exhibits exceptionally high insertion loss due to waveguide damage (lower left image). The small and non-monotonic differences between the three shorter waveguides of chip 1 indicate that insertion loss is dominated by coupling losses, bend losses, and scattering from the waffling. The negligible contribution of waveguide loss indicates that longer path length differences are necessary. The measurement results of chip 2 reinforce this conclusion, as the insertion loss does not increase monotonically with path length difference.

The measurements of chip 3 provide two different waveguide loss values. The difference in insertion loss between the shortest waveguide and the next shortest waveguide is 2 dB. The half millimeter path length difference indicates that waveguide loss is 40 dB/cm. The differences between the three longest waveguides are 0.5 dB every half millimeter, indicating a waveguide loss of 10 dB/cm. A least squares fit to all four data points splits the difference between these two loss values at 19 dB/cm. The inconsistency of these results, however, together with the inconsistency of chips 1 and 2, indicates that longer path length differences are necessary to overcome coupling and scattering losses.

The insertion losses in Figure 3.37 include losses due to fiber coupling, bends, and general waveguide loss, which may be formulated as

$$IL = L_{coupling} + L_{bends} + L_{waveguide}. \quad (3.1)$$

The general or straight waveguide loss includes loss due to scattering from waffling, scattering from sidewall roughness, reflection from supports, material absorption, etc. Eq. 3.1 implies that if W is taken to be the waveguide length, then

$$\frac{IL}{W} \geq \frac{L_{waveguide}}{W}, \quad (3.2)$$

provides an upper bound on the general waveguide loss. Figure 3.38 displays this upper bound for the four M6-V5M5V4 waveguides, for all three measured chips, taking into account the reduction in waveguide length due to polishing. The third waveguide of chip 1 provides the lowest upper bound of 38.8 dB/cm, which is comparable to the previously reported best loss in unmodified bulk CMOS of 38 dB/cm [55].

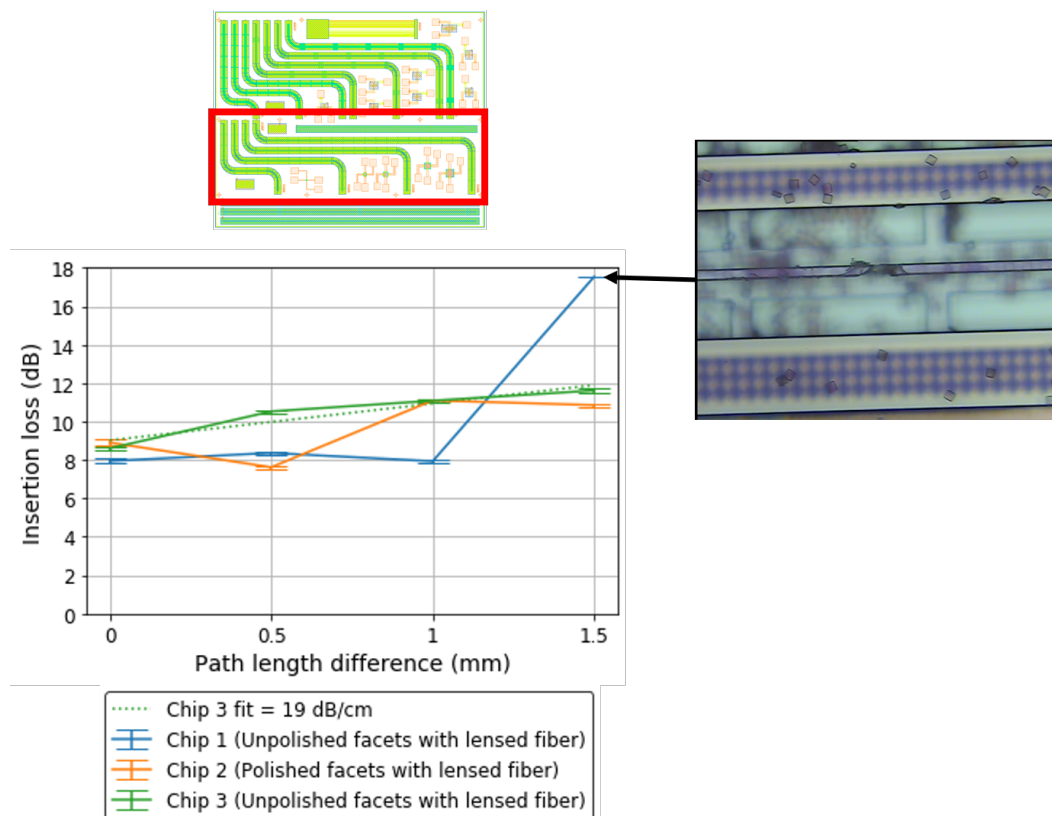


Figure 3.37: Measuring M6-V5M5V4 waveguides in TO2. Left top: Location of the tested waveguides in TO2. Left bottom: Measurements of insertion loss. Right: Microscope image of the ridge damage causing the excessive insertion loss of the longest waveguide of chip 1.

V5M5-V4M4V3

TO2 introduced a singlemode waveguide for the first time, using the V5M5-V4M4V3 geometry, as shown in Figure 3.39. As mentioned in Section 2.4, the rib geometry makes large area singlemode waveguides possible through a reduction in index contrast between the ridge and slab regions. This is illustrated clearly in Figure 3.40, where the index contrast of the V5M5-V4M4V3 geometry at 780 nm can be seen to be much lower than the index contrast of the M6-V5M5V4 (Fig. 3.5) or V5M5V4-M4V3 (Fig. 3.29) geometries at 1550 nm. A ridge width of 4 μm was used in TO2 in order to maximize confinement while maintaining singlemode operation (Fig. 3.41).

The fiber coupling efficiency is plotted in Figure 3.42 for a 2 μm diameter spot commensurate with a 780 nm lensed fiber, with optimal coupling of about 70% for the confined modes. The loss per centimeter due to reflections from supports is plotted in Figure 3.43, assuming that the waveguide supports have a period of

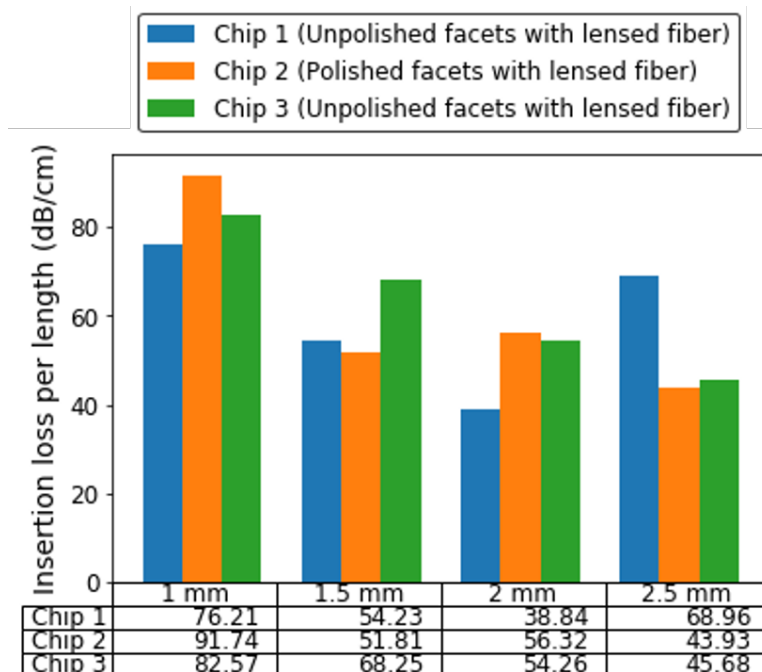


Figure 3.38: Upper bounds on waveguide loss derived from the insertion loss of each waveguide.

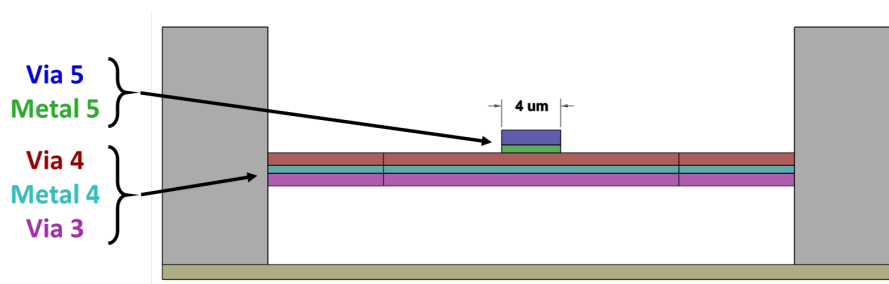


Figure 3.39: Schematic cross section of the V5M5-V4M4V3 waveguide geometry used in TO2.

100 μm . The slab extension used in TO2 is about 10 μm to match the waveguides fabricated in TO1, although the extension could be dropped to 5 μm with almost no penalty.

As mentioned in section 2.4, the disadvantage of low index contrast is high bend loss, although this loss can be mitigated using Euler bends and single-sided slabs. A comparison between a standard V5M5-V4M4V3 bend and these techniques is shown in Figure 3.44. As expected from the discussion in section 2.4, the standard bend suffers higher loss than both single-sided slab bends, and the Euler bend suffers the highest loss. The Euler single-sided slab bend provides the lowest loss with the

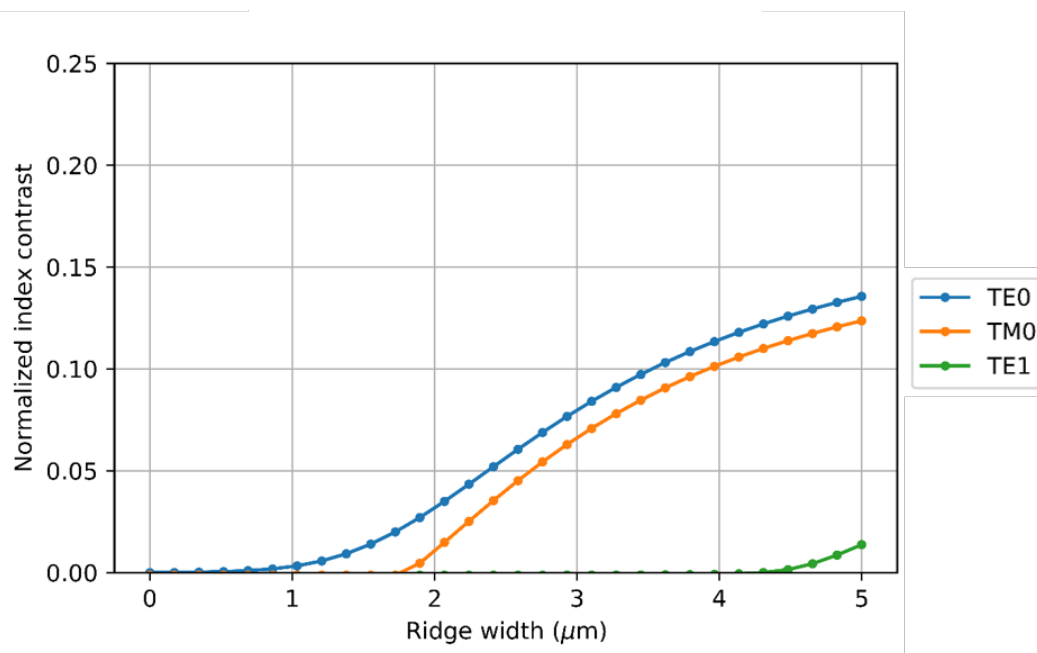


Figure 3.40: Normalized effective index of the confined modes of the V5M5-V4M4V3 waveguide across ridge widths at a wavelength of 780 nm.

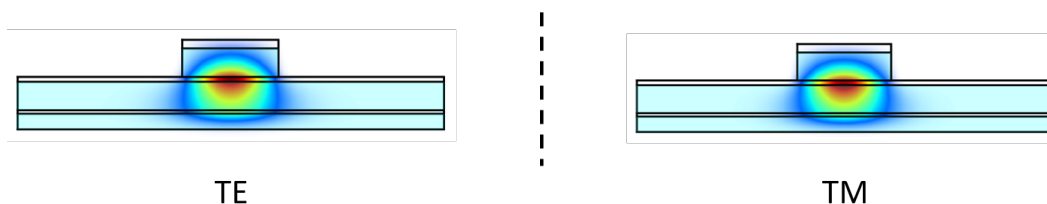


Figure 3.41: Confined modes of the V5M5-V4M4V3 waveguide at a wavelength of 780 nm and a ridge width of 4 μm .

most compact footprint, but the V5M5-V4M4V3 test structures in TO2 do not use these bends for simplicity, since the waveguide loss requires measurement of only the difference in insertion loss between waveguides.

Measurements of the V5M5-V4M4V3 waveguides were not entirely consistent, as shown in Figure 3.45, which indicates the need for greater path length differences. The M6-V5M5V4 waveguides next to them, with 300 μm support spacing, do exhibit increasing insertion loss with path length difference, but least squares fits indicate losses of 58-59 dB/cm. Moreover, the inconsistency of other waveguides on these chips indicates that these positive trends may be a coincidence, necessitating a new design with longer path length differences. The cavity designed as an etching and polymer deposition test structure was plagued by the waffling on all layers, which

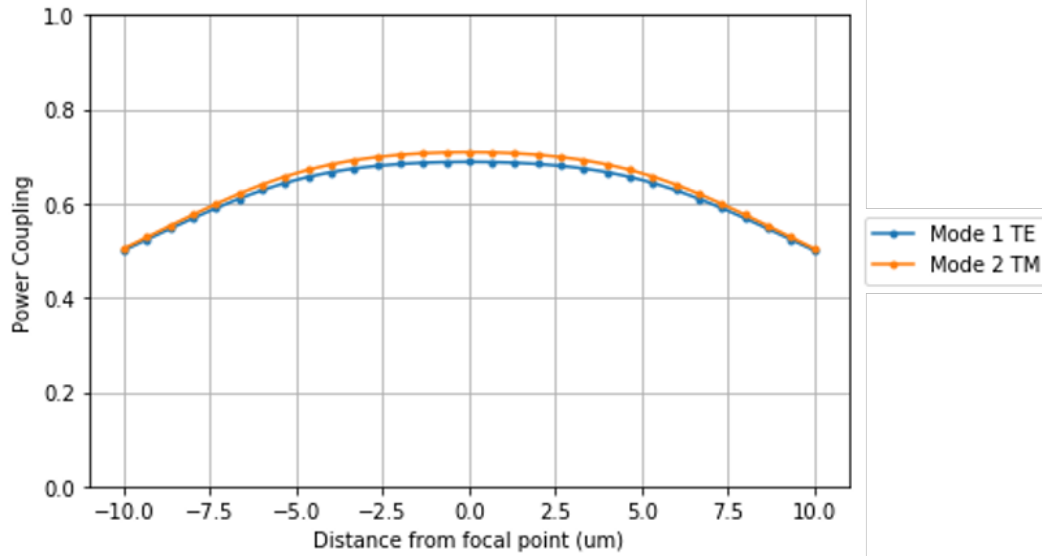


Figure 3.42: Simulated fiber coupling efficiency for all confined modes of a V5M5-V4M4V3 waveguide and a 780 lensed fiber with a 2 μm spot diameter. The ridge width of the waveguide is 4 μm .

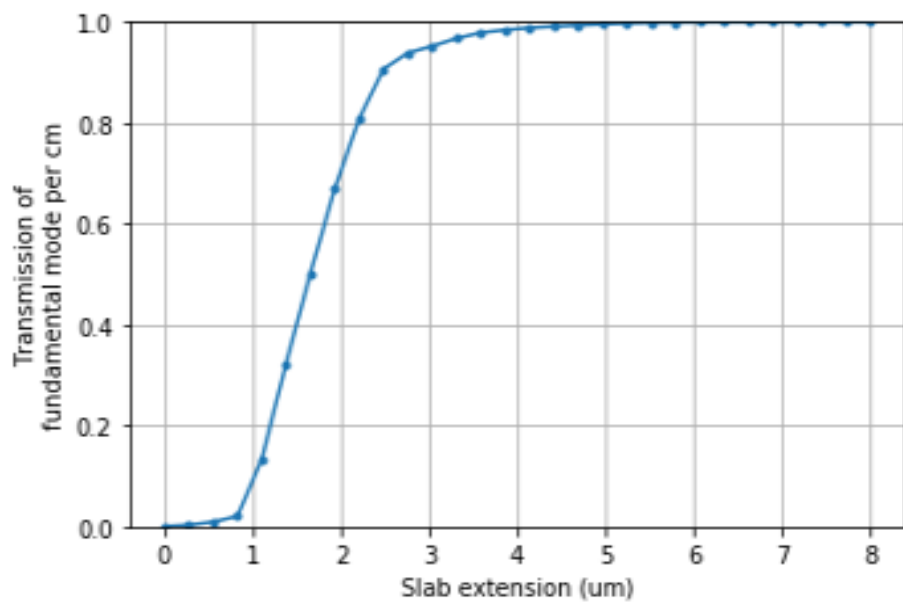


Figure 3.43: Transmission of the fundamental mode of a V5M5-V4M4V3 waveguide at 780 nm, assuming a support period of 100 μm , based on the extension of the waveguide slab beyond the 4 μm ridge width.

was difficult to remove from the cavity and attached channels, and prevented any testing (Fig. 3.46).

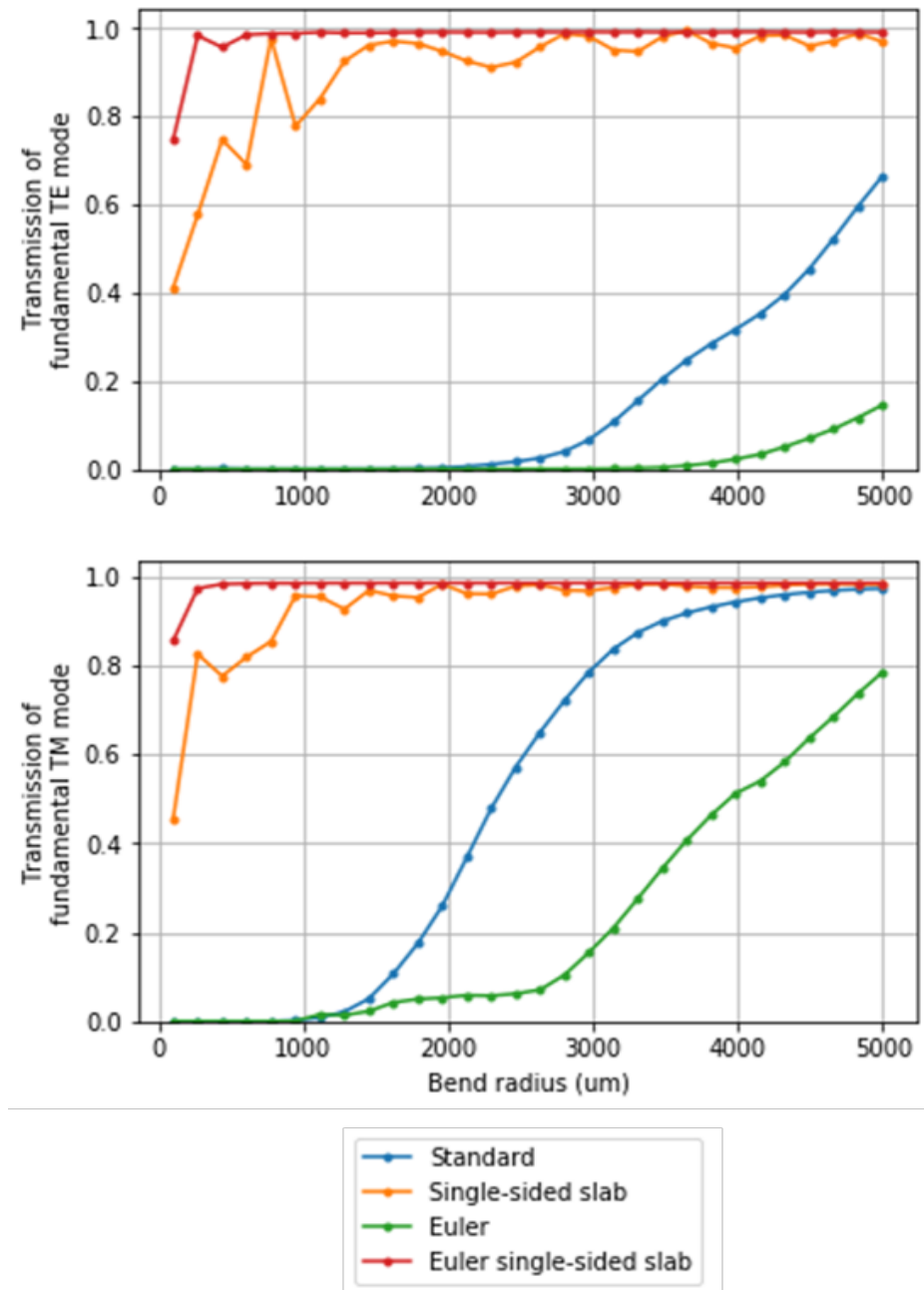


Figure 3.44: Bend loss comparison for several variations of the V5M5-V4M4V3 geometry, where the bend radius is defined relative to the area used by the standard bend. The standard bend refers to a waveguide bend with a slab on both sides of the ridge.

3.5 T03

The third chip taped out in the 180 nm process was intended to accomplish a variety of purposes, including attaining a conclusive waveguide loss measurement,

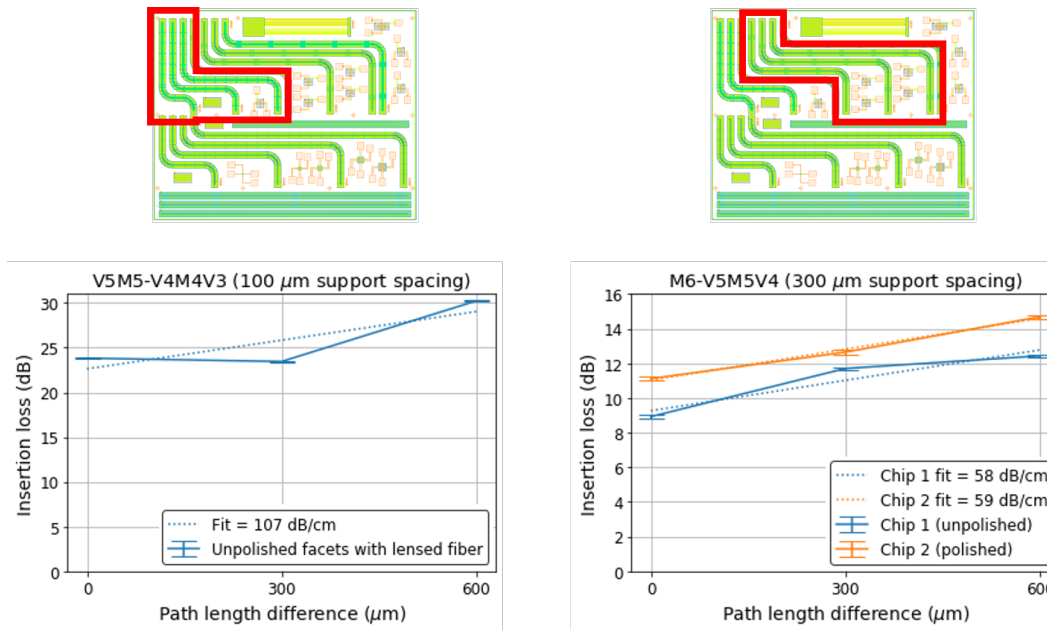


Figure 3.45: Additional waveguide measurements in TO2. Upper: Locations of the measured waveguides. Lower left: Insertion loss of V5M5-V4M4V3 singlemode waveguides. Lower right: Insertion loss of M6-V5M5V4 waveguides with 300 μm support spacing and 50 μm long supports.

demonstrating waveguide-photodiode coupling, and demonstrating the integration of electronics. The two main waveguide geometries used are M6-V5M5V4 and V5M5-V4M4V3. The former waveguides use a 5 μm ridge width, 10 μm support length, and 100 μm support spacing, unless noted otherwise. The latter waveguides use a 4 μm ridge width, 10 μm support length, and 100 μm support spacing, unless noted otherwise.

The chip included the following structures:

- Three M6-V5M5V4 waveguides with 5 mm path length differences
- One straight M6-V5M5V4 to serve as a reference
- Three M6-V5M5V4 waveguides with bend radii of 50, 75, and 100 μm
- Interlayer couplers, all with three layer slabs:
 - M6 \rightarrow V5M5 \rightarrow M6 \rightarrow V5M5 \rightarrow M6
 - V5M5 \rightarrow V4M4 \rightarrow V5M5 \rightarrow V4M4 \rightarrow V5M5
 - V5M5 \rightarrow V4M4 \rightarrow V5M5

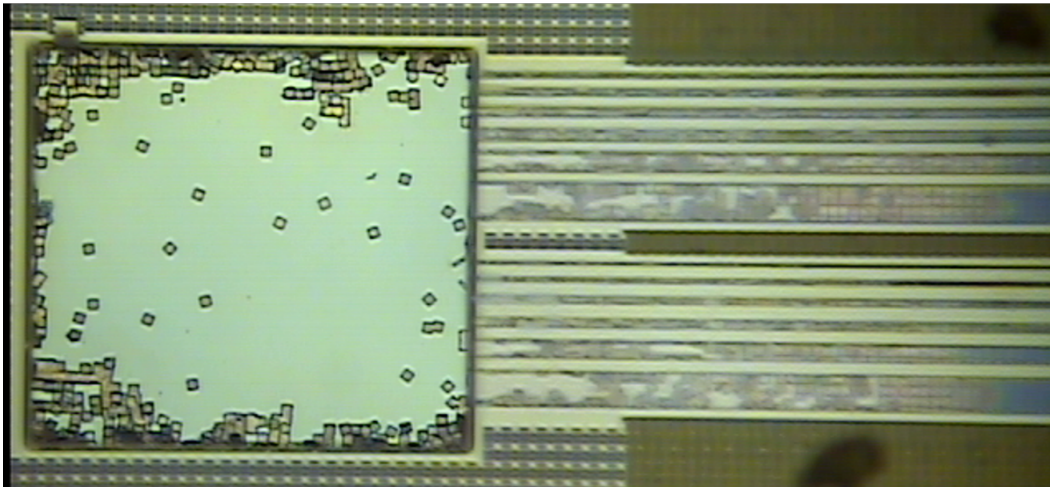
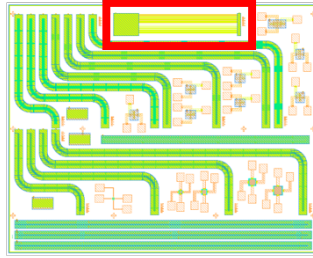


Figure 3.46: The etching and polymer deposition test structure after etching.

- V5M5 → V3M3 → V5M5
- V5M5-V4M4V3 waveguide coupled photodiodes, with photodiode areas of:
 - 16 x 150 μm
 - 50 x 150 μm
 - 10 segments of 16 x 23 μm
- V5M5-V4M4V3 waveguides coupled to TIAs through photodiodes.

The photodiode couplers are discussed in the following chapter.

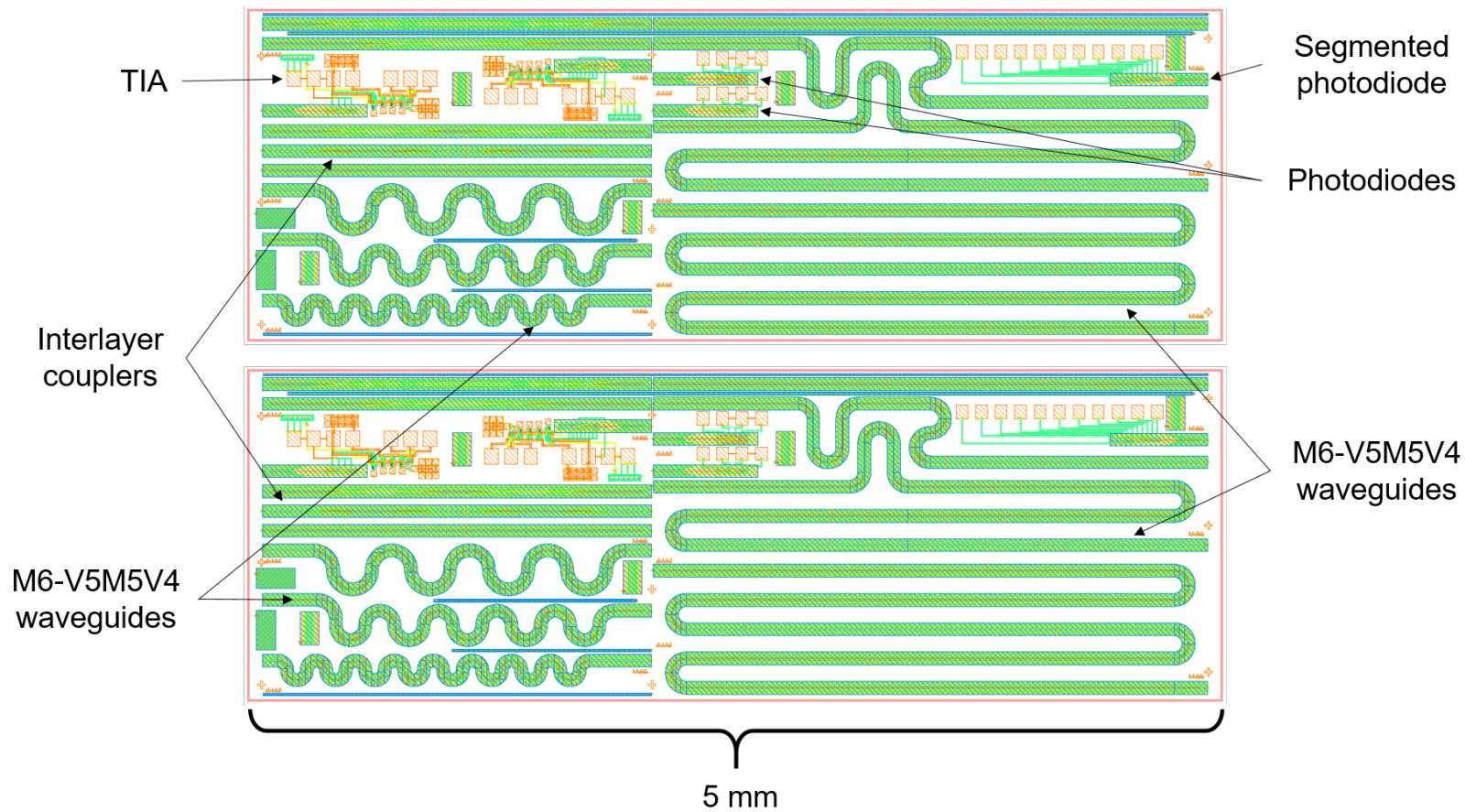


Figure 3.47: Layout of TO3.

Unfortunately, due to a design error, the ridges of the M6-V5M5V4 waveguides were not formed, and they could not be measured. Fortunately, the TO2 requirement that all via layers include waffling was dropped, enabling a return to designing with waffling on only the top via layer (Fig. 3.48). Although the waffling was generally removed cleanly during the etching, it caused problems with some of the interlayer couplers.

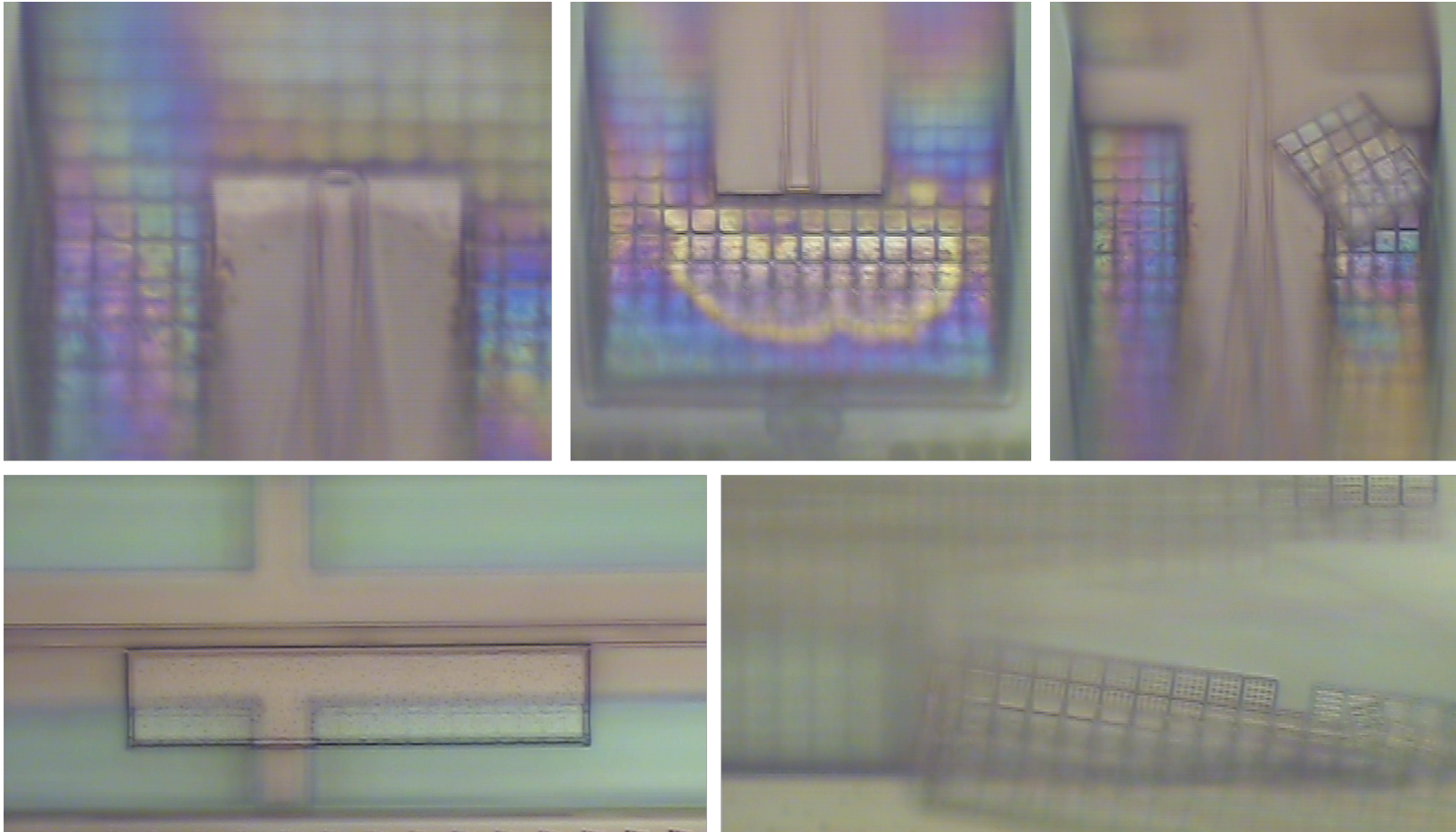


Figure 3.48: Microscope images of waffling in TO3. Note the transparency of the barrier layer in the lower left, and note the individual vias in the waffling image in the lower right.

A common DRC rule in electronics fabrication is that the metal surrounding a via must "enclose" it by a specified amount. Many of the vertical surfaces in TO0 had no overlap of the metal layers over the via layers, and so a large portion of the DRC errors in TO0 (which had more than 32 million DRC errors) were due to enclosure rules. These errors can be avoided by corrugating as many vertical surfaces in the chip as possible, as shown in the cavity sidewalls and slab edges of the waveguide in Figure 3.49. The corrugation results in strips of glass that line the cavity sidewalls, and unfortunately this strips of glass often peel off during the etching process. This can be clearly seen in the images of etch gauges in Figure 3.50.

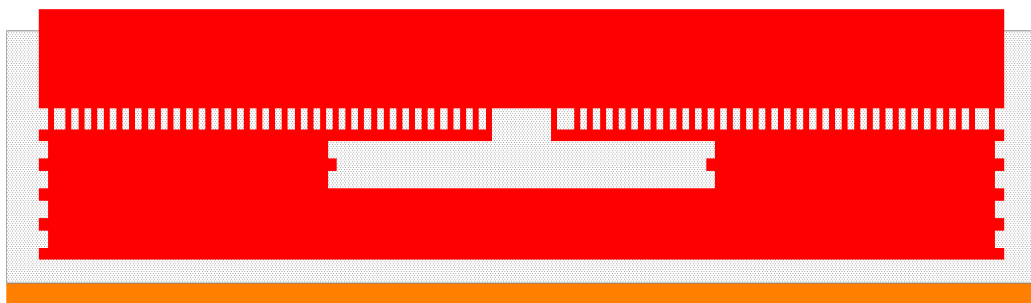


Figure 3.49: Example waveguide cross section.

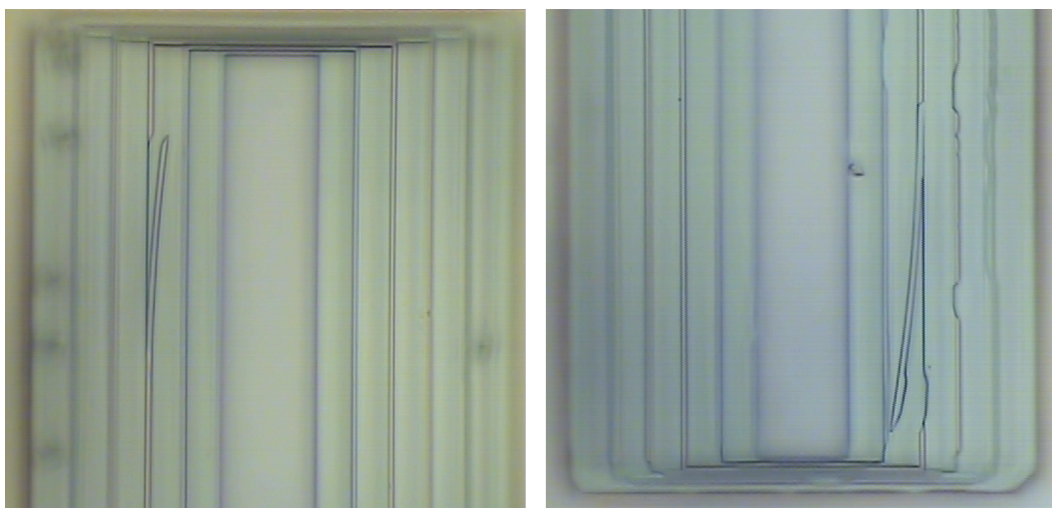


Figure 3.50: Microscope images showing peeling of glass overhangs.

Interlayer Couplers

One of the main advantages of subtractive photonics is the large design space opened up through the use of all of the back end layers. In the case of the 180 nm process

used in this work, there are six metal layers and five via layers². The large number of layers enables waveguides with exotic geometries, multilayer operation, optimized edge couplers, etc.

Given the large design space, there are a large number of possible ways to transition from one waveguide geometry to another. For a rib waveguide with a two layer ridge, one possibility is shown in Figure 3.51, in which the ridge and slab layers transition one layer at a time. The transition is depicted as a diagram in which the width of each layer is shown as a function of distance, as either ridge, slab, nothing, or a transition region. Another possibility is shown in Figure 3.52, in which two layers are tapered at a time. This results in a smoother transmission curve with coupler length, albeit with lower transmission at shorter lengths.

TO3 contains four different interlayer test structures, all shown in Figure 3.53. The first structure tests a transition between a M6-V5M5V4 multimode waveguide and a V5M5-V4M4V3 singlemode waveguide. The other three structures test transitions between a V5M5-V4M4V3 waveguide and other geometries with a two layer ridge and three layer slab.

Unfortunately, the interlayer couplers suffered from difficulties. Figure 3.54 is a typical view of the interlayer structures after etching. There is waffling residue scattered over the waveguides, some of the structures have waffling that has not been removed at all, there are barrier layers in various places, and the cavity sidewall corrugation is peeling.

The unremoved waffling is due to the insufficiency of the waffling algorithm, which was not designed to handle interlayer couplers. The difficulty can be seen in Figure 3.55, which shows a cross section of the bottom structure of Figure 3.53 at a point when the V5 ridge is tapering down. During this transition, the ridge width of the M5 layer is unchanged, so the oxide of the waffling of the V5 layer is merged with the oxide of the M5 ridge. Since the waffling has a tendency to be removed from the chip in sheets (Fig. 3.48), the waffling stuck to the M5 ridge leaves much of the surrounding waffling in place as well. This stuck waffling caused excessive loss in the interlayer couplers and they were not measurable.

The barrier layers only appear when certain layers are used. Figure 3.56 shows five cross sections to demonstrate that the brown barrier layer only appears when

²There is also an oxide open layer that is generally used together with the top metal layer, and so is not considered explicitly.

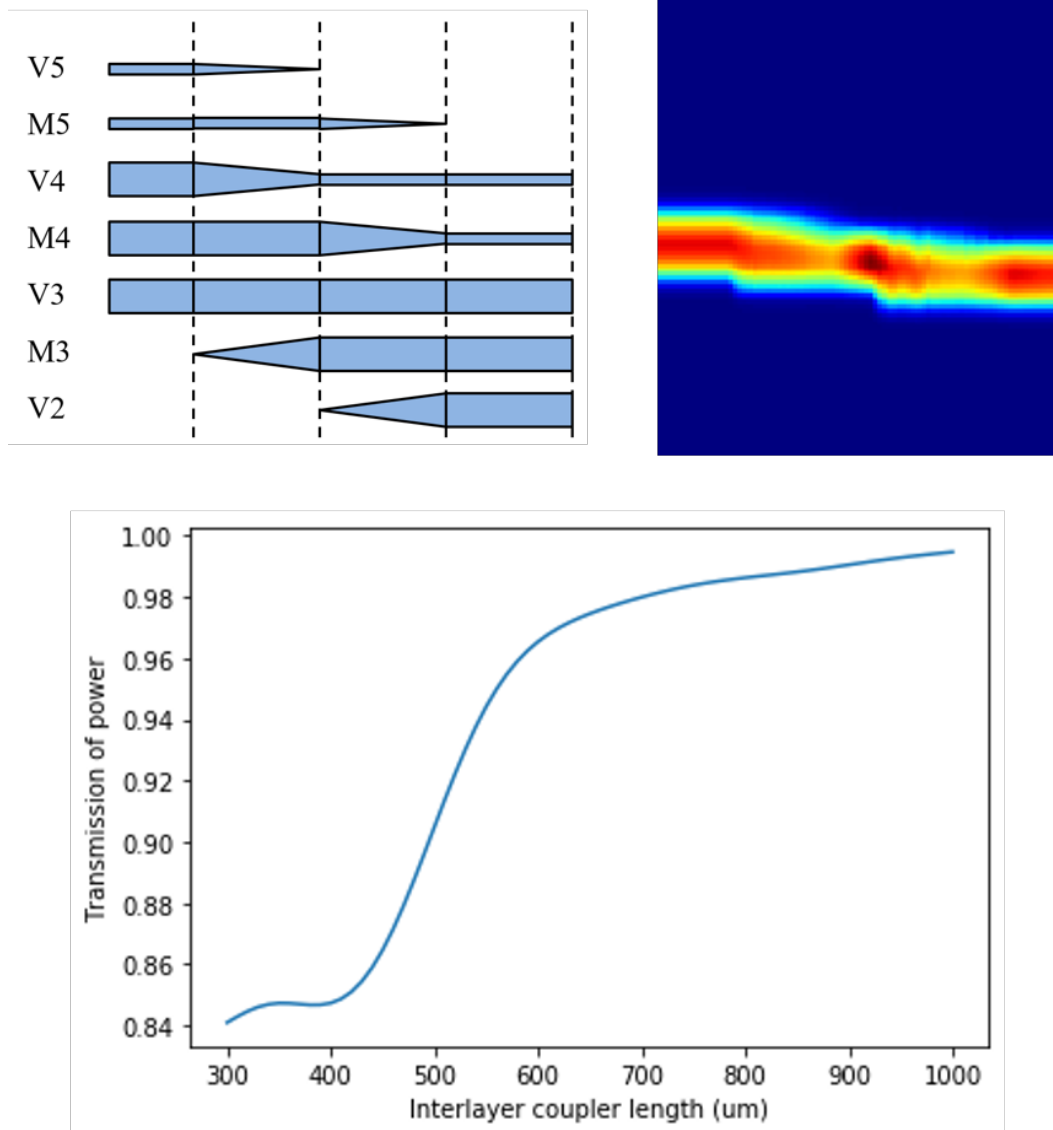


Figure 3.51: Example of an interlayer transition that tapers one layer at a time. Upper left: Transition diagram. Upper right: Side view of an FDTD simulation of the transition. Bottom: Transmission as the interlayer coupler length is varied.

M3 is the bottom oxide layer. This occurs on the supports of the waveguide in the microscope image of Fig. 3.56 due to the nature of the interlayer coupler design.

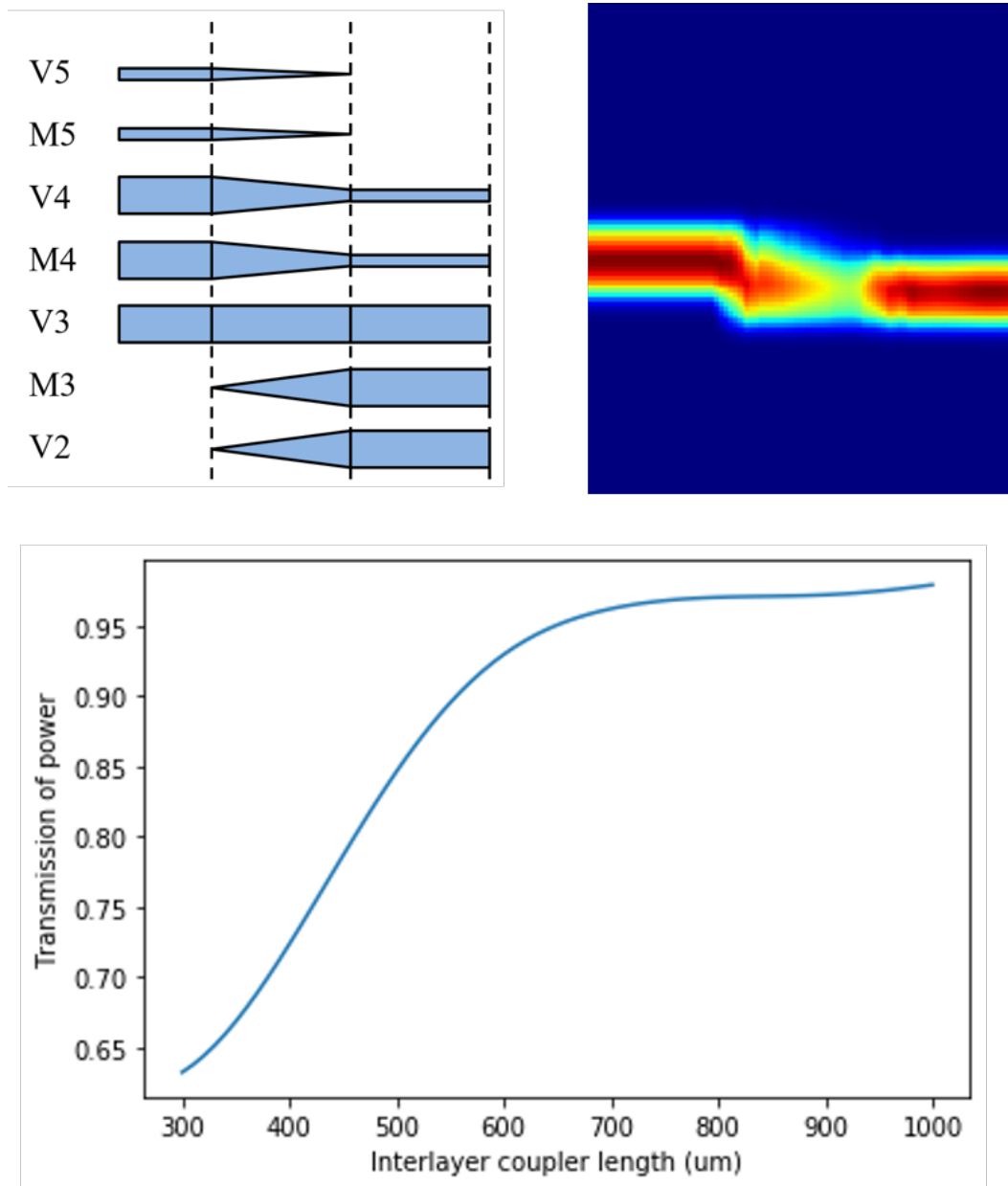


Figure 3.52: Example of an interlayer transition that tapers two layers at a time. Upper left: Transition diagram. Upper right: Side view of an FDTD simulation of the transition. Bottom: Transmission as the interlayer coupler length is varied.

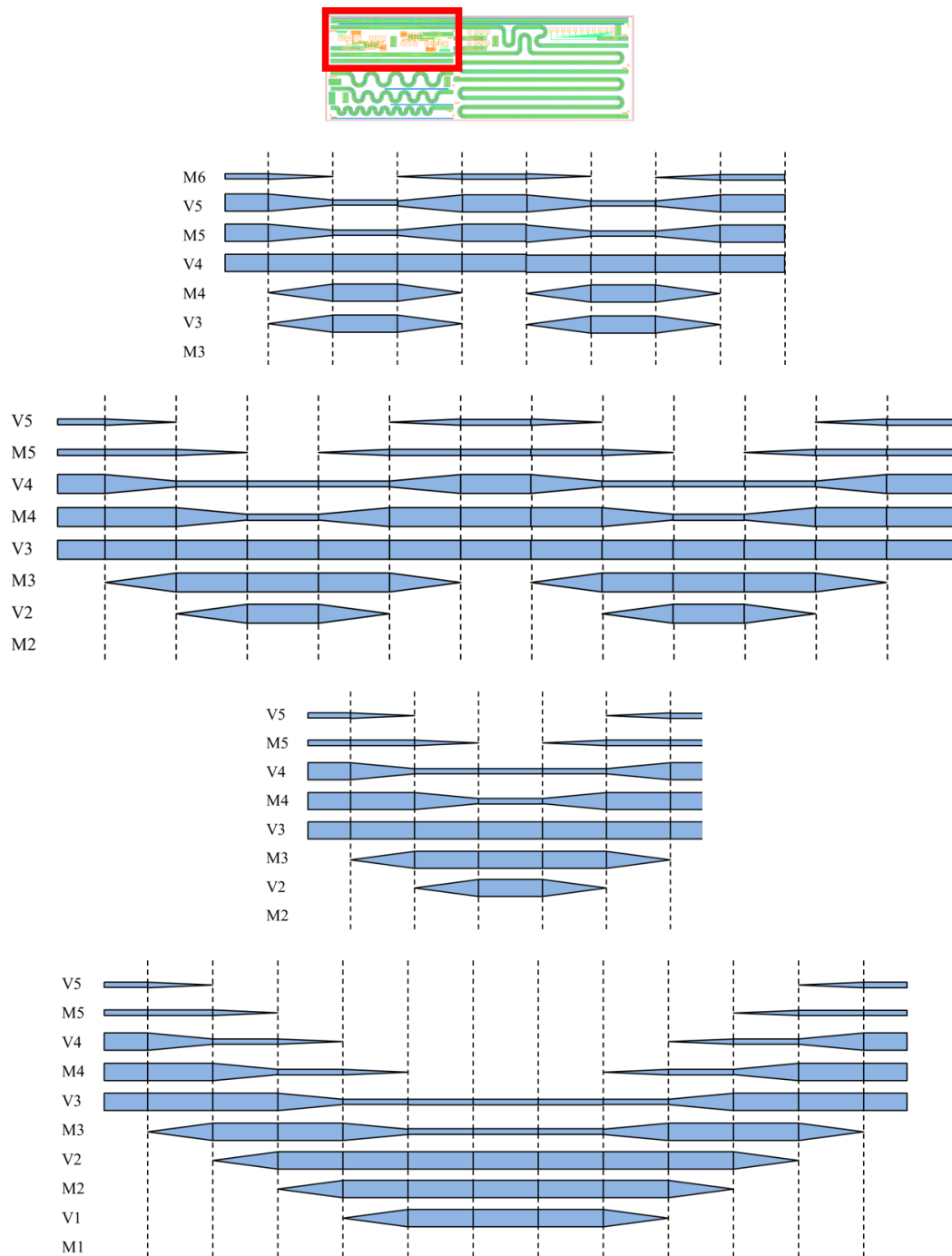


Figure 3.53: Transition diagrams depicting the interlayer transitions taped out in TO3.

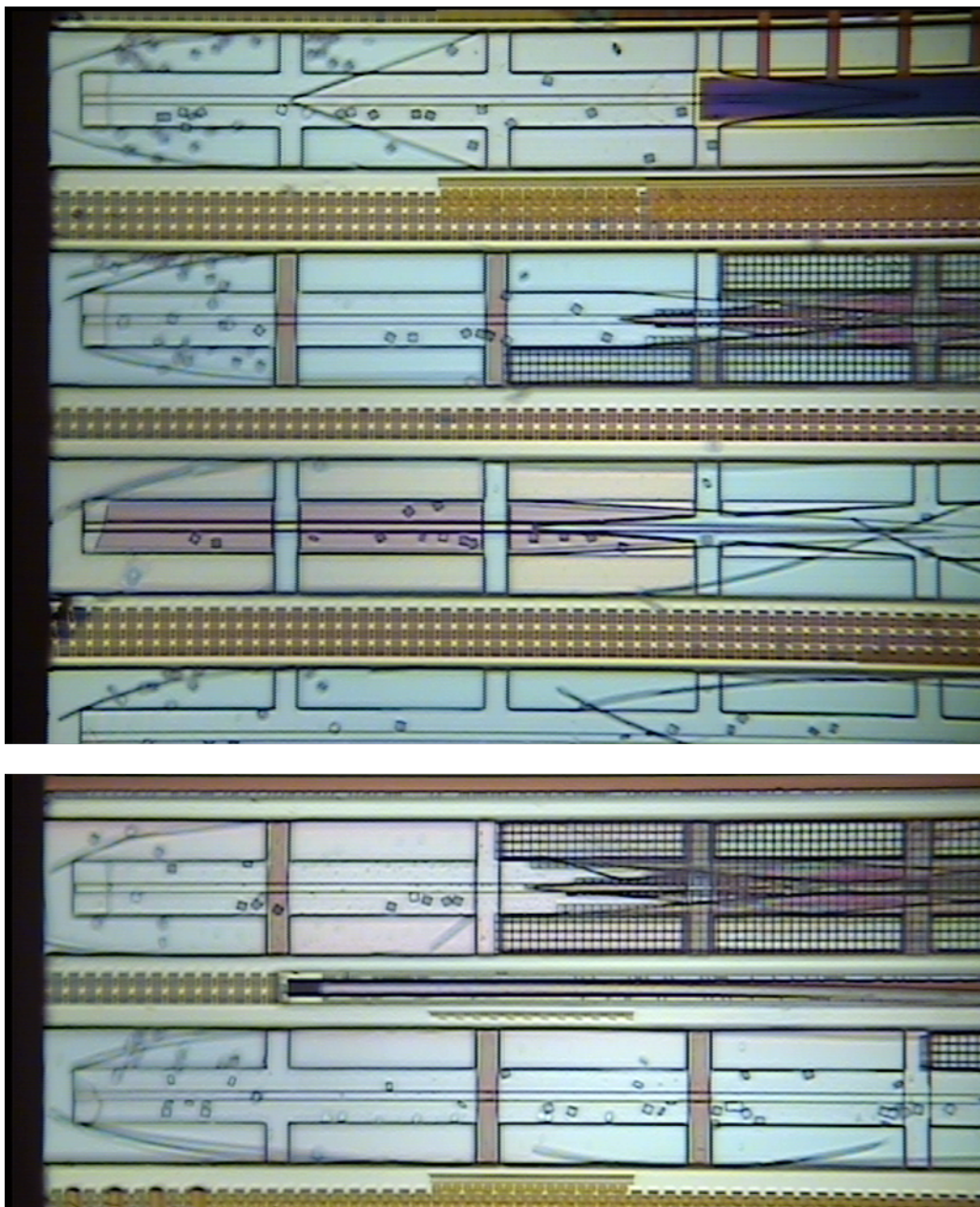


Figure 3.54: Microscope images showing waffling residue, peeling of sidewall corrugation, and barrier layers.

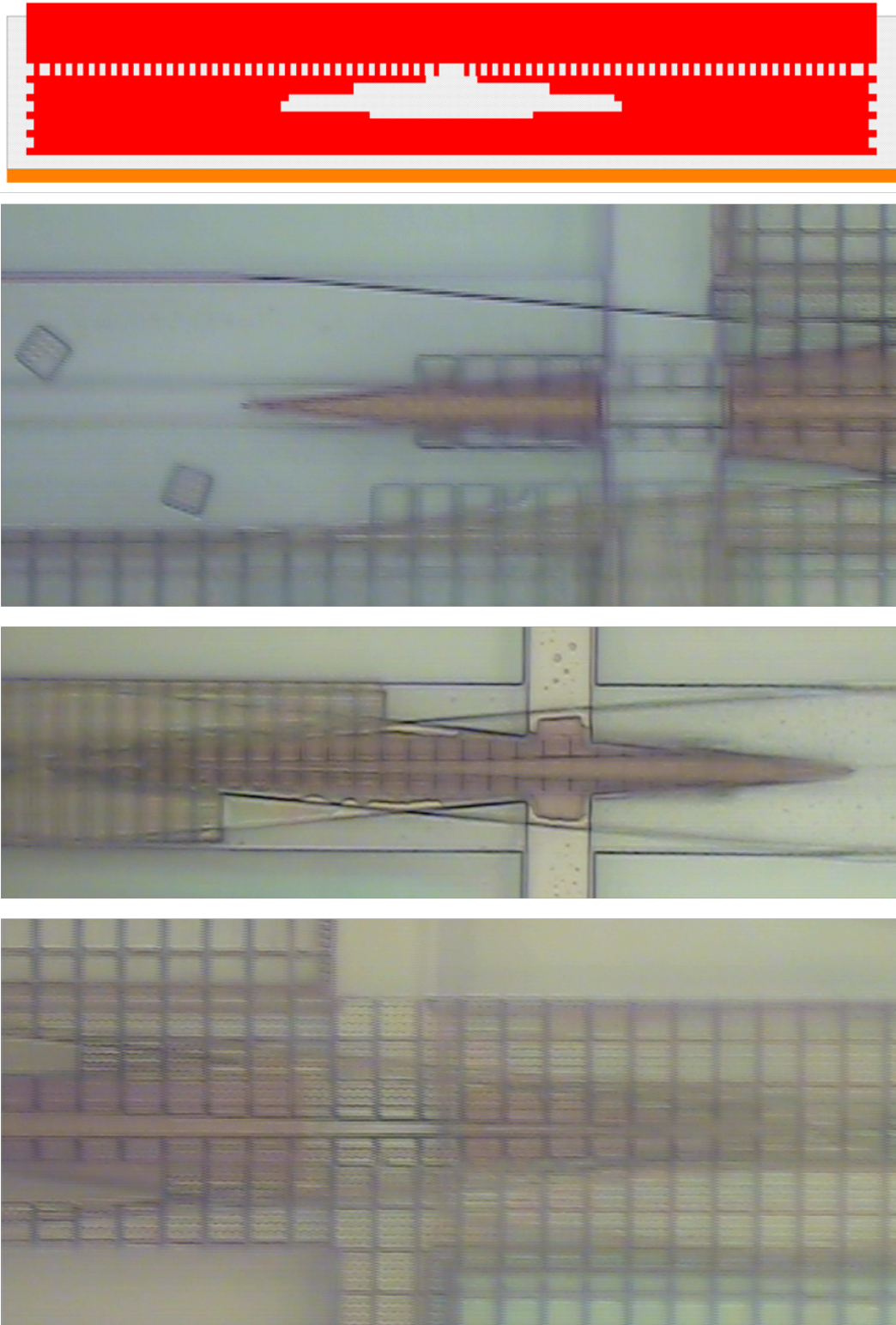


Figure 3.55: Waffling and interlayer couplers. Top: Cross section depicting the problem caused by using the waffling algorithm on an interlayer coupler. Bottom: Images of V5 tapering and the resulting stuck waffling.

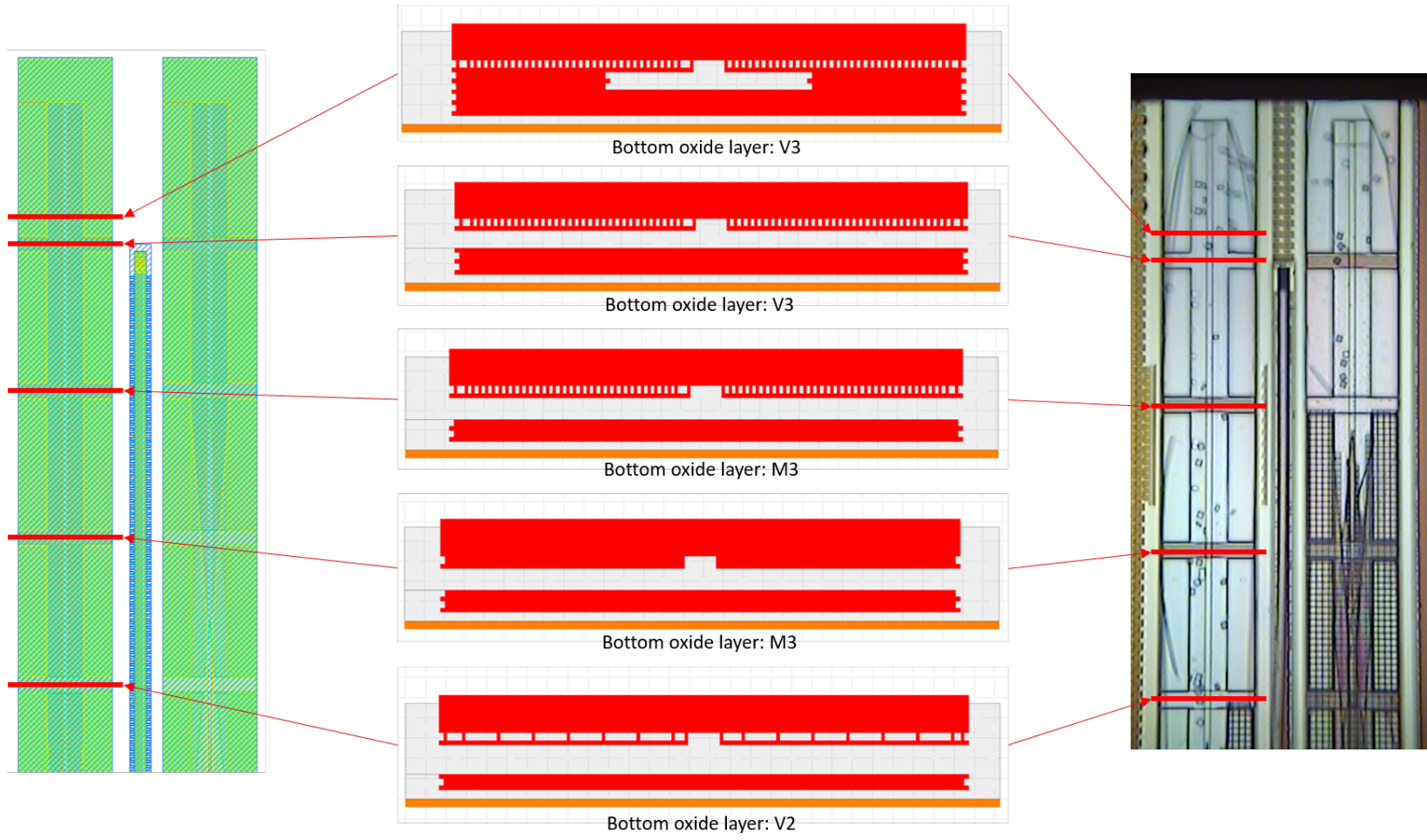


Figure 3.56: Examination of barrier layers on an interlayer coupler.

The layer dependency of the barrier is very clearly seen in Figure 3.57, in which the support has no discernible barrier layer but the surrounding waveguide taper does. The cross sections of Figure 3.58 show that only those areas in which M3 is the bottom oxide layer show the barrier layer. The support of Figure 3.57 has an extra layer in the slab due to the nature of the interlayer coupler design.

Figure 3.59 shows that there is an additional constraint on the barrier layer than just the bottom oxide layer. The geometry of the barrier layer in Figure 3.59 is restricted by the V4 layer, which, as the transition diagram shows, is the only layer besides M3 that is tapering.

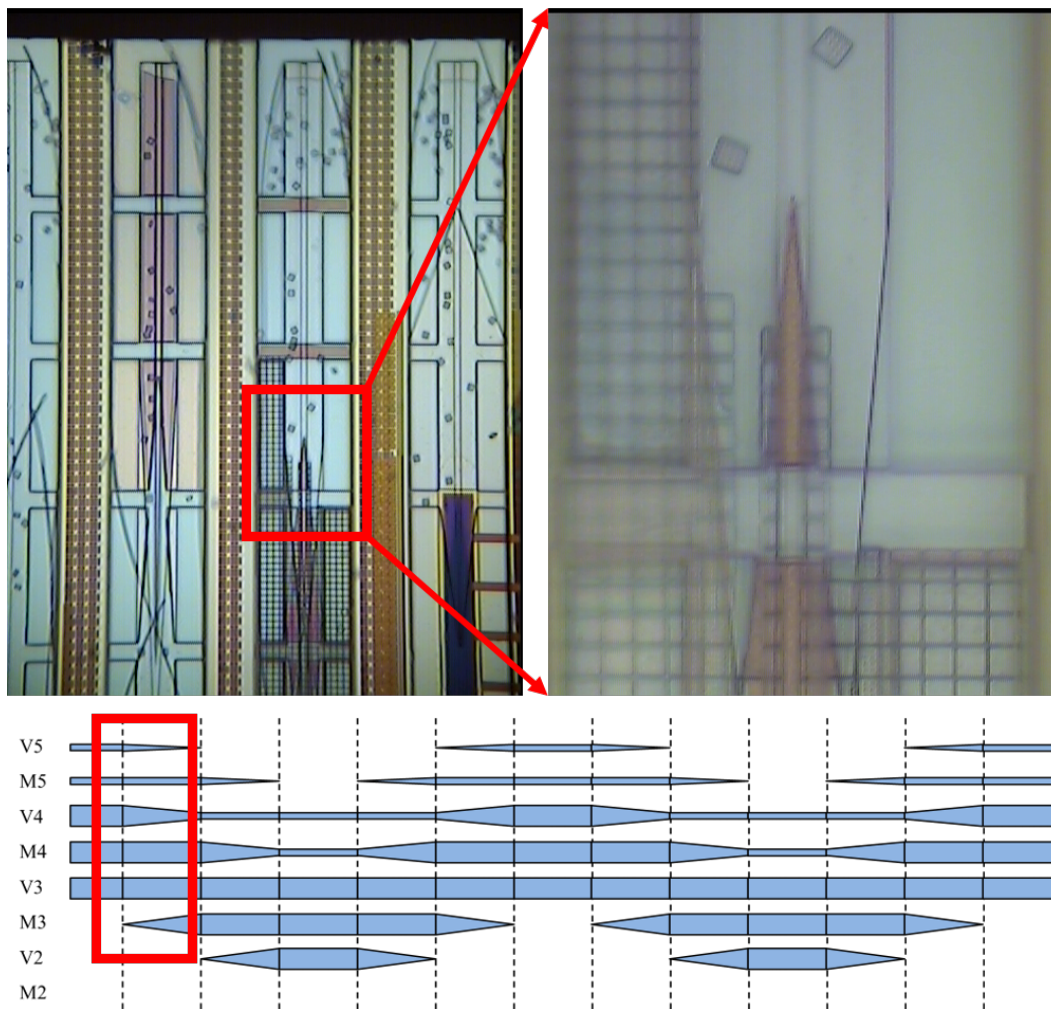


Figure 3.57: Close view of a waveguide taper to illustrate the layer dependency of the barrier.

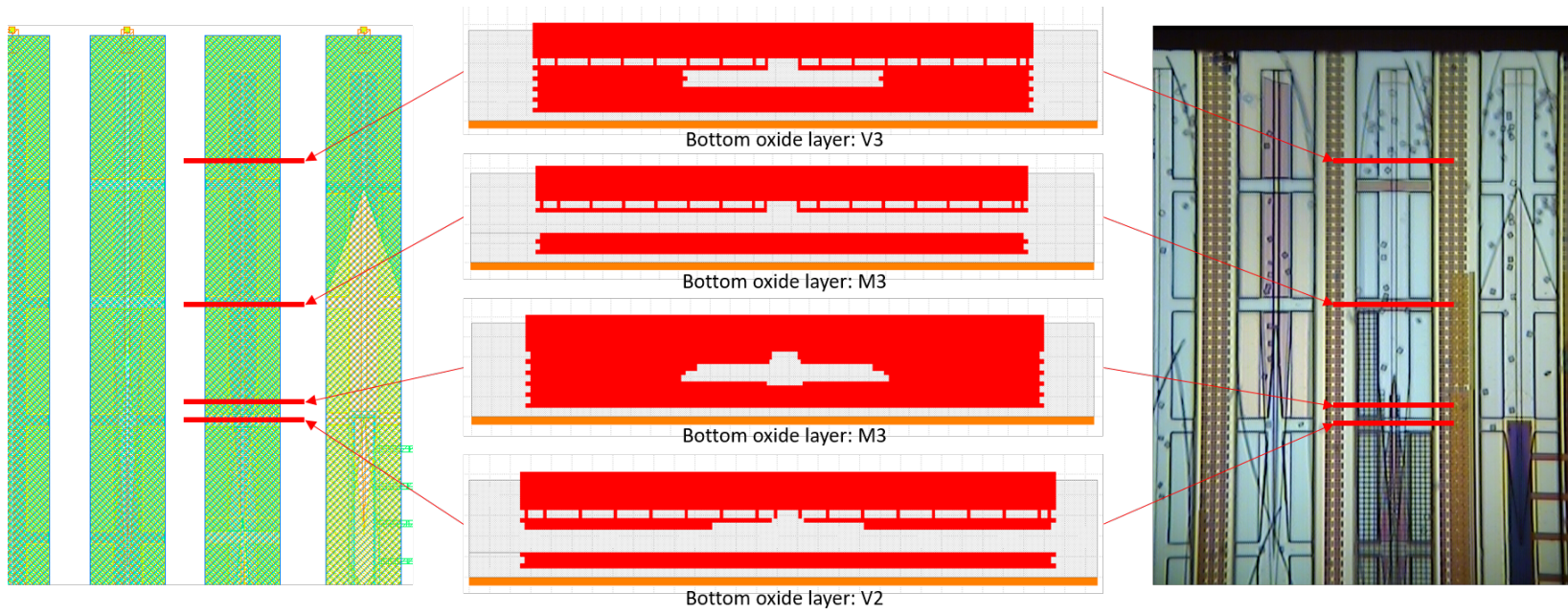


Figure 3.58: Examination of barrier layers on an interlayer coupler.

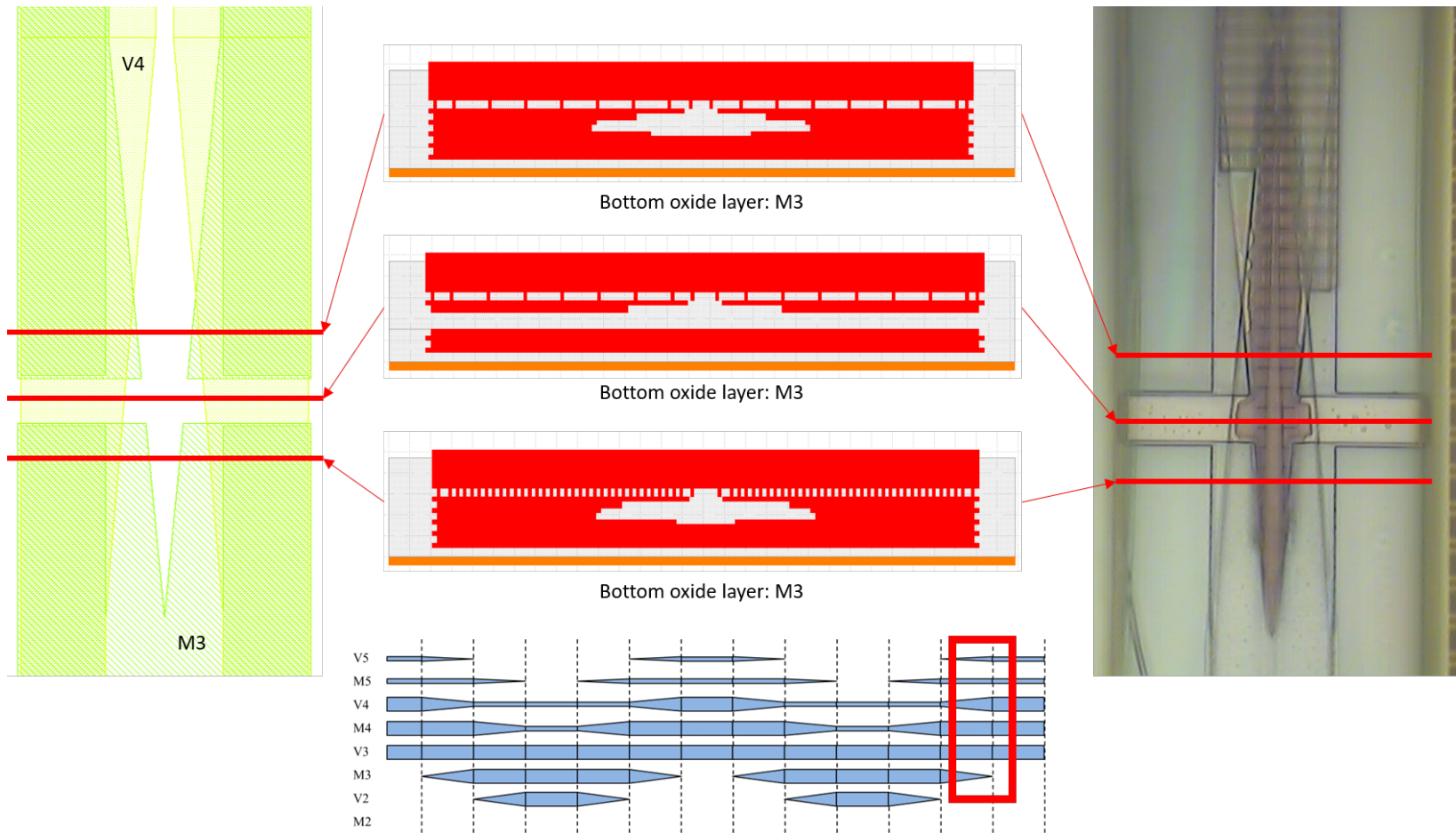


Figure 3.59: Microscope photo demonstrating that the barrier layer is constrained by both the M3 and the V4 layers.

Another example of layer dependency is shown in Figure 3.60. The barrier layer in Fig. 3.60 appears as a different color than the barrier layer in Fig. 3.57, reflecting the fact that it depends on different layers. As the transition diagram of Figure 3.60 and the cross sections of Figure 3.61 show, the barrier layer follows the contours of the V5 and M5 layers, but does not exist where the M4 and V3 layers are.

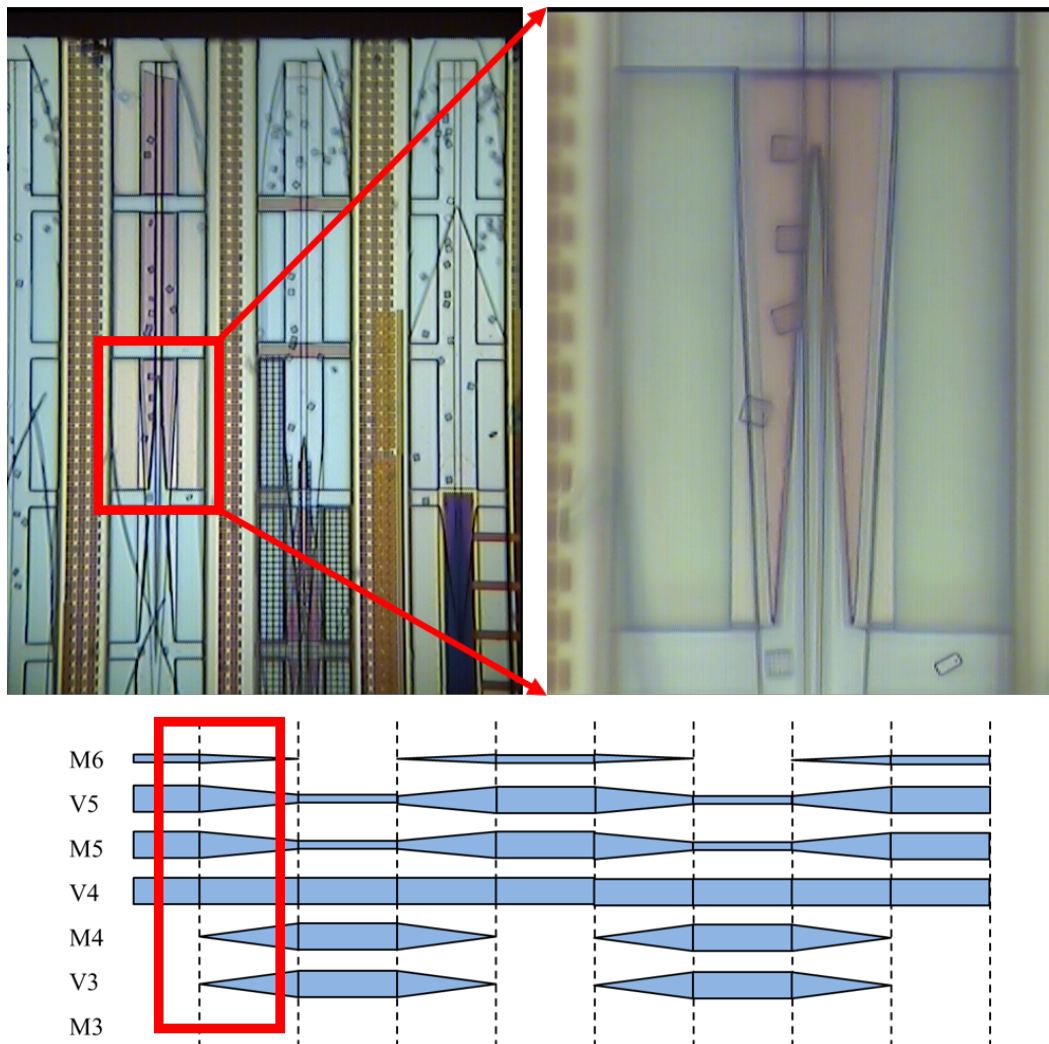


Figure 3.60: Close view of a waveguide taper to illustrate the layer dependency of the barrier.

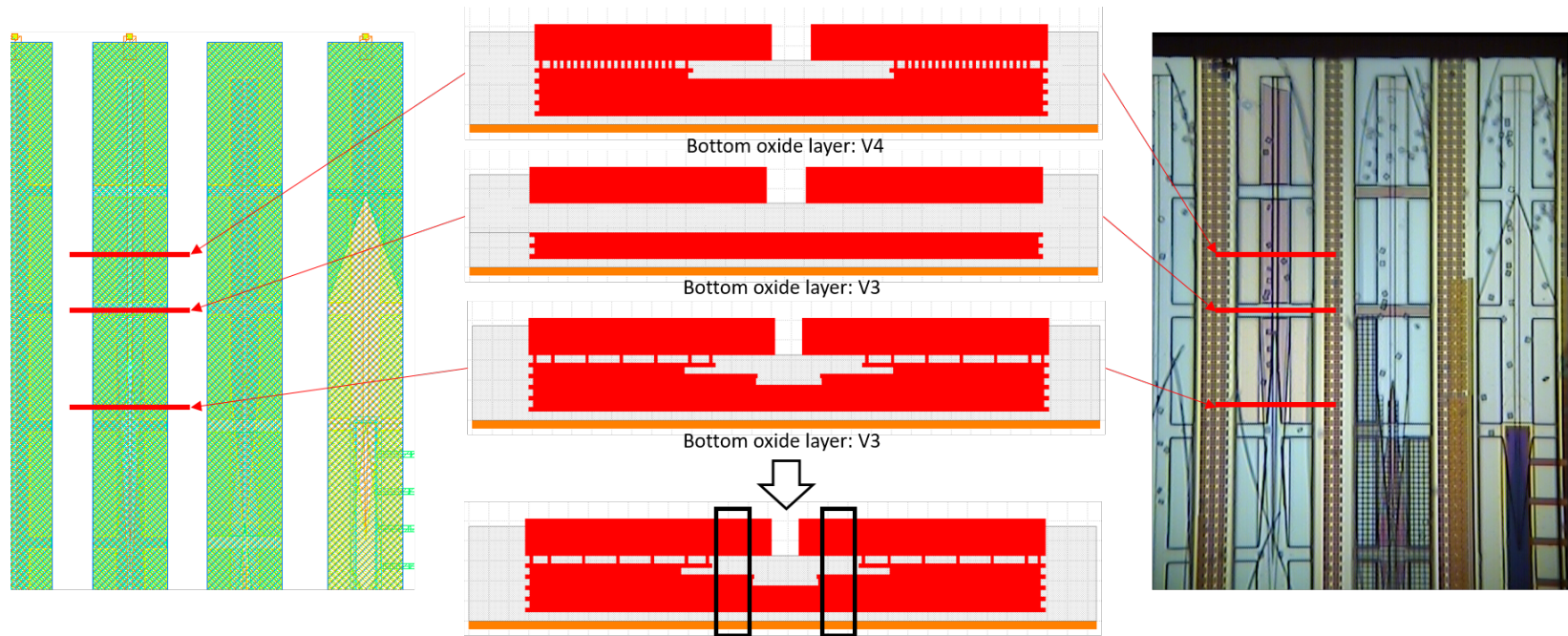


Figure 3.61: Examination of barrier layers on an interlayer coupler. The black boxes in the bottom cross section denote the place where the barrier layer exists.

The barrier layer observations may be summarized as follows:

- Barrier layers hold together the wafling and may be transparent (Fig. 3.48)
- An opaque barrier layer appears when M3 is the bottom oxide layer of a waveguide, but is constrained by the V4 layer (Fig. 3.59)
- An opaque barrier layer follows the geometry of the V5 and M5 layers in a M6-V5M5V4 waveguide, but is constrained by the V3 and M4 layers if those are added to the structure (Fig. 3.61).

3.6 TO4

Given previous difficulties with the cutback method for determining waveguide loss, the fourth chip taped out in the 180 nm process included waveguide structures for determining an upper bound on waveguide loss. The two main waveguide geometries used are M6-V5M5V4 and V5M5-V4M4V3. The former waveguides use a 5 μm ridge width, 20 μm support length, and 100 μm support spacing, unless noted otherwise. The latter waveguides use a 3 μm ridge width, 20 μm support length, and 100 μm support spacing, unless noted otherwise.

The chip included the following structures:

- A 4.3 cm long M6-V5M5V4 spiral
- A 3.5 cm long V5M5-V4M4V3 spiral
- Five M6-V5M5V4 waveguides with 1 mm path length differences
- Reflecting waveguide-photodiode couplers
- Photodiode test structures
- MEMS directional couplers for modulation.

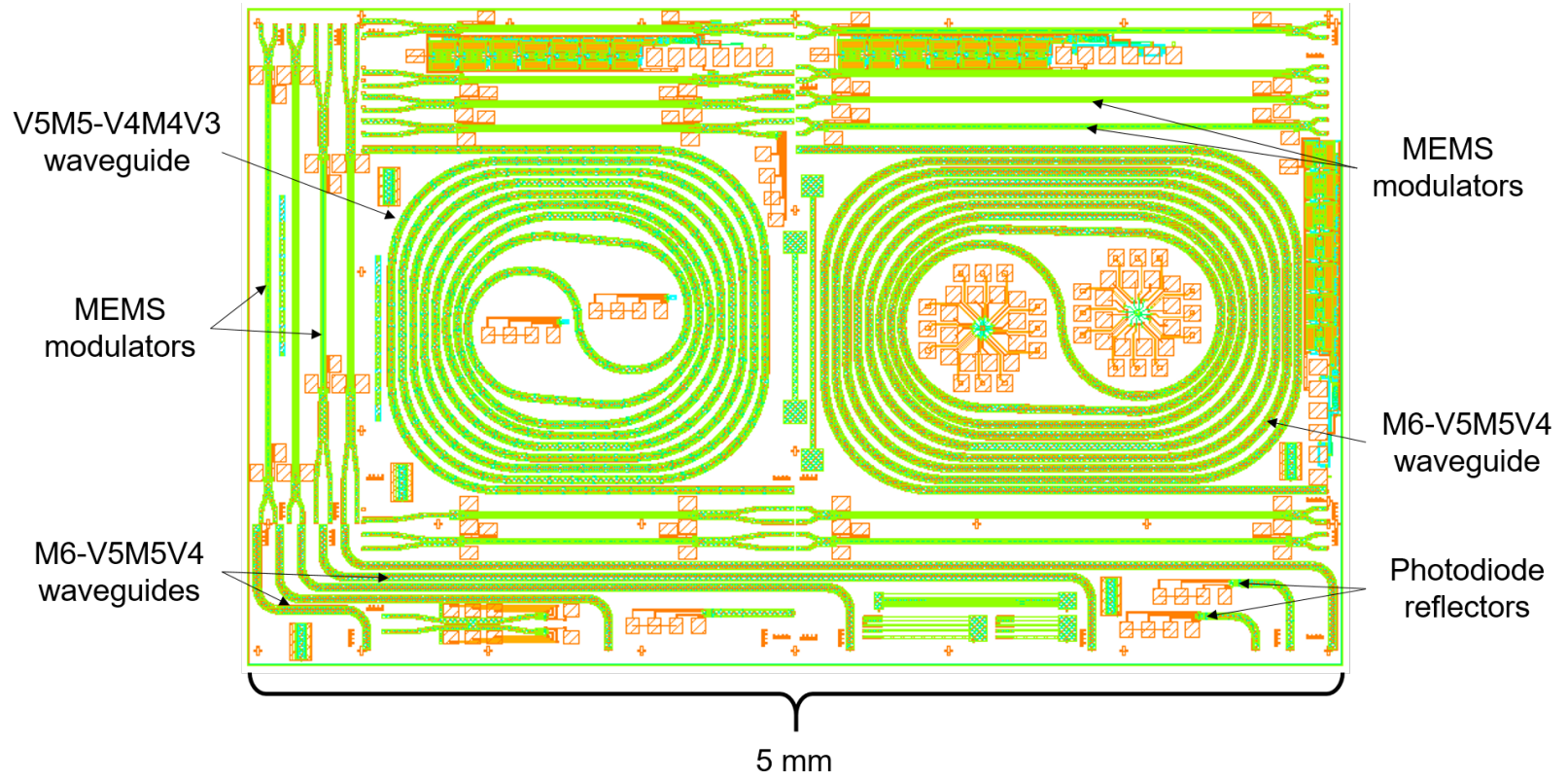


Figure 3.62: Layout of TO4.

M6-V5M5V4

The geometry of the M6-V5M5V4 waveguide was changed for TO4 to reduce the width of the waveguide (Fig. 3.63). The support width was reduced from $20\ \mu\text{m}$ to $7.5\ \mu\text{m}$, and the support length was increased from $10\ \mu\text{m}$ to $20\ \mu\text{m}$. The change in support length had a negligible effect on etching.

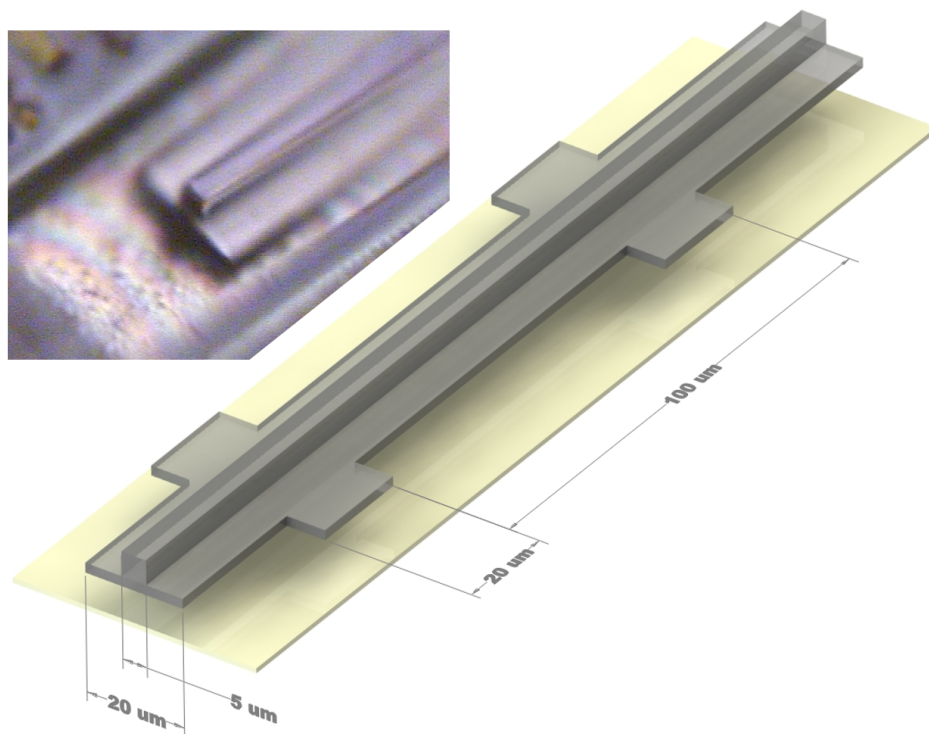


Figure 3.63: Structure of the M6-V5M5V4 waveguide used in TO4.

One structure using the M6-V5M5V4 waveguide is a 4.3 cm spiral, as shown in Figure 3.64. Similarly to TO3, the M6-V5M5V4 waveguide exhibited a purple barrier layer after being etched in AETA. Similarly to previous tapeouts, the barrier layer was removed cleanly using a solution of hydrogen peroxide and EDTA (Fig. 3.65).

The goal of the spiral structure was to determine an upper bound on the waveguide loss, as done in TO2, by dividing the insertion loss by the waveguide length (eq. 3.2). The measurement setup, which is similar to those used for past waveguide measurements, is shown in Figure 3.66. The lensed fibers used in the upper bound measurement are designed for operation at 1550 nm, have a spot size of $5\ \mu\text{m}$, and a working distance of $26\ \mu\text{m}$ (Fig. 3.67).

The results of the upper bound measurement are shown in Figures 3.68 and 3.69.

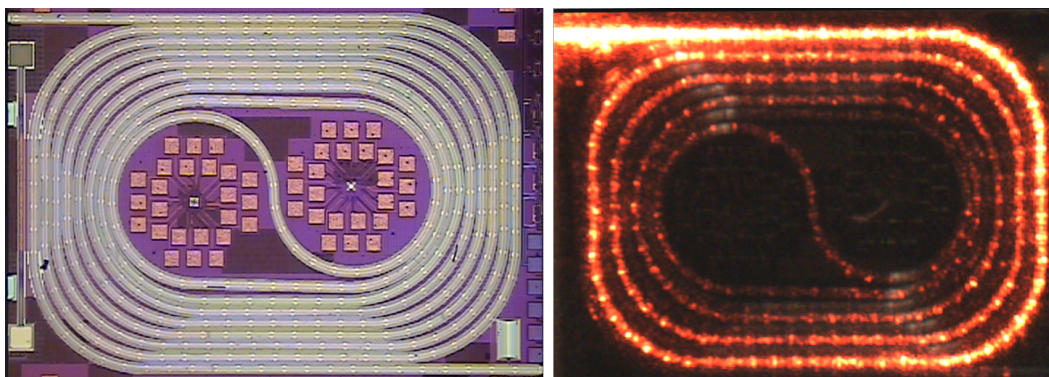


Figure 3.64: Images of the M6-V5M5V4 spiral. Left: Microscope photo. Right: The spiral illuminated with 635 nm light.

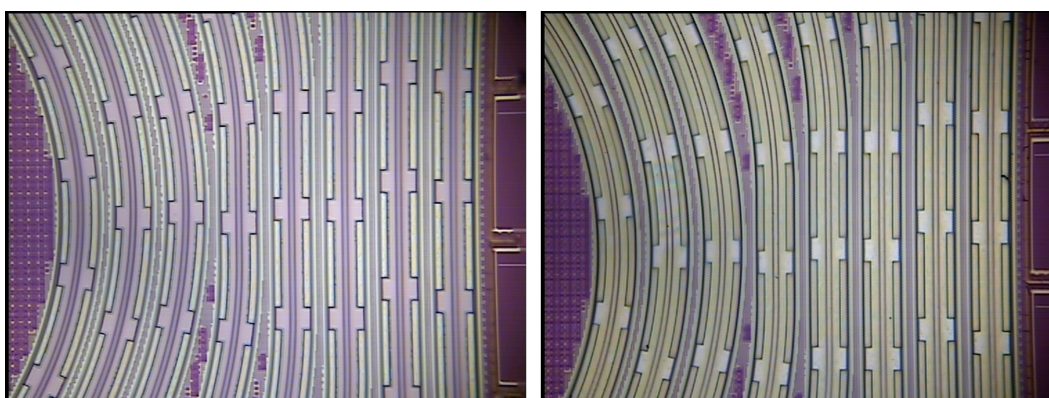


Figure 3.65: Etching comparison of the M6-V5M5V4 spiral. Left: The spiral after etching in AETA for 4 hours at 80°C. Right: The spiral after additional etching in AETA and H₂O₂/EDTA.

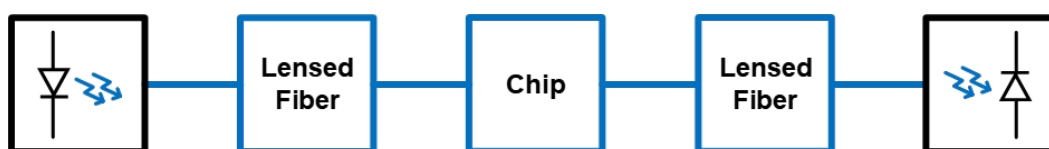


Figure 3.66: Block diagram of the spiral measurement setup.

The best results are 4.1 dB/cm at 1550 nm, 4.8 dB/cm at 780 nm, and 7.8 dB/cm at 635 nm. The dashed lines in Fig. 3.68 provide a reference to the two previously reported results in unmodified bulk CMOS [47], [55]. There do not seem to be any previously reported results for waveguide loss at visible wavelengths in unmodified bulk CMOS.

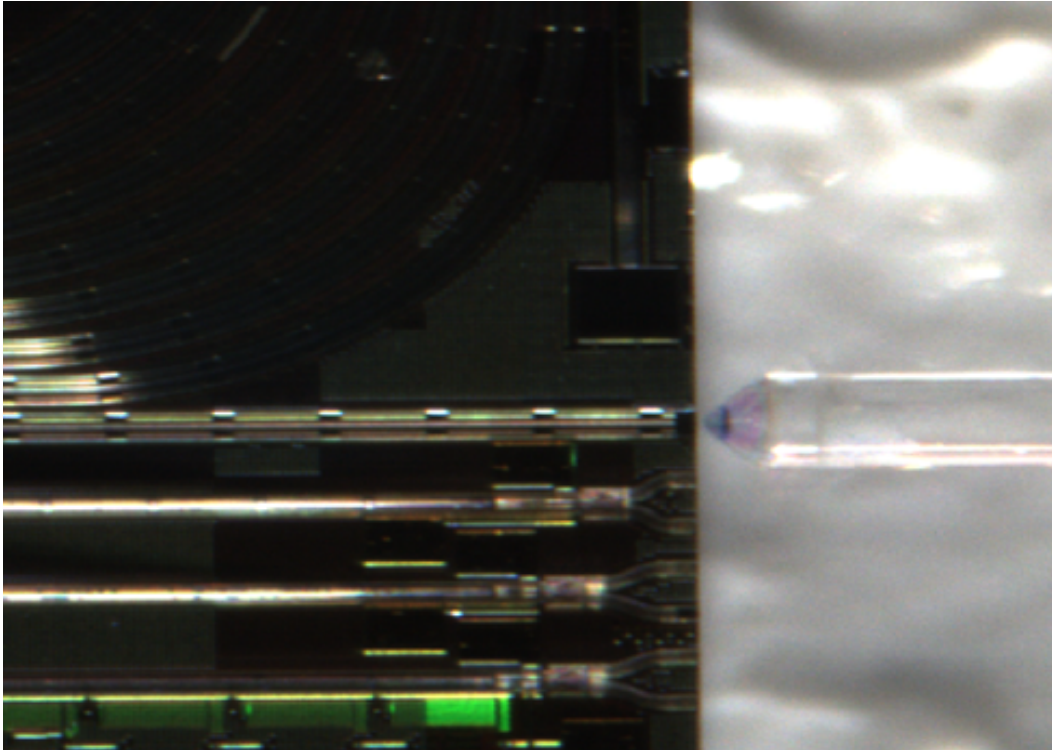


Figure 3.67: Edge coupling setup with a 1550 nm lensed fiber.

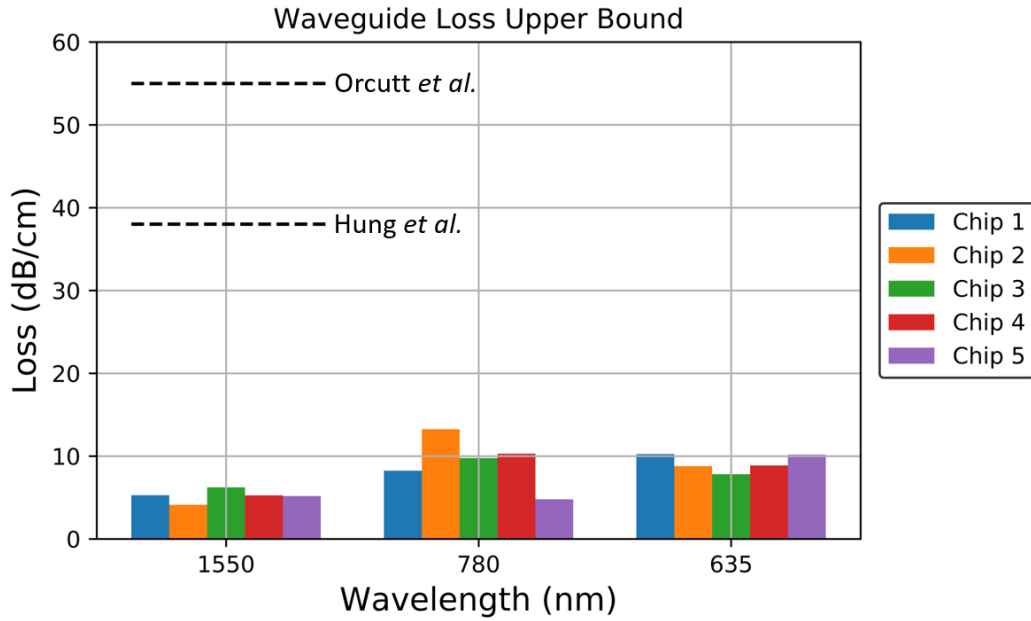
V5M5-V4M4V3

The V5M5-V4M4V3 waveguide used in TO4 uses the same geometry as the V5M5-V4M4V3 waveguide in TO3 except the ridge was reduced to $3\ \mu\text{m}$, which preserves the single-mode behavior (Fig. 3.40). The V5M5-V4M4V3 spiral was measured using a similar setup as the M6-V5M5V4 waveguide (Fig. 3.70), but with the addition of an EDFA due to the higher loss of 20.7 dB/cm at a wavelength of 1550 nm. Results were not obtained for visible wavelengths due to a lack of suitable amplification. This high upper bound is most likely a result of the bend loss in the spiral, since a standard V5M5-V4M4V3 waveguide was used (no Euler bends or single-sided slab).

MEMS Modulators

One possibility for a monolithic modulation mechanism is that of mechanical movement. The MEMS industry has explored mechanical movement in the context of CMOS platforms extensively, and MEMS modulators for silicon photonics have been demonstrated [90]. The question arises as to whether mechanical movement can be used in the context of subtractive photonics to enact modulation.

Consider a beam with a central mass depicted in Figure 3.71. The electrostatic force



Wavelength (nm)	1550	780	635
Average loss (dB/cm)	5.2	9.3	9.2
Standard deviation (dB/cm)	0.7	2.8	0.9

Figure 3.68: Upper bound on the loss of the M6-V5M5V4 waveguide for various wavelengths and chip samples. The dashed lines refer to the waveguide loss reported in Orcutt *et al.* [47] and Hung *et al.* [55].

Chip	Loss at 1550 nm (dB/cm)	Loss at 780 nm (dB/cm)	Loss at 635 nm (dB/cm)
1	5.3	8.2	10.3
2	4.1	13.2	8.8
3	6.2	9.8	7.8
4	5.3	10.3	8.9
5	5.2	4.8	10.2

Figure 3.69: Table of the values plotted in Fig. 3.68.

between the central mass and a nearby electrode is [91]

$$F_e = \frac{\epsilon AV^2}{2(d-y)^2}. \quad (3.3)$$

where A is the overlapping area between the electrode of the central mass and the

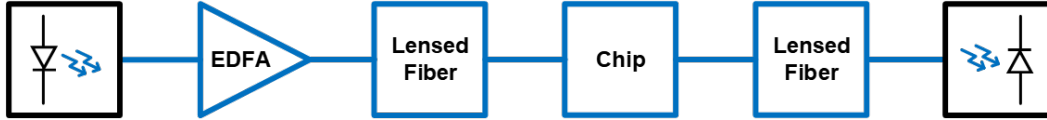


Figure 3.70: Block diagram of the V5M5-V4M4V3 spiral measurement setup.

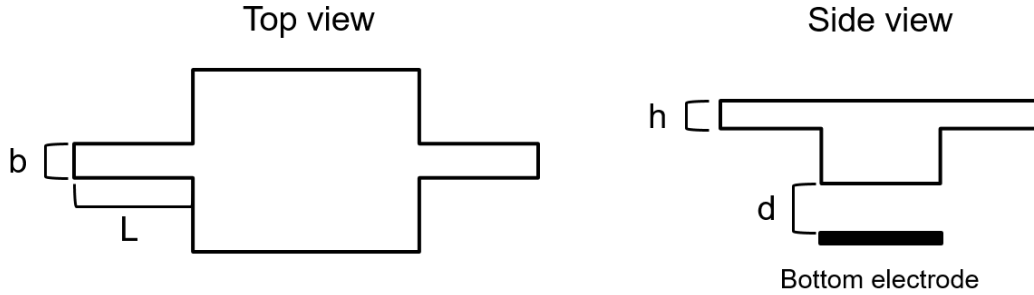


Figure 3.71: A beam with a central mass suspended above an electrode.

electrode below it.

The elastic force is

$$F_k = -ky. \quad (3.4)$$

Normalizing the distance and the force, Figure 3.72 plots these two forces with a variety of voltages applied for the electrostatic force. For $V > V_{pull-in}$, the electrostatic force is greater than the elastic force for all displacements and the central mass snaps to the electrode, which is called pull-in. For $V < V_{pull-in}$, there are two points of intersection between F_e and F_k . The left intersection is a stable point since the elastic force prevents greater displacements, and the electrostatic force prevents smaller displacements. The balance between the forces is formulated mathematically by writing [91]

$$\frac{\varepsilon AV^2}{2(d-y)^2} - ky = 0. \quad (3.5)$$

The stability is formulated mathematically by requiring $\partial F/\partial y < 0$ [91], resulting in

$$\frac{\varepsilon AV^2}{(d-y)^3} - k < 0. \quad (3.6)$$

Substituting 3.5 in 3.6, we find that

$$y < \frac{1}{3}d. \quad (3.7)$$

Consequently, if the displacement of the electrodes ever exceeds 1/3 of the starting displacement, the stability is lost.

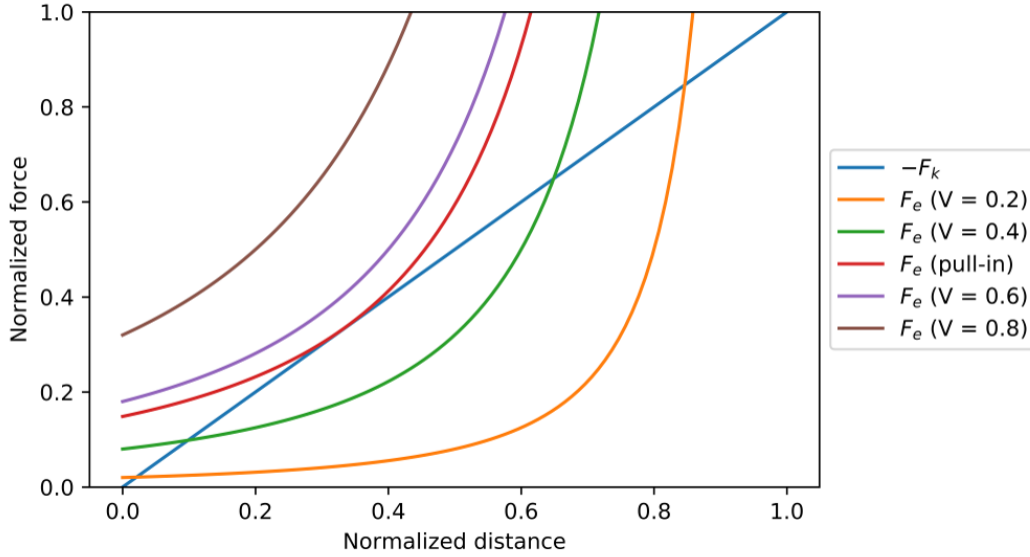


Figure 3.72: Elastic force and electrostatic force demonstrating the end of stability at the pull-in voltage.

The spring constant of the of the beam with central mass is known to be [91]

$$k = \frac{2Ebh^3}{L^3} \quad (3.8)$$

where E is the Young's modulus and the other parameters are defined in Figure 3.71. Using Equations 3.3, 3.4, and 3.8, the pull-in voltage can be solved for:

$$V^2 = \frac{16Ebh^3d^3}{27\epsilon L^3 A}. \quad (3.9)$$

Since h and d are severely constrained by the BEOL dielectric stackup, the most important parameter in Equation 3.9 is L , the length of the support arm. This variable can always be adjusted to bring the voltage into a desired range, no matter the distance between the electrodes or the overlapping area of the electrodes. This is critical since the minimum distance between the electrodes is determined by the minimum feature size of the process, and the area of the electrodes is mainly determined by the optical coupling length.

The modulation is accomplished using a directional coupler. A cross section of the coupling region is shown in Figure 3.73. The upper portion of each arm is a V5M5-V4M4 waveguide, with an isolation layer using V3 underneath. The lower portion of each arm holds the electrodes, which must be fully encased in glass to prevent them from being etched. Since the glass thickness on each electrode and

the gap distance are the same, this insulation serves a dual purpose of ensuring that the electrodes will never move more than one third of the distance between them, ensuring stability by Equation 3.7. In the design of Figure 3.73, the electrodes are formed on the V2 and M2 layers, so the lower glass consists of V1-M3. The gap distance and electrode insulation thickness are both 350 nm.

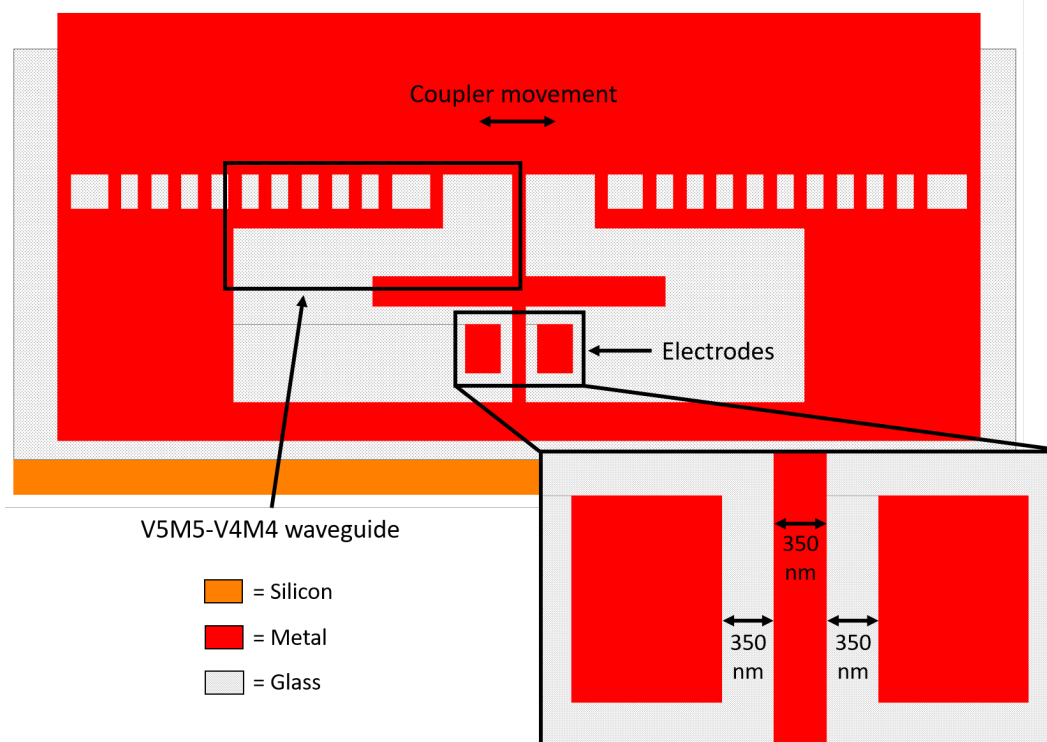


Figure 3.73: Cross section of the coupling portion of the MEMS directional coupler.

The waveguides transition into the coupling region using the layout shown in Figure 3.74. Both the real metal layout (the tapeout geometry) and the oxide structures formed through etching (the complement of the tapeout geometry) are shown, although only layers V5 and M2 are shown for simplicity. The V5 layer shows the path of the waveguide ridge layer, while the M2 layer shows how the electrode connection is made from the coupling region through the support, fully encased in glass. L , the critical parameter that enables voltage scaling, is also shown. A 3D visualization of the coupling region is shown in Figure 3.75, and a micrograph of the fabricated transition region is shown in Figure 3.76.

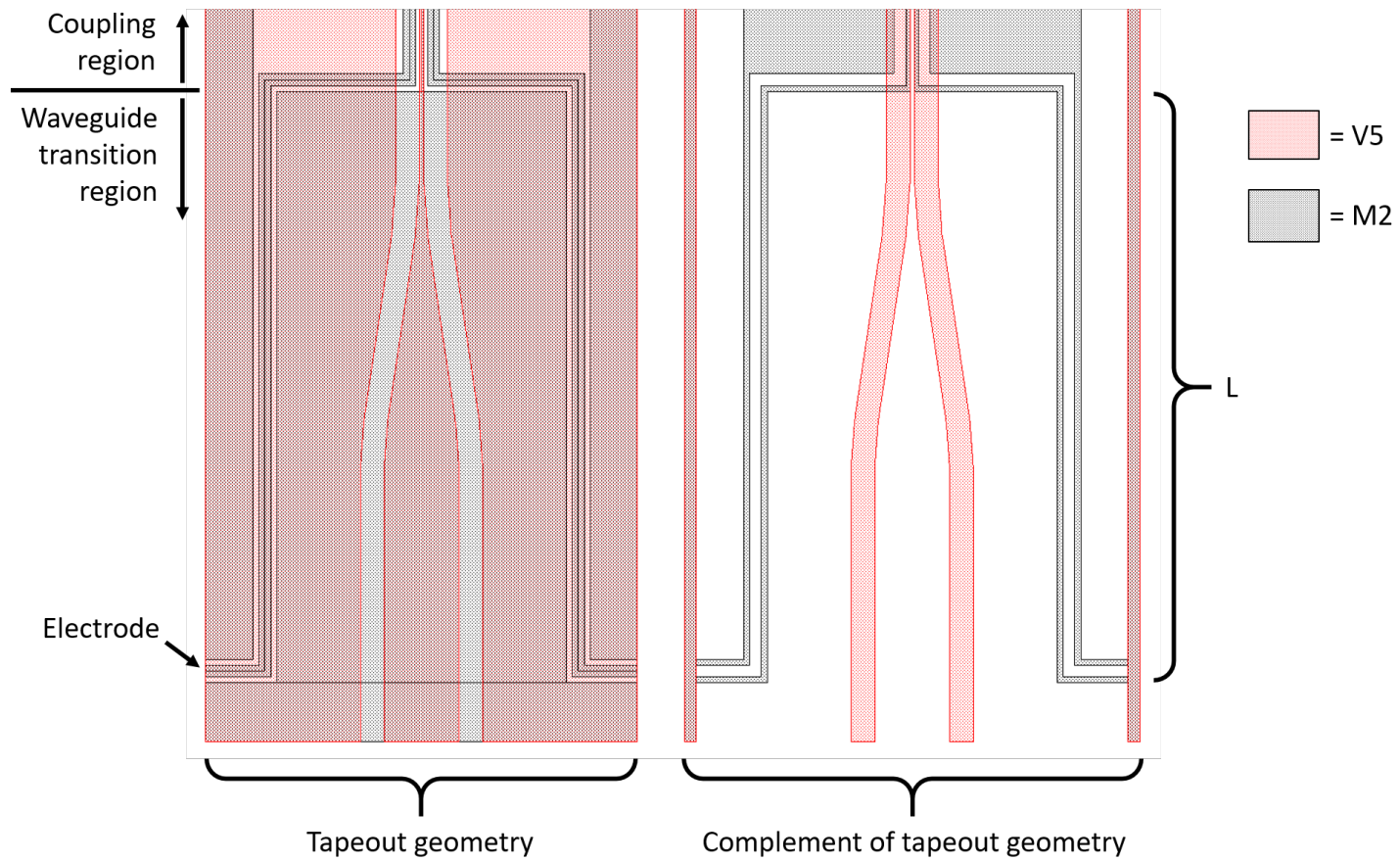
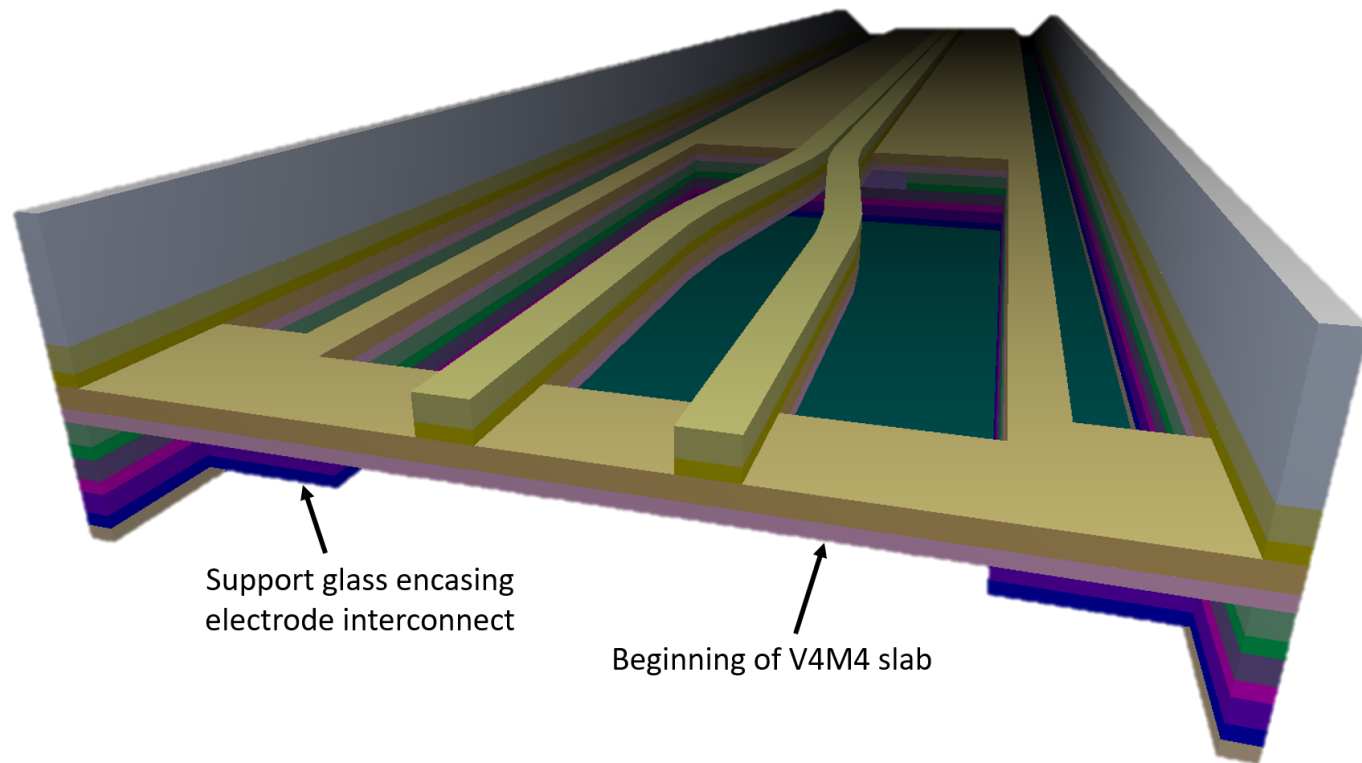


Figure 3.74: The V5 and M2 layers of the MEMS directional coupler. Both the tapeout geometry and its complement are shown to highlight both the metal and dielectric structures.



Support glass encasing
electrode interconnect

Beginning of V4M4 slab

Figure 3.75: A visualization of the oxide structure left after metal etching.

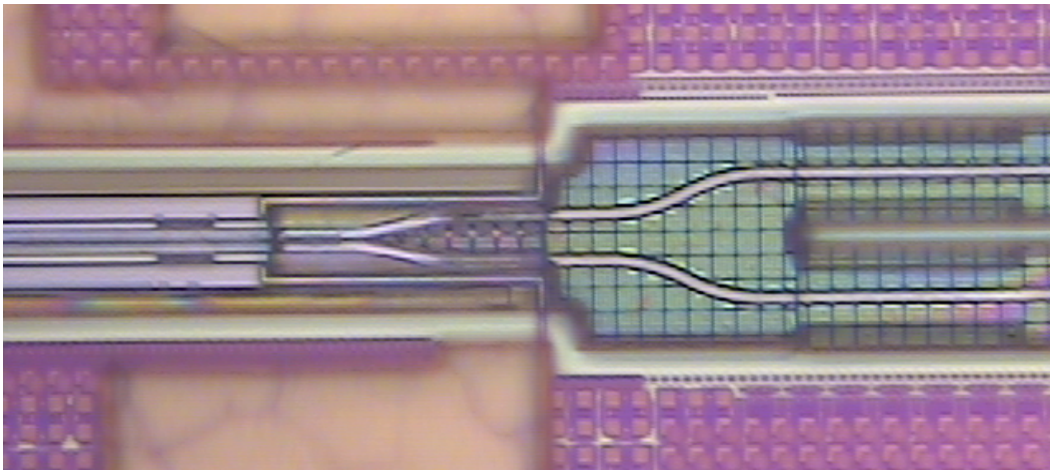


Figure 3.76: Micrograph of the transition region of the MEMS modulator. Note that the waffling has not been removed from the waveguides.

The MEMS directional coupler works by switching between coupling and double coupling, which are defined visually in Figure 3.77. TO4 used three different classes of couplers, each distinguished by its coupling length, which are tabulated in Table 3.1. The 260 nm gap of class 1 couplers is based on the approximate minimum features size of the BEOL, based on via dimensions. The 350 nm gap of the class 2 and 3 couplers was chosen merely to serve as a more conservative alternative. The small difference in coupling length of classes 2 and 3 was chosen due to layout space constraints.

The coupling and double coupling lengths of each coupler are determined based on FDTD simulations, and plotted against gap size in Figures 3.78, 3.80, and 3.82. On each of these plots, a gray dotted line denotes the physical length of the coupling region. The switching mechanism requires changing the gap size between the gap required for coupling and the gap required for double coupling. This gap change is shown in the figures, as well as tabulated in Table 3.1.

In order to remain within the output voltage range of an on-chip voltage multiplier, the voltage required to achieve the necessary gap change should be less than 20 V. Figures 3.79, 3.81, and 3.83 show the distance the electrodes can move for a given applied voltage. Note that these plots curve sharply upwards at the higher end of the voltage range. This asymptotic voltage corresponds to the pull-in voltage. The voltage range has been specifically chosen using the support length parameter, L in Equation 3.9 above, in order to meet the 20 V requirement.

Unfortunately, the MEMS modulators suffered from two problems in post-processing.

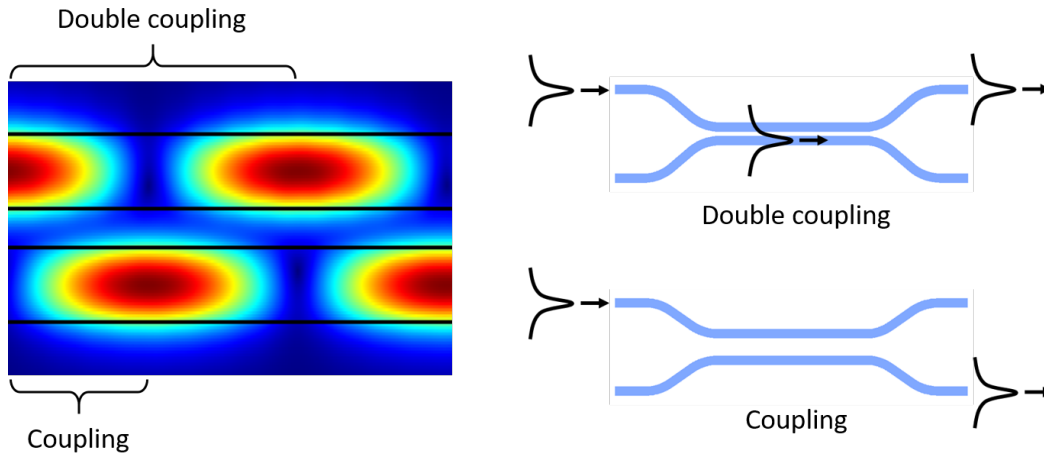


Figure 3.77: Cross section of the coupling portion of the MEMS directional coupler.

Table 3.1: Classes of MEMS directional couplers.

Class	Gap (nm)	Coupling length (mm)	Gap change (nm)
1	260	1	75
2	350	2	90
3	350	1.8	110

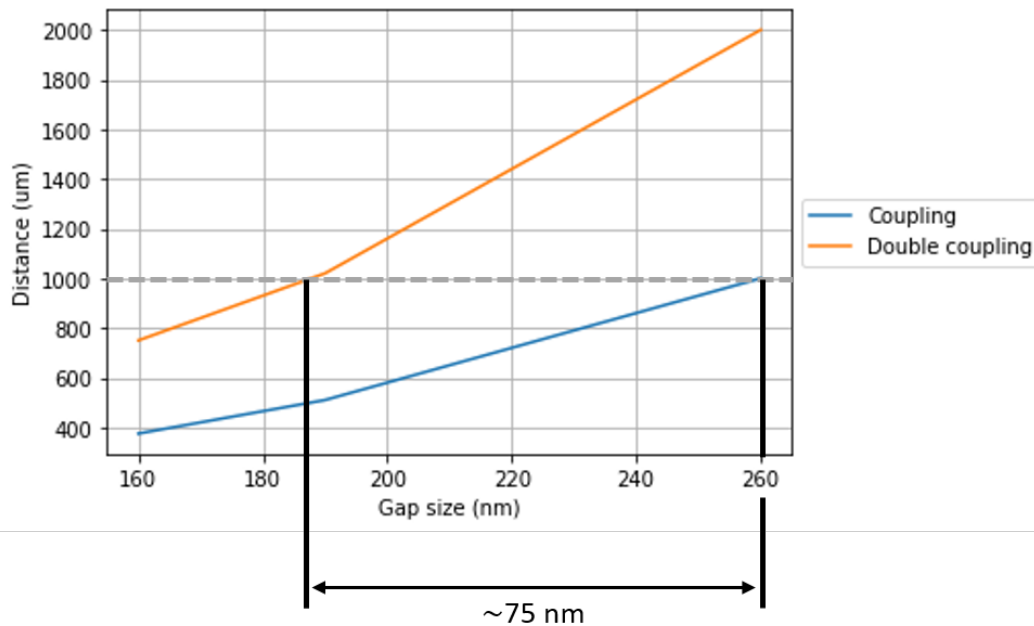


Figure 3.78: Optical coupling lengths of the class 1 MEMS modulator based on gap size.

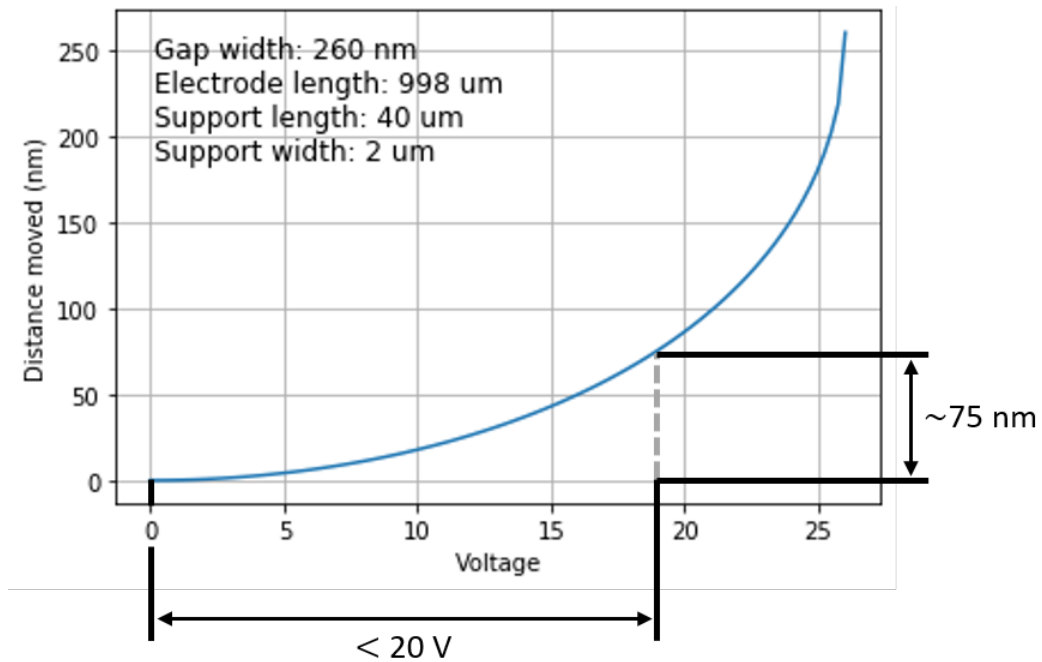


Figure 3.79: Displacement of the electrodes of the class 1 MEMS modulator with applied voltage.

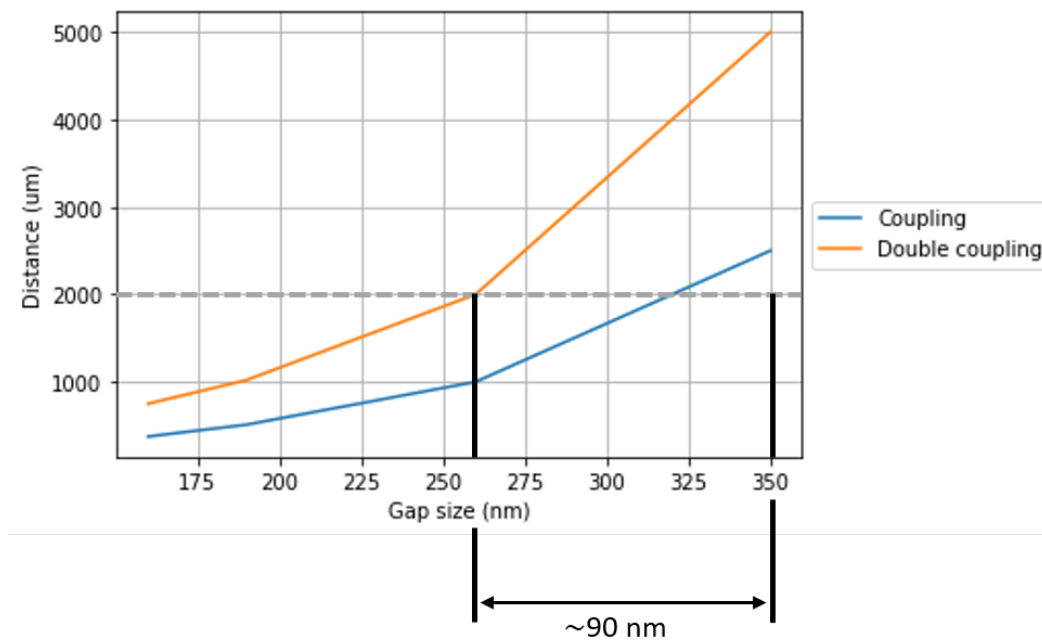


Figure 3.80: Optical coupling lengths of the class 2 MEMS modulator based on gap size.

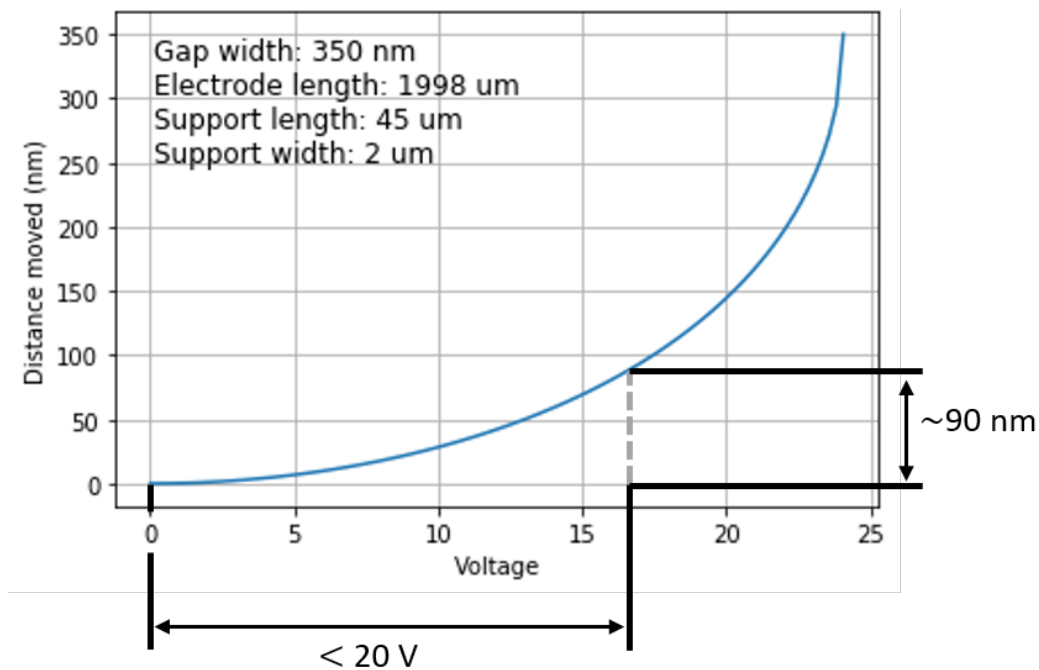


Figure 3.81: Displacement of the electrodes of the class 2 MEMS modulator with applied voltage.

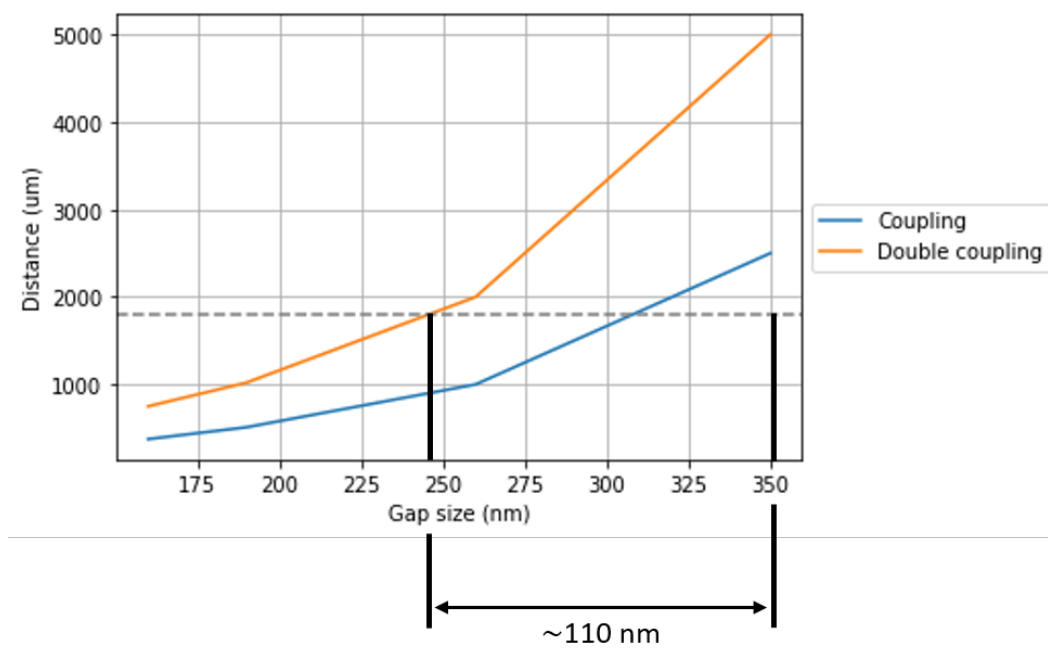


Figure 3.82: Optical coupling lengths of the class 3 MEMS modulator based on gap size.

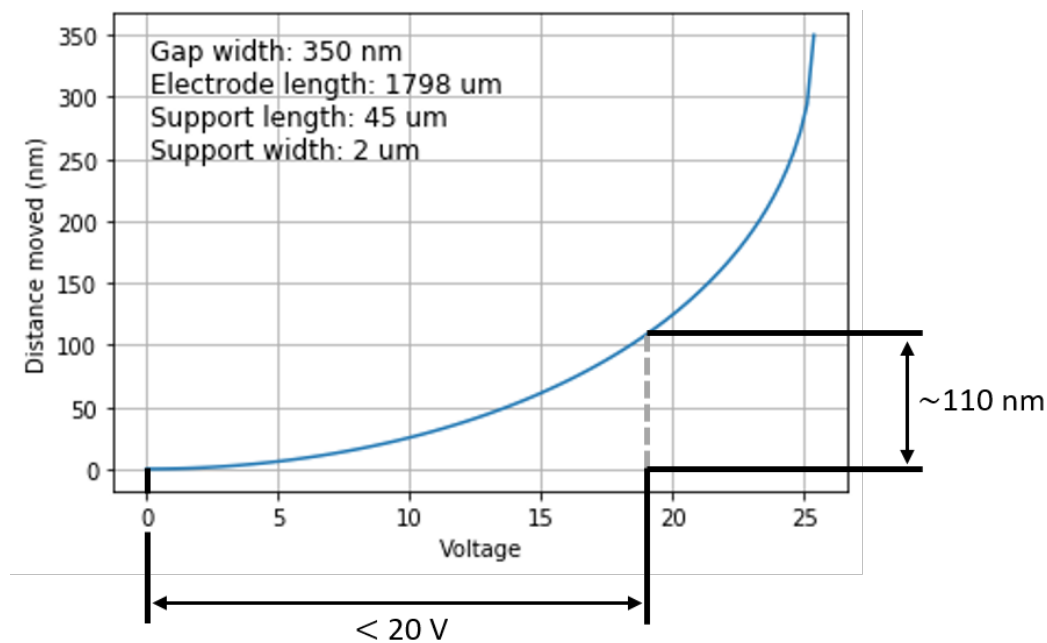


Figure 3.83: Displacement of the electrodes of the class 3 MEMS modulator with applied voltage.

First, the modulators bent upward after etching, as shown in Figure 3.84. This may be due to the compressive stress of the fluorosilicate glass [86]. If the stress is intrinsic rather than thermal, it is likely that the lower layers experience greater compressive stress than upper layers [92], which would lead to upward bending after etching.

The second problem was delamination of the ridge. Figure 3.85 show modulators and photodetectors in the middle of etching. Removal of the waffling also results in removal of the ridge, with no residue of the ridge remaining on the slab. No etching experiment was able to remove the waffling without also removing the ridge for all devices using V5M5-V4M4 waveguides.

At the time of TO4 it was supposed that due to poor adhesion at the M5-V4 interface, which is also an interface between fluorosilicate glass and silicon dioxide in this particular process, the 2 μm ridge width was insufficient. It was supposed that a 3 μm ridge *would* be sufficient since the V5M5-V4M4V3 spiral, which had a 3 μm wide ridge, suffered no adhesion difficulties, demonstrating that the problem could not possibly lie with the M5-V4 interface in particular. TO6 would prove this assumption wrong.

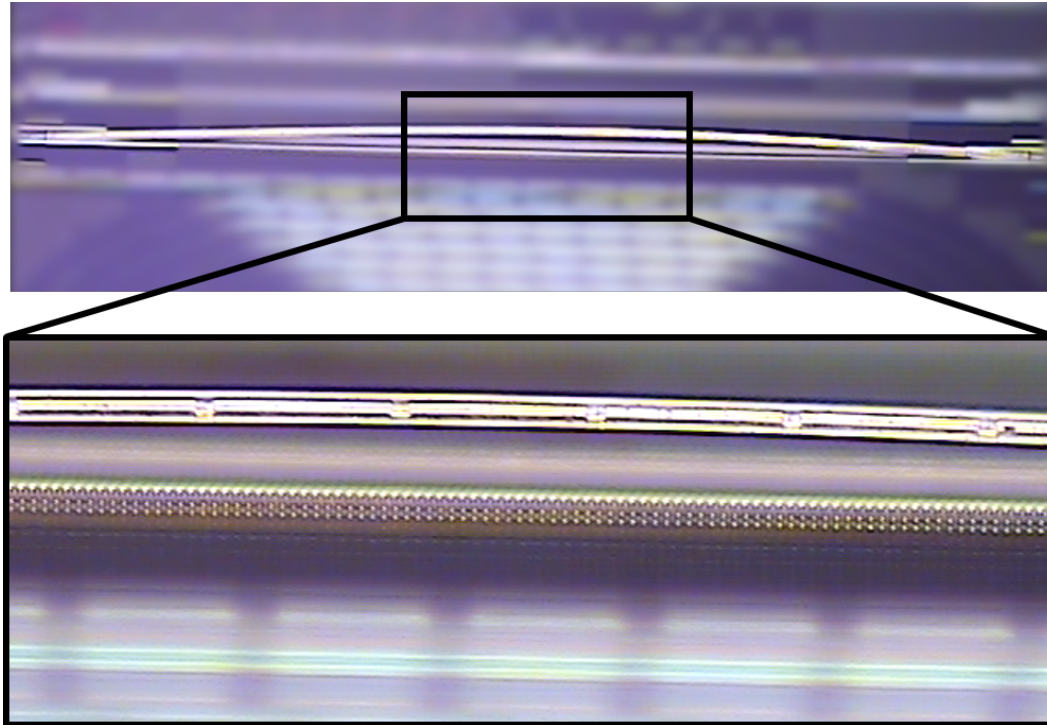


Figure 3.84: Microscope image of a MEMS modulator bent upward.

3.7 TO5

TO5 was the fifth chip taped out in the 180nm process. It is not in the purview of this thesis.

3.8 TO6

Note: The work described in this section was performed in collaboration with Samir Nooshabadi and Debjit Sarkar.

The sixth chip taped out in the 180nm process (Fig. 3.86) was intended to incorporate all of the lessons learned from previous tapeouts. The main focus was on a coherent receiver, which detects the amplitude and phase of an incoming optical signal. After the signal is coupled onto the chip, it is split in half. The system requires that an LO is also coupled onto the chip, which is split in directional coupler 1, the outputs of which are 90° out of phase. Each LO then enters one input of directional couplers 2 and 3, while each half of the signal enters the other input of directional couplers 2 and 3. The signals at the outputs of directional couplers 2 and 3 are 180° out of phase and so can be conveniently amplified by a differential amplifier chain.

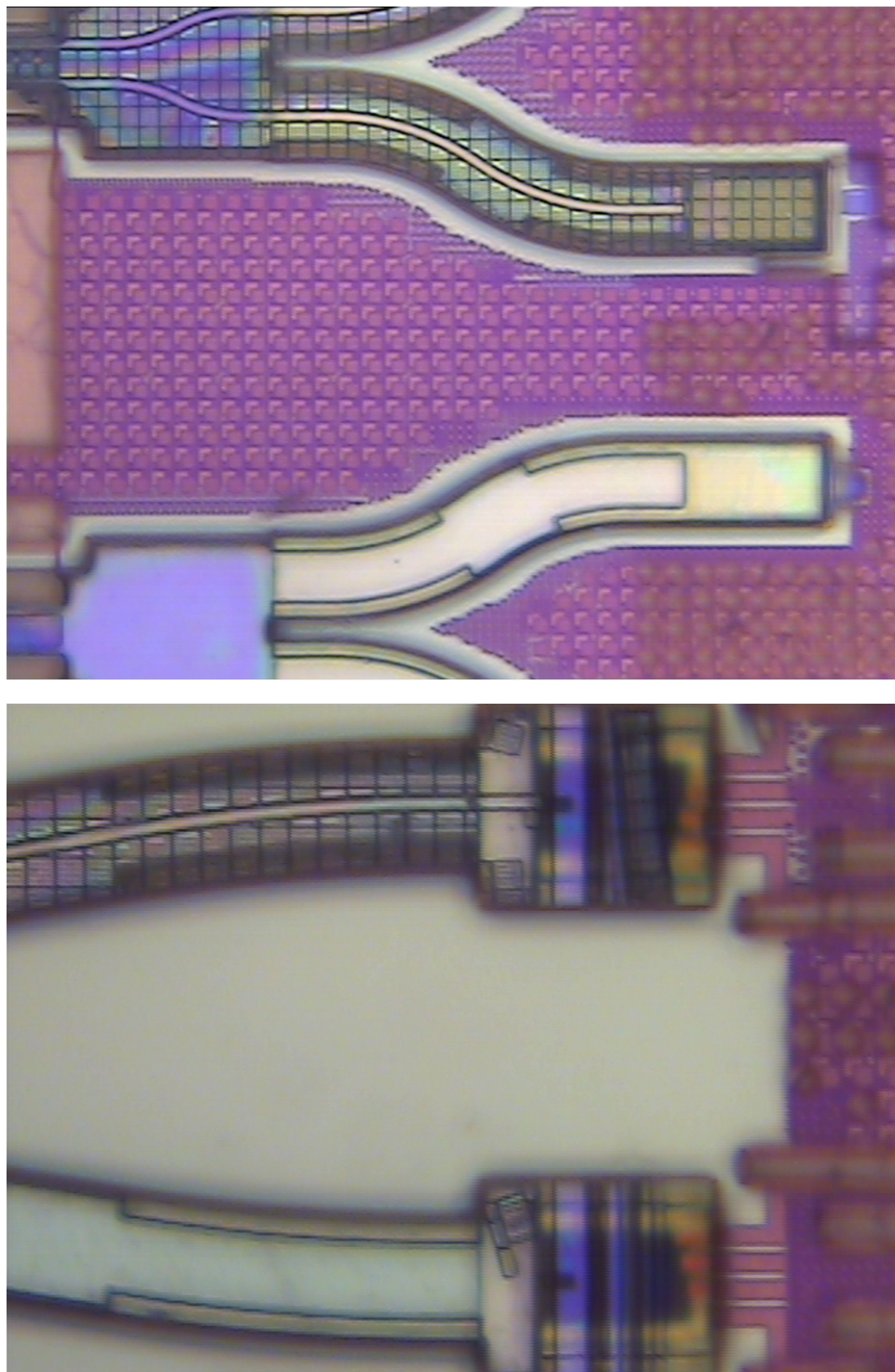


Figure 3.85: Demonstration of waveguide ridges removed with the surrounding waffling, both on modulators (top) and photodetectors (bottom).

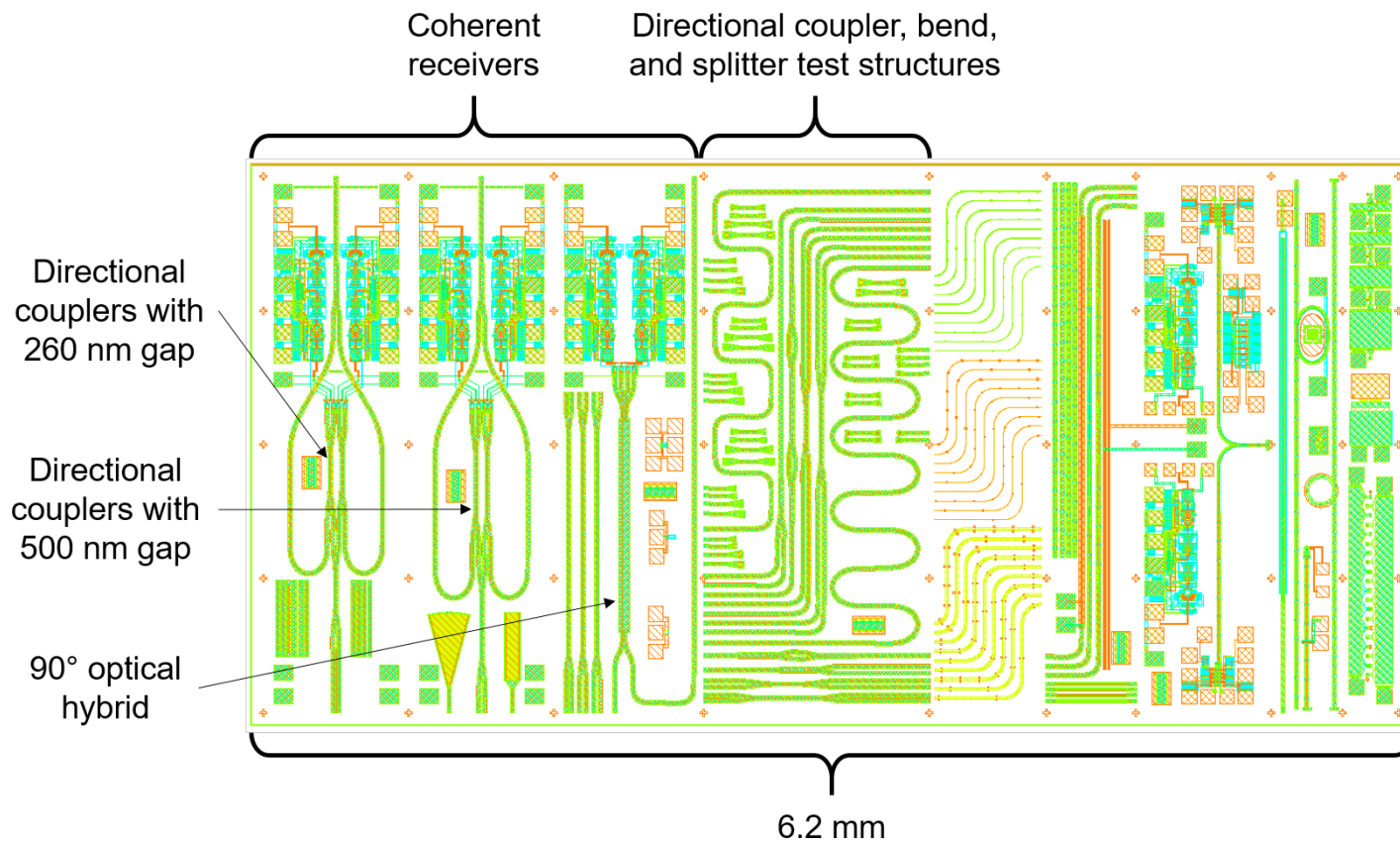


Figure 3.86: Layout of TO6.

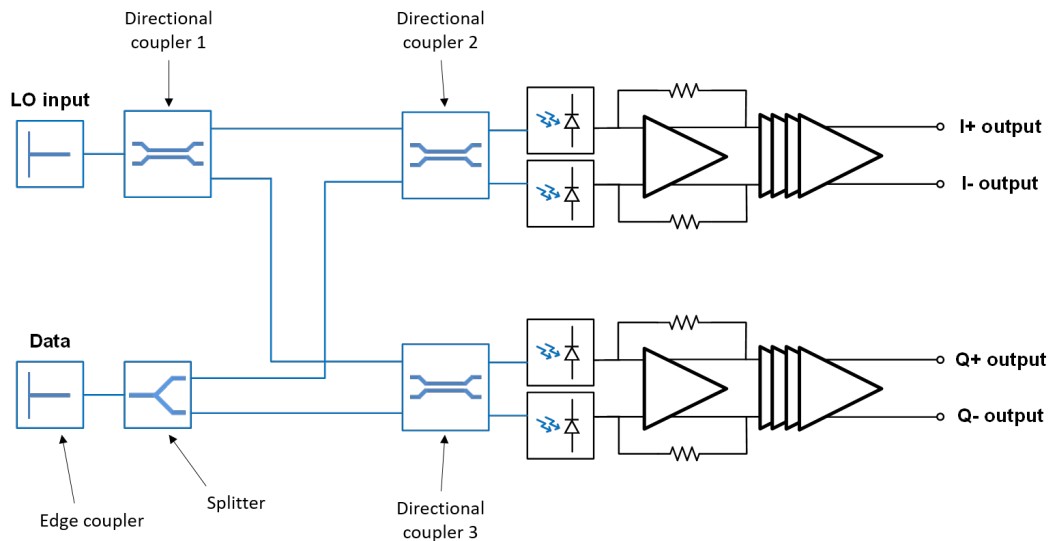


Figure 3.87: System diagram of the proposed coherent receiver.

The first item to be addressed in the design of this system is wavelength, which is considered in greater depth in the following chapter. Succinctly, shorter wavelengths are absorbed better by CMOS photodiodes, but optical amplifiers are only available down to a wavelength of 780 nm. Therefore 780 nm was chosen as the design wavelength.

The first device to be optimized for this system is the waveguide itself. Single-mode waveguides are essentially a requirement for phase sensitive applications, since the fundamental mode and higher order modes can become π out of phase in distances of only tens of micron. However, the single-mode waveguide previously explored in this project suffers from enormous bend loss at reasonable bend radii 3.44. Although a single-sided slab may be used to decrease losses, and an Euler single-sided slab can be used prevent coupling to higher order modes through the bend, a bend radius of at least a couple hundred micron is required to achieve close to full transmission through such a bend.

The V5M5-V4M4 waveguide geometry provides a promising alternative, given its superior loss characteristics for Euler single-sided bends (see Fig. 2.29). Figure 3.88 shows that the waveguide is also single-mode for ridge widths up to about $2.5 \mu\text{m}$, although the ridge width choice must also take adhesion into account. Consider Figure 3.89. On the top is the V5M5-V4M4V3 waveguide geometry used for the spiral in TO4, which did not suffer from ridge removal. On the bottom is the V5M5-V4M4 waveguide geometry used for the MEMS modulators, which did

suffer from ridge removal during post-processing. The differences are the extra slab layer of the former waveguide and the smaller ridge width of the latter waveguide. It was considered reasonable to assume that the additional V3 slab layer of the V5M5-V4M4V3 waveguide could not affect the adhesion of the M5-V4 interface, given that it is two layers away from the interface. Therefore, it was concluded that the reason for the ridge removal of the V5M5-V4M4 waveguide was that the reduced width provided insufficient adhesion (unfortunately, this hypothesis proved incorrect, which will be discussed below). The V5M5-V4M4V3 waveguide provided evidence that a $3\ \mu\text{m}$ wide ridge did provide sufficient adhesion, since it used the same V5M5 ridge. Although a $3\ \mu\text{m}$ wide ridge pushes the V5M5-V4M4 waveguide into the multimode regime, as shown in Figure 3.88, the TE1 and TM1 modes are poorly confined and will experience high losses in bends, rendering the V5M5-V4M4 waveguide with a $3\ \mu\text{m}$ wide ridge effectively single-mode. Single-mode behavior could also be obtained with a $2.5\ \mu\text{m}$ wide ridge, but given the uncertainty of ridge adhesion at this ridge width, as well as the quasi-single-mode behavior at $3\ \mu\text{m}$, it was concluded that the risks of a $2.5\ \mu\text{m}$ wide ridge outweighed the benefits, and a $3\ \mu\text{m}$ wide ridge was chosen for TO6.

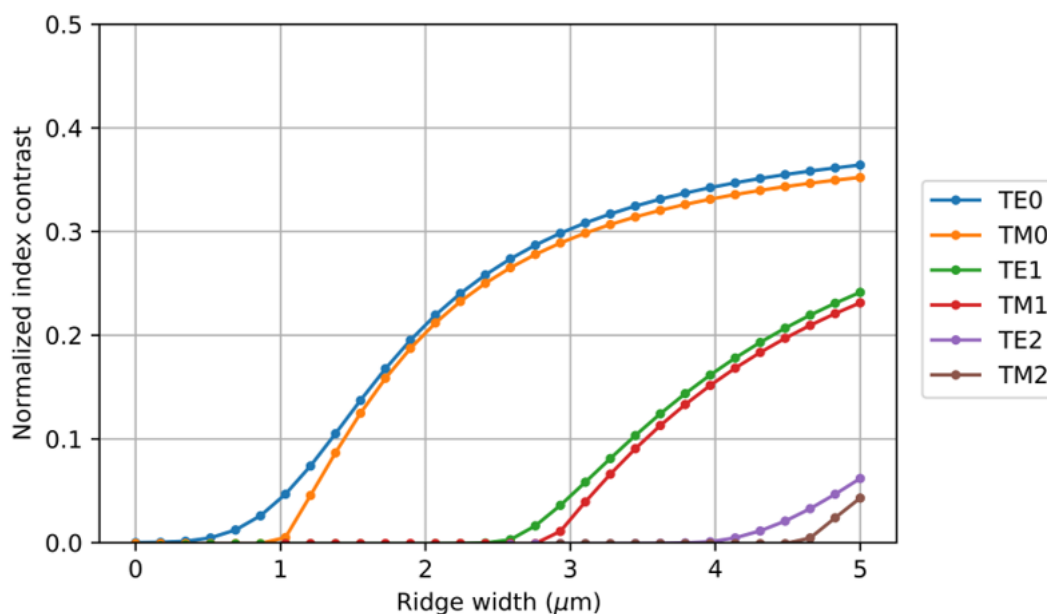


Figure 3.88: Mode confinement of the V5M5-V4M4 waveguide at 780 nm.

The modes of the V5M5-V4M4 waveguide with a ridge width of $3\ \mu\text{m}$ at a wavelength of 780 nm are shown in Figure 3.90. The fundamental TE and TM modes at this

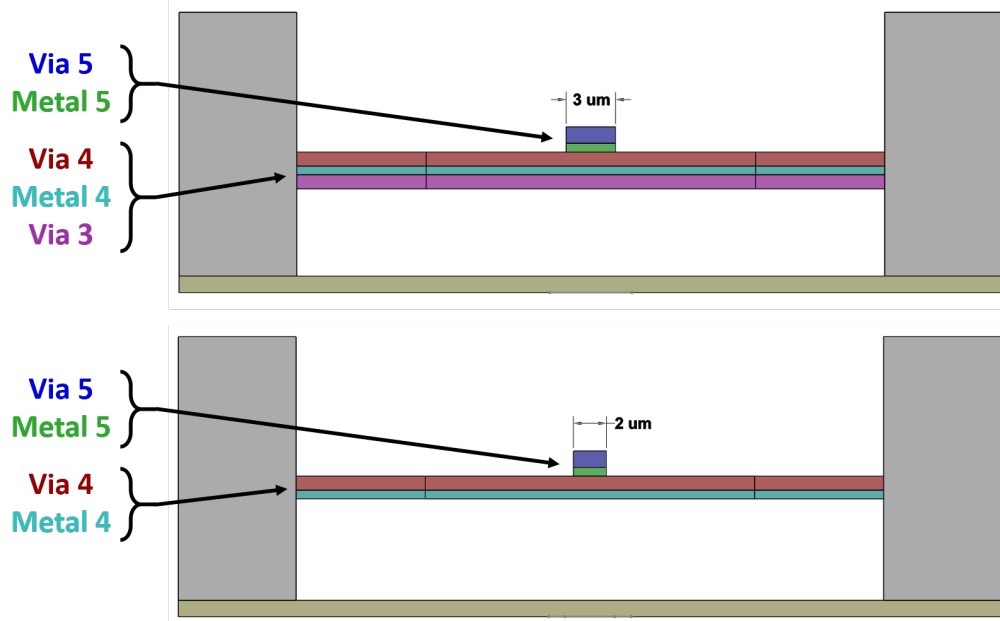


Figure 3.89: Comparison of waveguide geometries. Top: Cross section of the V5M5-V4M4V3 waveguide used in the spiral of TO4, which had adequate ridge adhesion. Bottom: Cross section of the V5M5-V4M4 waveguide used in the MEMS modulators of TO4, which had insufficient ridge adhesion.

ridge width support low loss fiber coupling, as shown in Figure 3.91, when a lensed 780 nm fiber with a spot size of $2 \mu\text{m}$ is used.

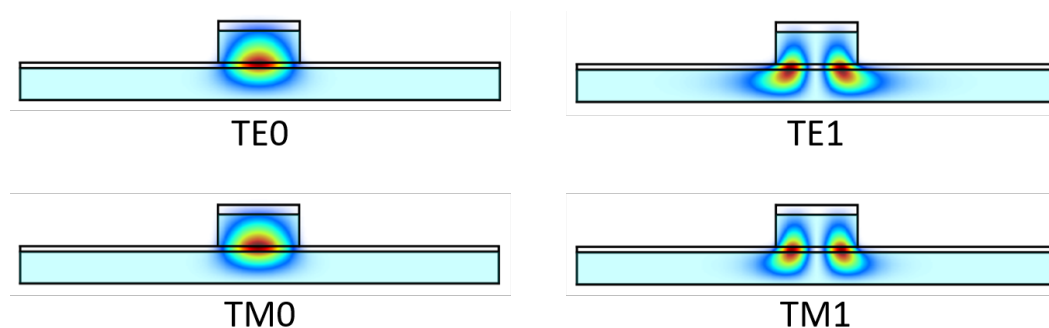


Figure 3.90: Modes of the V5M5-V4M4 waveguide with a ridge width of $3 \mu\text{m}$ at 780 nm.

The directional coupler uses a shared slab to reduce coupling lengths and improve mechanical strength (Fig. 3.92), which means that the slabs of the input waveguides merge and a gap exists only between the ridges. Two variants were simulated with the full dielectric stack, one with a 500 nm ridge gap and the other with a 260 nm

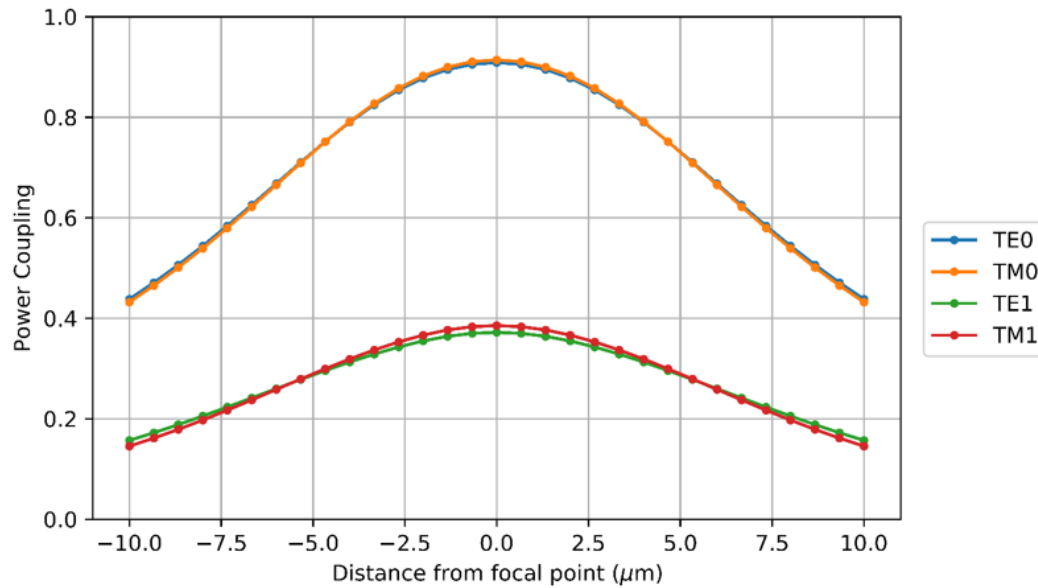


Figure 3.91: Simulated fiber coupling efficiency of a 780 nm lensed fiber with a spot size of $2\ \mu\text{m}$ coupling to the confined modes of a V5M5-V4M4 waveguide.

ridge gap. The former was simulated to have an insertion loss of 0.079 dB while the latter had a simulated insertion loss of 0.044 dB.

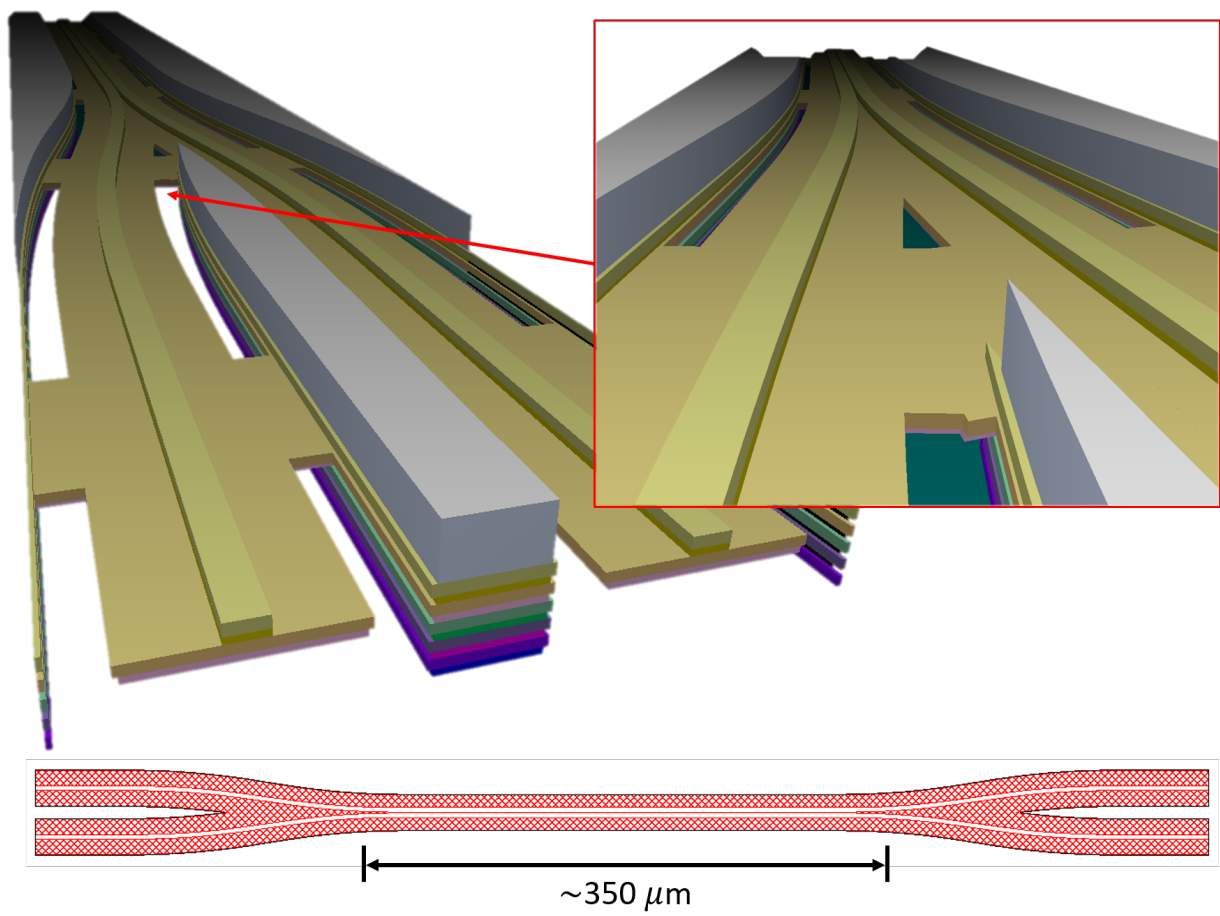


Figure 3.92: A visualization of the V5M5-V4M4 directional coupler, with the inset image showing a close view of the shared slab.

The splitter was of the MMI variety, and was simulated with the full dielectric stack to have an insertion loss of 0.035 dB. Figure 3.93 shows the dielectric stack used in the EME simulation and a sweep of the MMI length, with the optimal transmission occurring at a length of 138 μm .

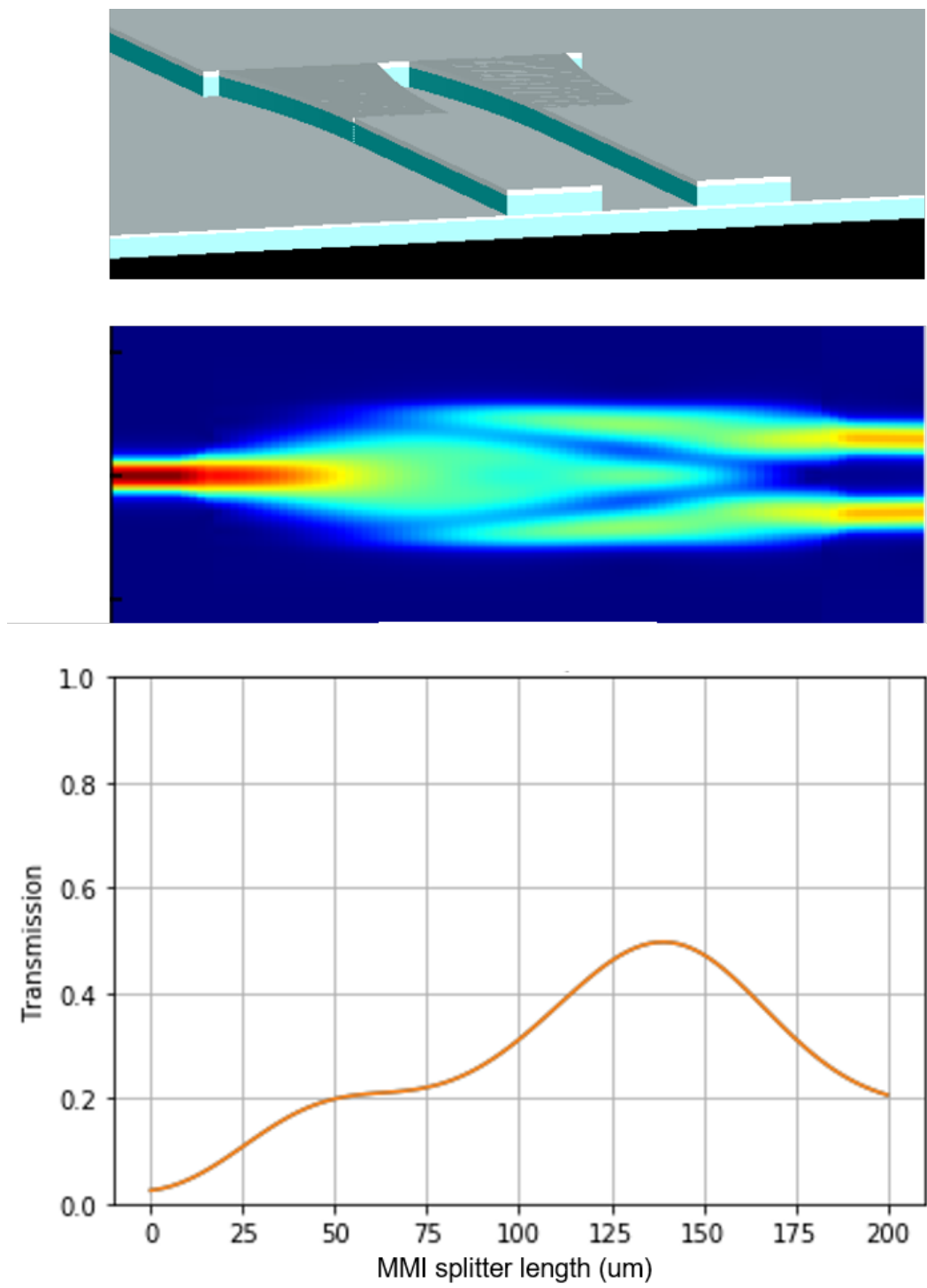


Figure 3.93: MMI splitter optimization. Top: A view of the EME simulation setup using the full dielectric stack. Middle: EME simulation of the MMI splitter at its optimal length of $138 \mu\text{m}$. Bottom: Plot of the EME optimization of the MMI splitter.

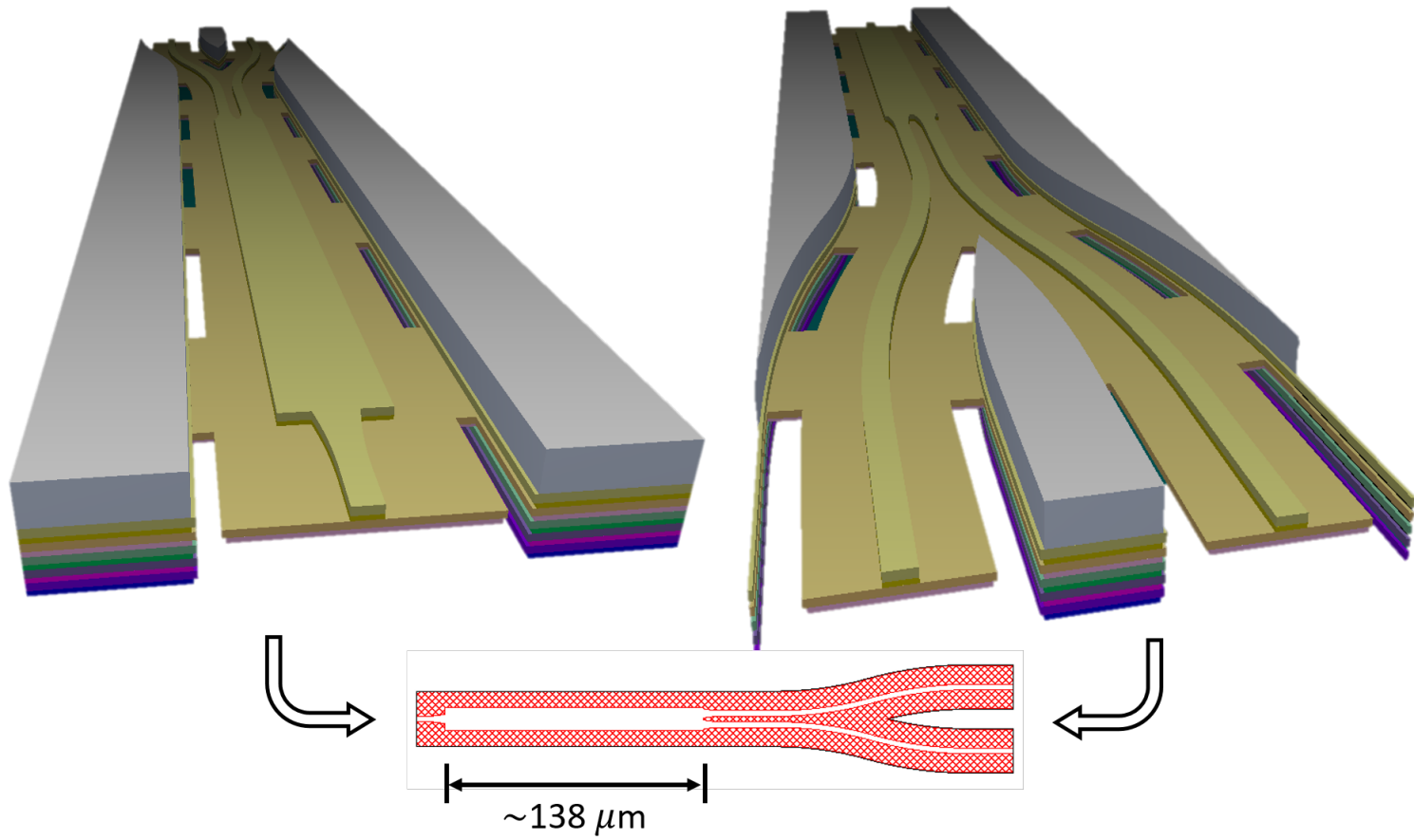


Figure 3.94: A visualization of the V5M5-V4M4 MMI splitter.

The directional coupler based coherent receiver shown in Figure 3.87 suffers from the drawback that there are currently no phase shifters in the subtractive photonics platform. As a result, path length mismatches could cause a relative phase shift between the outputs of directional coupler 1, which need to be in quadrature for the coherent receiver to work. To reduce the risk of path length mismatch, a coherent receiver with an MMI 90° hybrid was also included in the tapeout, replacing three directional couplers and the MMI splitter with one block (Fig. 3.95).

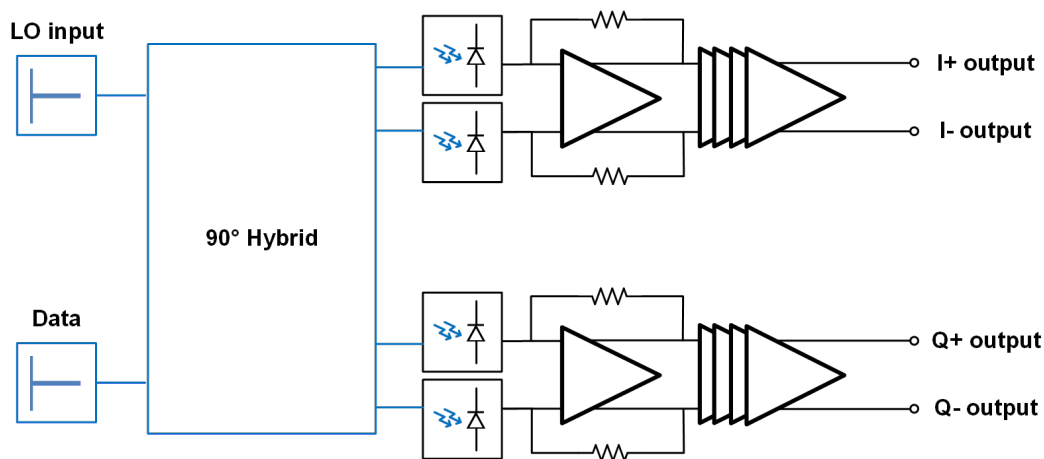


Figure 3.95: System diagram of the coherent receiver using an MMI hybrid.

The MMI hybrid was simulated using the full dielectric stack in an EME simulator. The output of the hybrid in the simulation setup is shown in Figure 3.96. Figure 3.96 also shows cross sections of the MMI hybrid when illuminated at both input ports. The length of the hybrid was determined using an EME sweep, as shown in the plots in the bottom left of Figure 3.96, with a final optimum at a length of 1.11 mm. The plots in the lower right of Figure 3.96 demonstrate that the use of 25 modes in the EME simulation was sufficient for an accurate simulation.

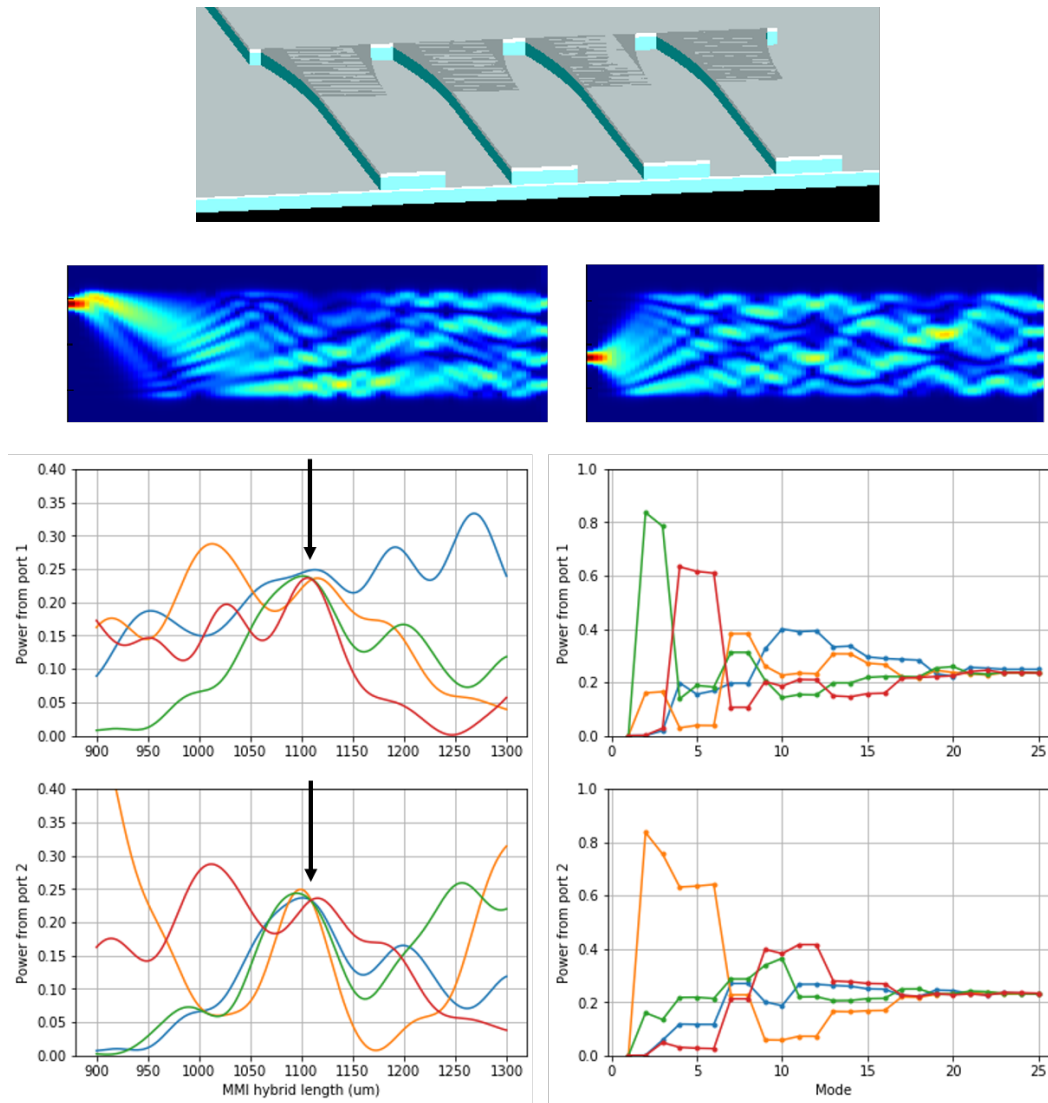


Figure 3.96: MMI hybrid optimization. Top: A view of the EME simulation setup using the full dielectric stack. Middle: EME simulations of the MMI hybrid at its optimal length of 1.11 mm. Bottom left: Plot of the EME optimization of the MMI hybrid length. Bottom right: Mode convergence of the EME simulation at the optimal length of 1.11 mm.

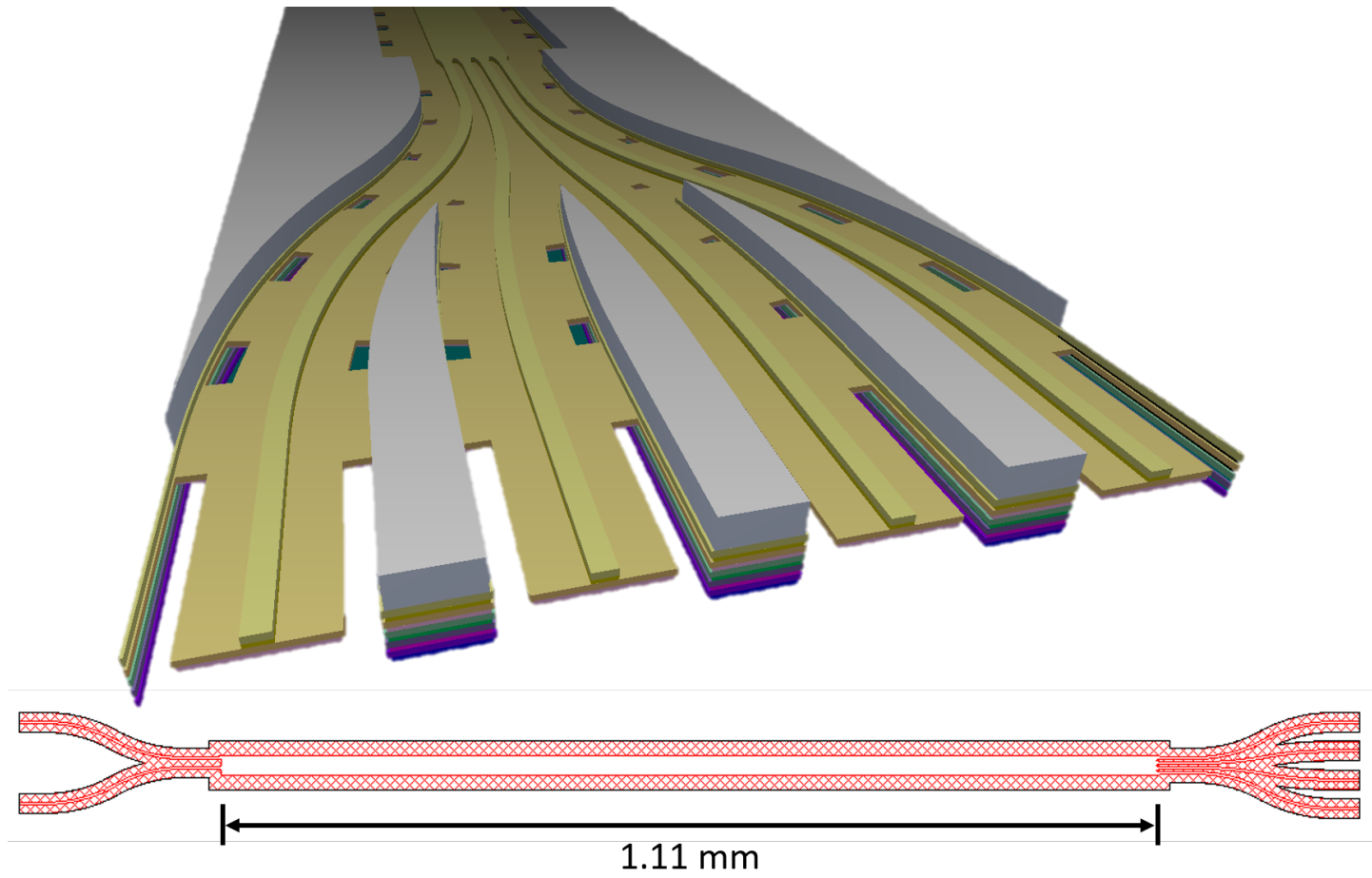


Figure 3.97: A visualization of the quadrature output of the V5M5-V4M4 MMI hybrid.

Unfortunately, the conclusion reached above, that a $3\ \mu\text{m}$ wide ridge on a V5M5-V4M4 waveguide would ensure sufficient adhesion, turned out to be incorrect. As Figure 3.98 shows, the ridges on the waveguides were removed during post-processing. Specifically, when the $\text{H}_2\text{O}_2/\text{EDTA}$ solution was used to remove the barrier layer, the ridge was also removed. This removal occurs even when the $\text{H}_2\text{O}_2/\text{EDTA}$ solution is kept at room temperature, and the chip is imaged without removing it from the etchant, as shown in Figure 3.99, which eliminates any possibility of ridge damage due to handling after the etching. To give assurance that these etching problems are not due to design, Figure 3.100 shows a waveguide region as it progress through the stages of etching. The ridges are present before etching and survive the aluminum etching, but are removed before the barrier layer can be etched. Given these facts, the question arises as to what could cause the ridge to be removed if the width was not a factor.

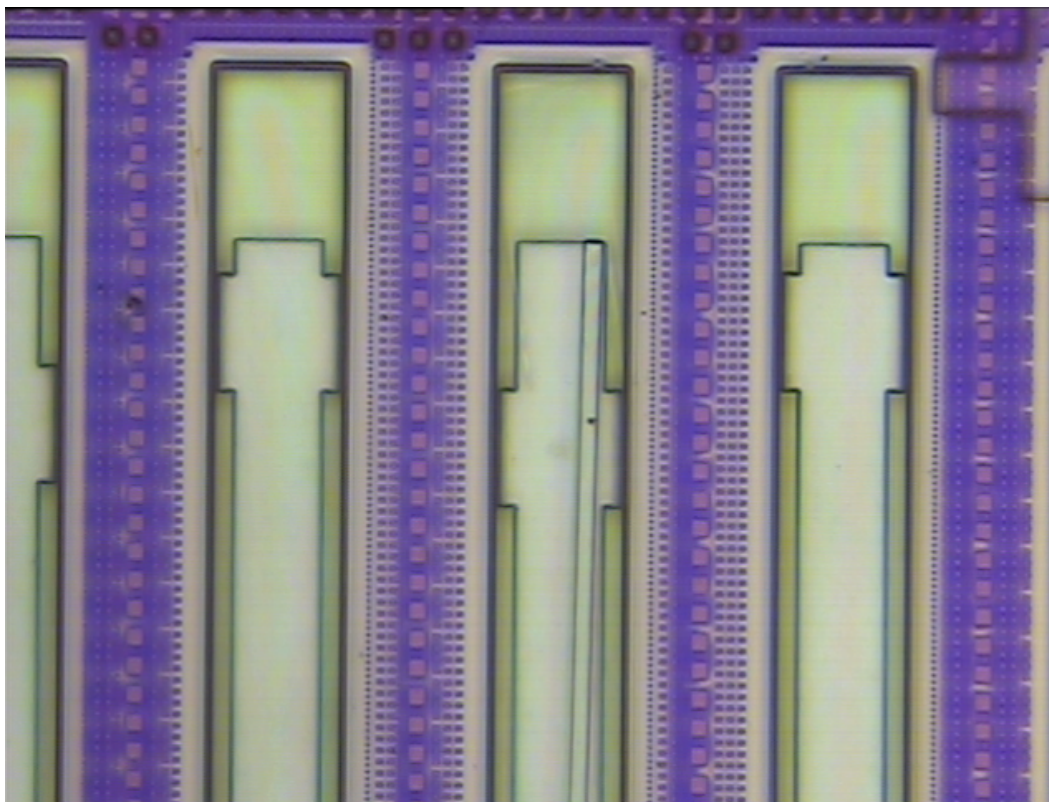


Figure 3.98: Microscope image of V5M5-V4M4 waveguides after barrier layer etching. Most of the waveguides show no evidence of a ridge, while the one remaining ridge has slipped from its designed position.

Consider again a comparison of waveguide cross sections, this time shown in Figure 3.101. The top waveguide is the V5M5-V4M4V3 waveguide used in the

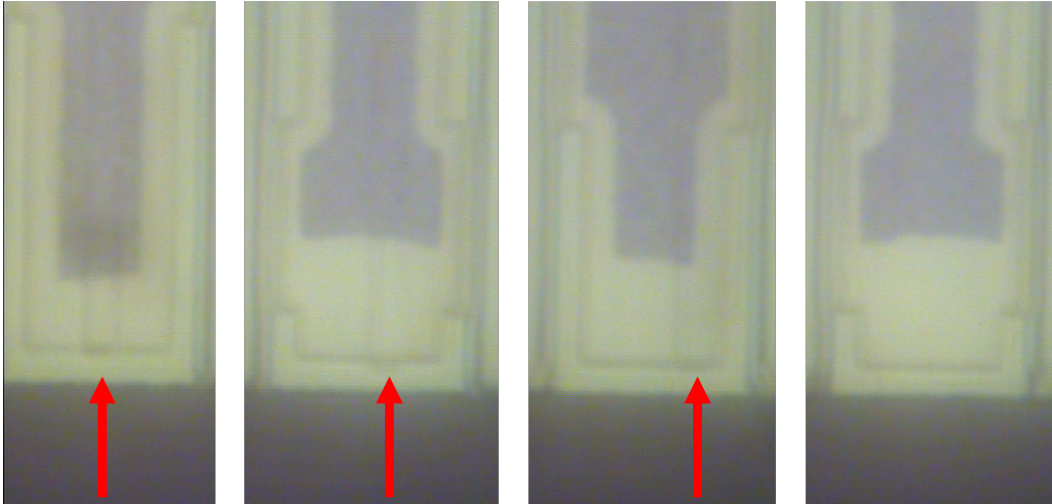


Figure 3.99: Microscope images of waveguides submerged in etchant. This chip underwent 3.5 hours of AETA etching at 80°C to remove the aluminum, and then 3 hours of H₂O₂/EDTA etching at room temperature with no stirring. These images show several waveguides on this chip while they are still in the H₂O₂/EDTA. The red arrows indicate the position of the remaining ridges.

single mode spiral of TO4. As mentioned above, this waveguide suffered from no ridge peeling with a 3 μm wide ridge. The bottom waveguide is a V5M5V4-M4V3 waveguide fabricated on TO6. It also demonstrated no ridge peeling with a 3 μm wide ridge. The middle waveguide is the V5M5-V4M4 waveguide used in TO6 that suffered from ridge removal. Since the V5M5-V4M4 geometry combines the two layer ridge of the V5M5-V4M4V3 waveguide with the two layer slab of the V5M5V4-M4V3 waveguide, and since both of those waveguides survived etching perfectly intact, the reason for the ridge removal of the V5M5-V4M4 geometry cannot lie with either of these attributes independently. Moreover, the V5M5-V4M4 waveguide shares with the V5M5-V4M4V3 waveguide the M5-V4 interface, which rules out any cause due to the interface in particular. And as mentioned above, all waveguides share the same ridge width, ruling out width as an independent cause. The only remaining causes lie with the unique combination of attributes of the V5M5-V4M4 waveguide. For example, it may be that for the M5-V4 interface in particular, a two layer slab provides sufficient mechanical flexibility to weaken the interface and allow the ridge to peel off during the H₂O₂/EDTA etch. Further experiments are required to test this hypothesis.

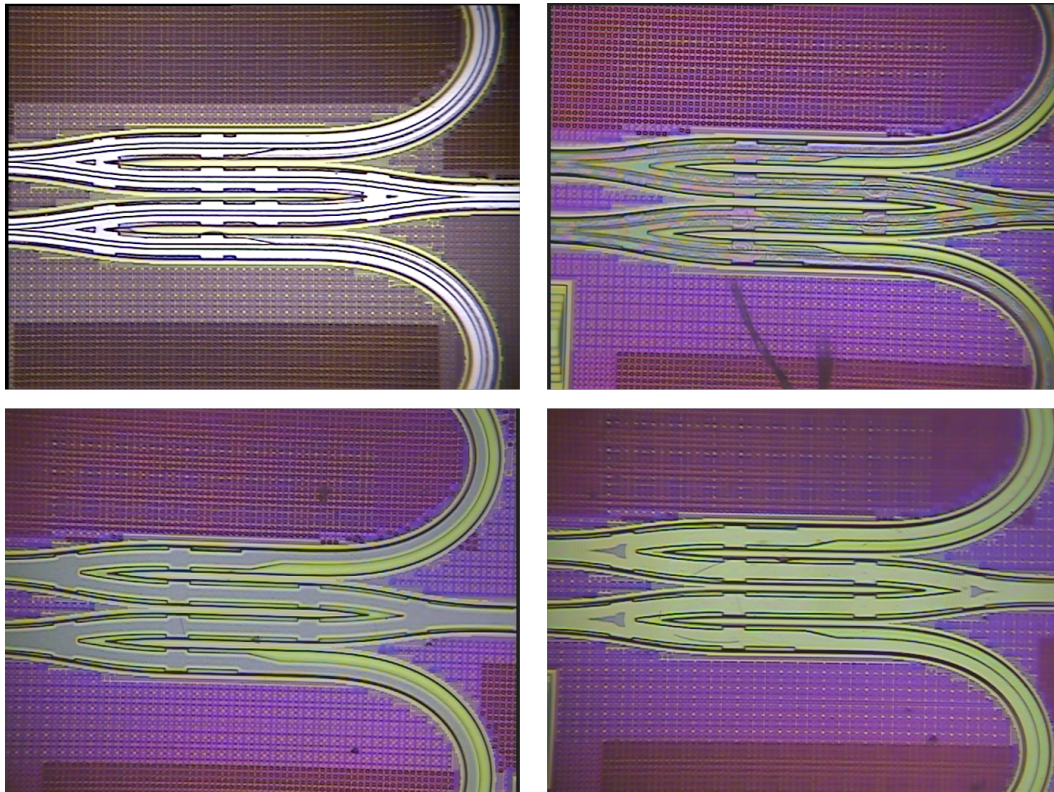


Figure 3.100: Microscope images of a waveguide region throughout the etching process. Upper left (step 1): Before etching. Upper right (step 2): After 9 hours of AETA at 80°C. Lower left (step 3): After 2 hours of H₂O₂/EDTA at room temperature. Lower right (step 4): After 6 hours of H₂O₂/EDTA at room temperature.

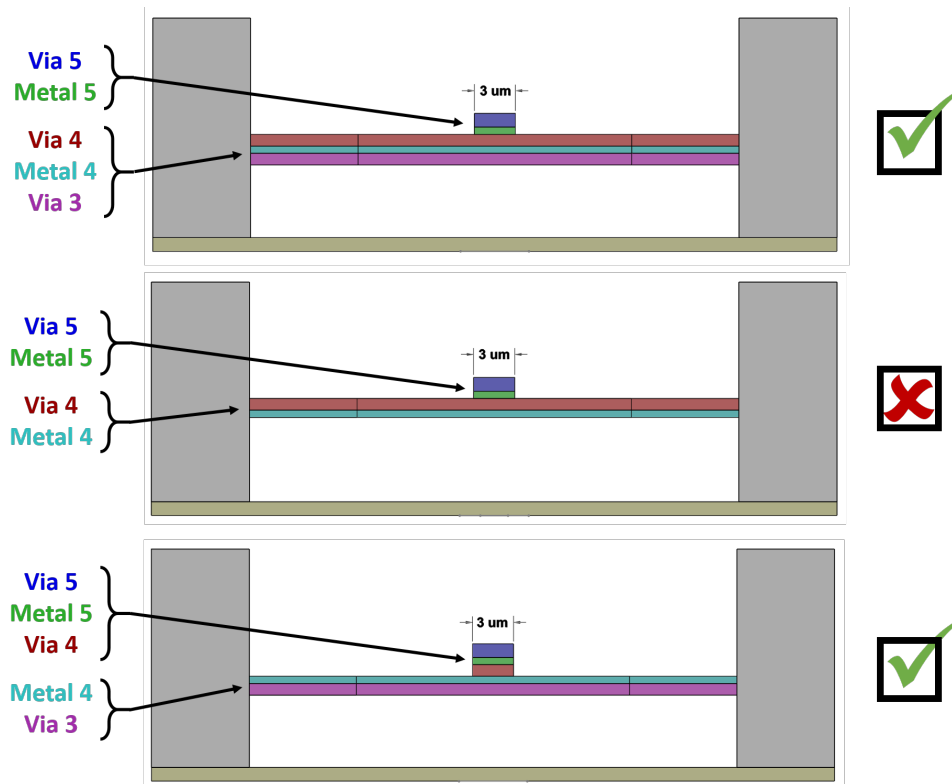


Figure 3.101: Waveguide geometry comparison. Top: V5M5-V4M4V3 waveguide used in TO4 - no ridge problems. Middle: V5M5-V4M4 waveguide used in TO6 - ridge removed during post-processing. Bottom: V5M5V4-M4V3 waveguide used in TO6 - no ridge problems.

PHOTODIODES AND PHOTODIODE COUPLERS

4.1 General Considerations

CMOS photodiodes provide a convenient method of detecting visible light carried in the waveguides. They are well-studied and can be fabricated with zero DRC errors [71]. Light in the near infrared can also be detected in bulk CMOS using Schottky photodiodes [72], but they are not studied in this thesis, under the assumption that the methods applied to CMOS photodiodes also apply to Schottky photodiodes.

Both the 65 nm and the 180 nm processes used in thesis include a deep N-well layer, which adds a p-n junction to the FEOL. The junctions available for photodetection include the P-substrate and the deep N-well (PSUB/DNW), the deep N-well and the P-well (DNW/PW), and the P-well and the highly doped N region (PW/N+). Other possible junctions around an N-well include PSUB/NW and NW/P+, but the analysis and measurements in [71] shows that the PW/N+ junction provides the highest bandwidth, so that is the junction most often used. The doping concentrations are estimated from [93], [94] and shown in Table 4.1.

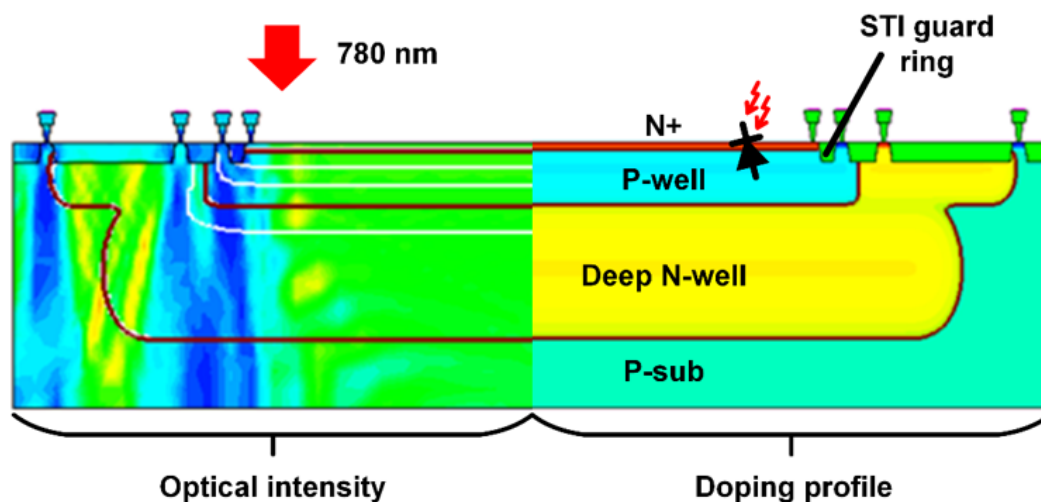


Figure 4.1: TCAD simulation of a bulk CMOS photodiode.

There are a variety of tradeoffs associated with CMOS photodiodes. The bandwidth of a CMOS photodiode can be limited by several mechanisms, including the RC time constant, carrier transit time, diffusion current, and avalanche gain. These bandwidth

Table 4.1: Assumed doping concentrations

Region	Doping concentration (cm ⁻¹)
N+/P+	3×10^{20}
NW/PW*	1×10^{18}
NW/PW/DNW**	1×10^{16}
PSUB	1×10^{15}

*Close to the N+/P+ region

**Close to the bottom of NW/PW or the top/bottom of DNW

limitations trade off with responsivity, and each trade-off is different for each junction. The depletion region width plays a large part in these tradeoffs, and can be calculated for each junction based on the doping concentrations given in Table 4.1.

Following [95], the depletion width of a p-n junction is

$$W = \sqrt{\frac{2\epsilon}{q} \left(\frac{N_A + N_D}{N_A N_D} \right) \psi_{bi}}, \quad (4.1)$$

where ψ_{bi} is the built-in potential and the other symbols have their usual meaning.

For nondegenerate semiconductors, the built-in voltage is

$$\psi_{bi} = \frac{kT}{q} \ln \left(\frac{N_D N_A}{n_i^2} \right). \quad (4.2)$$

For degenerate semiconductors, the built-in potential must be calculated from the Fermi potentials using

$$q\psi_{bi} = E_g - (q\phi_n + q\phi_p). \quad (4.3)$$

An estimate for the Fermi potentials in degenerate semiconductors is given in [95] as

$$q\phi_n \approx -KT \left[\ln \left(\frac{n}{N_C} \right) + 2^{-3/2} \left(\frac{n}{N_C} \right) \right] \quad (4.4)$$

for n-type and

$$q\phi_p \approx -KT \left[\ln \left(\frac{p}{N_V} \right) + 2^{-3/2} \left(\frac{p}{N_V} \right) \right] \quad (4.5)$$

for p-type, where N_C is the effective density of states in the conduction band and N_V is the effective density of states in the valence band. In silicon at room temperature, $N_C = 2.8 \times 10^{19}$ and $N_V = 2.65 \times 10^{19}$.

For nondegenerate semiconductors the Fermi potentials can be calculated as

$$q\phi_n = kT \ln \left(\frac{N_C}{n} \right) \quad (4.6)$$

for n-type and

$$q\phi_p = kT \ln \left(\frac{N_V}{p} \right) \quad (4.7)$$

for p-type, which may be useful for a junction between a degenerate semiconductor and a nondegenerate semiconductor, as is the case for N+/PW CMOS photodiodes.

Using Equations 4.1, 4.2, and 4.3, the depletion region width for various junctions can be calculated. The results for the junctions under consideration in this thesis are tabulated in Table 4.2.

Table 4.2: Depletion region widths

Junction	Depletion region width (nm)
N+/PW	39
PW/DNW	427
DNW/PSUB	959

The absorption of light in silicon can be modeled using the Lambert-Beer's law,

$$I(z) = I_0 e^{-\alpha z}, \quad (4.8)$$

where α is the absorption coefficient of silicon and the light is incident onto the silicon along the z -axis. The penetration depth is the value of z for which the intensity drops to I_0/e , and is tabulated for silicon according to various wavelengths in Table 4.3.

Table 4.3: Absorption coefficients and penetration depths in silicon (from [96]).

Wavelength (nm)	α (μm^{-1})	$1/\alpha$ (μm)
980	0.0065	153.8
850	0.06	16.7
780	0.12	8.3
635	0.38	2.6
565	0.73	1.4
465	3.6	0.28
430	5.7	0.18

The most important wavelengths for the purposes of this thesis are 635 nm, 780 nm, and 850 nm, because of the ease with which sources and modulators can be procured. The shortest wavelength at which an optical amplifier can be procured is 780 nm¹, but 635 nm is still a useful wavelength due to the visibility of light propagation. Consequently, the wavelengths under consideration will be restricted to these three,

¹Based on my search of vendor offerings at the time of writing.

with the understanding that system design will be restricted to 780 nm and 850 nm in order to ensure access to optical amplifiers. The large penetration depths of 8.3 μm and 16.7 μm for 780 and 850 nm, respectively, indicate that the efficiency of CMOS photodiodes will be low.

The efficiency is often quantified using the quantum efficiency, defined as [95]

$$\eta = \frac{I_{ph}/q}{P_{opt}/h\nu}, \quad (4.9)$$

where I_{ph}/q , the photocurrent divided by the charge of each carrier, is the number of carriers generated, and $P_{opt}/h\nu$, the incident optical power divided by the energy of each photon, is the number of incident photons. In other words, the quantum efficiency provides the number of carriers generated per photon. For a p-n junction photodiode, using a one dimensional analysis, the quantum efficiency is found to be [95]

$$\eta = (1 - R) \left[1 - \frac{e^{-\alpha W}}{1 + \alpha L} \right] \quad (4.10)$$

where R is the reflection coefficient, W is the width of the depletion region, and L is the diffusion length. It is assumed that within the depletion region every photon excites a carrier, and dark current is neglected.

The quantities that are measured in experiments and relevant to system design are the incident optical power and the photocurrent, so often the responsivity is used instead of the quantum efficiency. It is defined as

$$\mathcal{R} = \frac{I_{ph}}{P_{opt}} = \frac{\eta q}{h\nu} = \frac{\eta \lambda}{1.24} \quad (4.11)$$

where λ is specified in μm . As a result, the responsivity of a p-n junction is

$$\mathcal{R}_{pn} = \frac{\lambda(1 - R)}{1.24} \left[1 - \frac{e^{-\alpha W}}{1 + \alpha L} \right]. \quad (4.12)$$

To apply this model to CMOS photodiodes, we assume that all of the junctions are one sided, which is accurate for the N+/PW and DNW/PSUB junctions, and a reasonable simplifying assumption for the PW/DNW junction, considering the doping profile given in [93]. Then for the N+/PW junction, the diffusion length of electrons in the p-well is limited by the thickness of the p-well, which is about 800 nm [93]. For the PW/DNW junction, the diffusion length of holes in the deep n-well is limited by the thickness of the deep n-well, which is about 2 μm [93]. For the DNW/PSUB junction, the entire substrate is available for collecting carriers,

so the full 58 μm diffusion length of electrons in the p-substrate [97] can be used. Assuming the depletion region widths tabulated in Table 4.2, and neglecting the reflection coefficient, the responsivity for each of the three junctions is plotted in Figure 4.2. Clearly the DNW/PSUB junction is superior based on responsivity alone, but bandwidth and avalanche gain are additional factors that need to be considered.

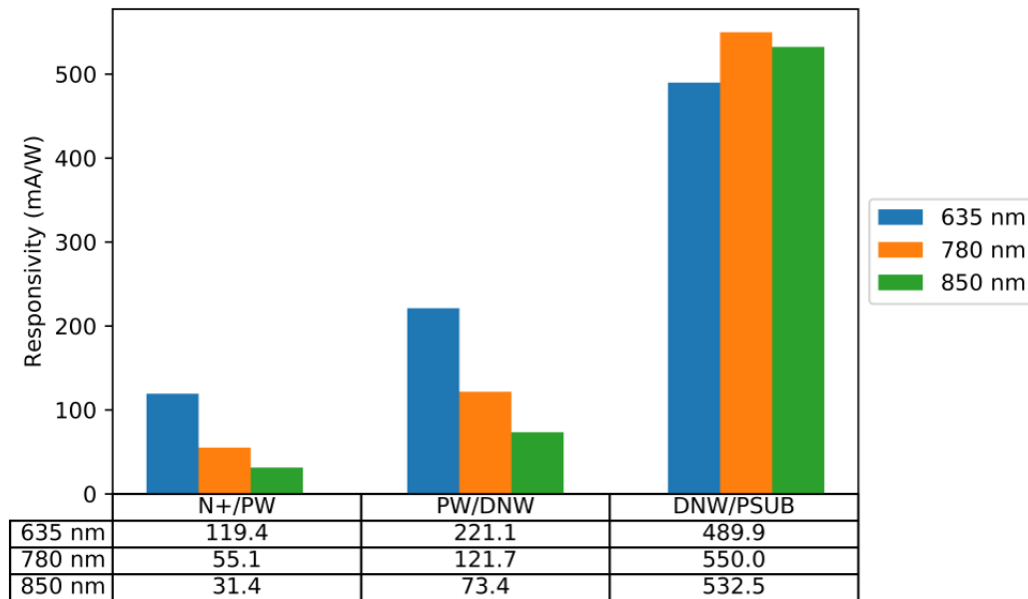


Figure 4.2: Responsivity of each of the three CMOS photodiode junctions across the wavelengths of 635, 780, and 850 nm.

The bandwidth of p-n junction photodiodes is limited by transit time, capacitance, and the delay due to diffusion current. A simple photodiode model can be used to obtain the following estimate for the bandwidth of a p-n junction considering only transit time [95]:

$$f_{3dB-transit} = \frac{2.4v}{2\pi W} \quad (4.13)$$

where v is the carrier velocity and W is the depletion region width. Assuming that holes are the limiting factor, the hole velocity can be written as

$$v_p = \frac{\mu_p \mathcal{E}}{1 + \mu_p \mathcal{E}/v_s} \quad (4.14)$$

where \mathcal{E} is the electric field, $\mu_p = 500$ [$\text{cm}^2/\text{V}\cdot\text{s}$] is the mobility of holes in silicon, and $v_s = 1 \times 10^7$ [cm/s] is the saturation velocity in silicon [95].

The reverse bias junction capacitance can be estimated using the depletion region widths W tabulated in Table 4.2 above. Then the bandwidth is

$$f_{3dB-RC} = \frac{W}{2\pi R\epsilon A}, \quad (4.15)$$

where R is the series resistance, ϵ is the permittivity of silicon, and A is the area of the photodiode.

The bandwidth limitations caused by diffusion current are strongly dependent on the layout of the wells [97]. To provide a general comparison between the bandwidth limitations caused by transit time, capacitance, and diffusion current, however, it is useful to use a bandwidth based on the time required for a carrier to diffuse a distance L [95],

$$\tau_{diff} = \frac{4L^2}{\pi^2 D} \quad (4.16)$$

where D is the diffusion constant. As with the responsivity calculation above, we assume all of the junctions are one-sided and that the diffusion lengths in the wells are limited by the thicknesses of the wells. We assume that for the N+/PW and PW/DNW junctions, the bandwidth is limited by the time required for minority carriers to diffuse from the bottom of the PW and DNW, respectively. If we take τ_{diff} as a time constant, the bandwidth is [98]

$$f_{3dB-diff} = \frac{1}{2\pi\tau_{diff}}. \quad (4.17)$$

The bandwidths for each junction due to transit time, capacitance, and diffusion current are plotted in Figure 4.3. The transit time bandwidth assumes a bias of 1.8 V. The RC bandwidth assumes 50Ω resistance and a 50 μm × 50 μm photodiode area.

Figure 4.3 demonstrates several key points. Transit time is far from the limiting factor for all three junctions and so can be ignored. The RC limit could be an issue for high speed designs, especially for the smaller junctions, but the series resistance and photodiode area can be dropped to push the bandwidth up. The method of subtractive photonics greatly eases the cost of dropping photodiode area, as the fiber alignment is fully separated from the photodiode coupling. The significant limiting factor is the diffusion current. In particular, the DNW/PSUB junction exhibits a modeled bandwidth of just 430 kHz. Other models predict bandwidths of 2.23 MHz [97], "a whopping 5 MHz" [99], 10 MHz [100], and "in the MHz range" [71]. [98] notes that the typical cutoff for diffusion current is 1 MHz. Nevertheless, the high responsivity of the DNW/PSUB junction is enticing, so a variety of methods

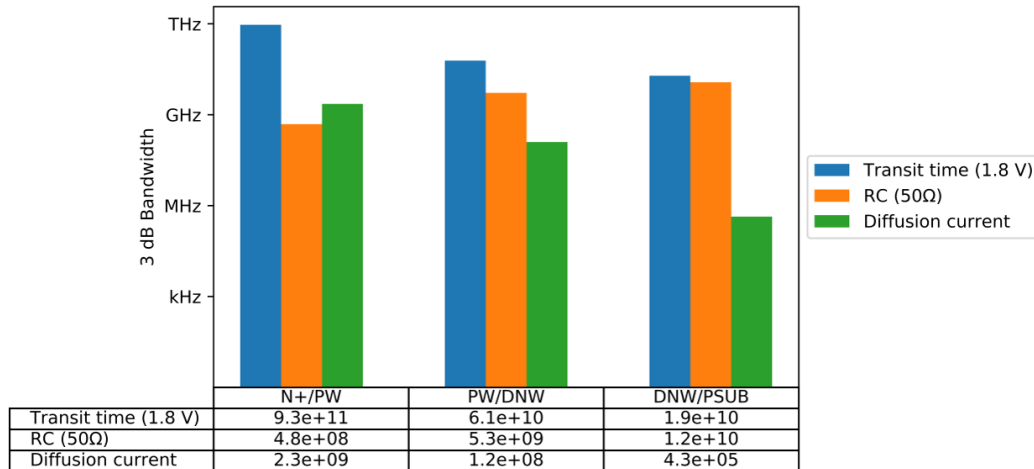


Figure 4.3: Estimated junction bandwidths due to transit time, capacitance, and diffusion current.

for cancelling the diffusion current have been proposed, including the differential photodiode and speed-enhanced photodiode [97]. However, since these methods inherently require removing the diffusion current from the signal, the responsivity is significantly lowered compared to a standard NW/PSUB junction. Since a high speed photodiode will inevitably suffer from reduced responsivity, an easier method is to simply use the DNW/PSUB junction as a guard to block diffusing carriers [97]. This leaves the N+/PW and PW/DNW junctions as photodiode options, which is just the usual gain-bandwidth tradeoff, since the PW/DNW junction offers higher responsivity while the N+/PW junction offers higher bandwidth. The question now arises as to how avalanche gain affects this gain-bandwidth tradeoff.

The bandwidth of an avalanche photodiode can be written as [98]

$$f_{3dB-APD} = \frac{0.36}{\tau_d(1 + M/M_o)}, \quad (4.18)$$

where τ_d is the transit time and M is the avalanche gain. $M_o = \alpha/\beta$, where α is the electron ionization coefficient and β is the hole ionization coefficient. α and β depend on the applied field E , and are given by

$$\alpha(\beta) = Ae^{-B/E}, \quad (4.19)$$

where A and B are constants that depend on the field E and the material properties. A and B for silicon can be obtained from [101], and so assuming a voltage V_A is applied to each of the junctions, M_o can be obtained using the depletion region widths for each junction, as tabulated in Table 4.4. V_A for each junction was chosen

to model a high field just before the onset of avalanche gain, as will be shown later. For the N+/PW and PW/DNW junctions now under consideration, it is clear that M/M_o will be much greater than one, and so the bandwidth of the APD will vary inversely with the gain. Consequently, avalanche gain does not fundamentally alter the gain-bandwidth tradeoff between these two junctions.

Table 4.4: Ratio of ionization coefficients for each junction under an applied voltage V_A .

Junction	V_A (V)	M_o
N+/PW	9	1.3
PW/DNW	14	5.3
DNW/PSUB	14	110

However, another consideration comes into play for avalanche gain in CMOS photodiodes. This is the necessity of a guard ring to prevent premature breakdown [71], [95]. For junctions without a guard ring, any imperfection in the planarity of the junction can cause field enhancement, which will decrease the voltage at which the junction breaks down. In CMOS photodiodes, the PW/DNW and DNW/PSUB junctions suffer from premature breakdown at the edges of the wells. In contrast, a guard ring can be formed for the N+/PW (or P+/NW) junction using shallow trench isolation, as shown in Figure 4.1. The oxide ensures that the N+/PW junction is entirely planar across the area of the photodiode.

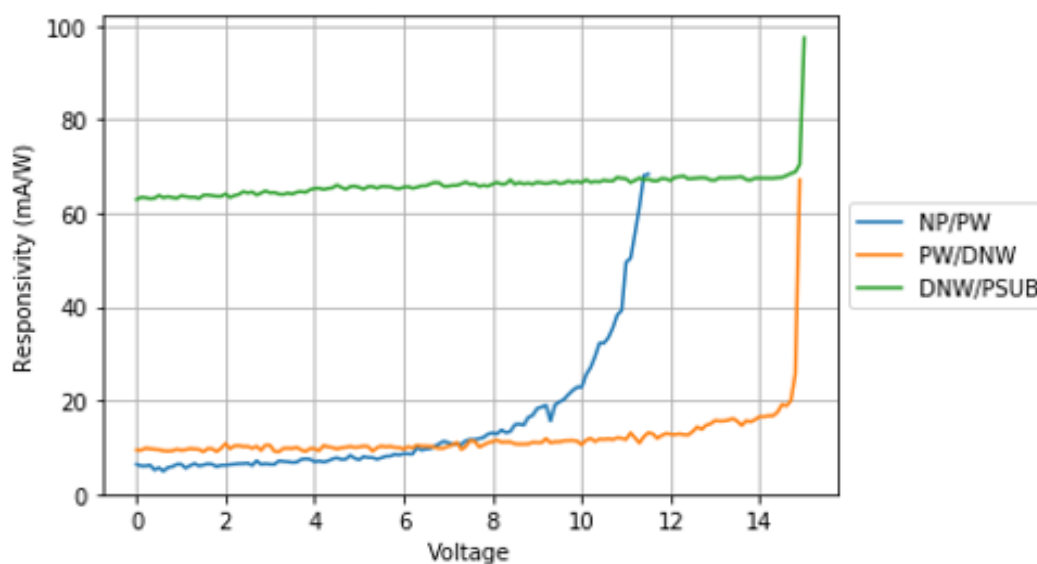


Figure 4.4: Comparison of the response of each junction of a CMOS photodiode to increasing reverse bias.

The effect of the guard ring can be seen quite clearly in the measurement shown in Figure 4.4. The N+/PW junction exhibits smooth avalanche gain, while the other two junctions exhibit a sharp breakdown threshold, which is consistent with premature breakdown at the edges of the wells. This difference between the junctions strongly suggests that N+/PW is the optimal junction for system design, as it possesses the option of using avalanche gain, the highest intrinsic frequency cutoff (due to transit time and diffusion current), and an extrinsic frequency cutoff (capacitance) that can be mitigated using the on-chip waveguides.

As mentioned above, the ability to use on-chip waveguides to guide light to very small CMOS photodiodes is a critical benefit that the method of subtractive photonics provides over competing systems using bulk CMOS photodiodes. Although it is possible to use small photodiodes with direct fiber-photodiode coupling in these competing systems (since the core of a 780 nm fiber is about $5\ \mu\text{m}$, and the spot size of a lensed 780 nm fiber is about $2\ \mu\text{m}$), the fiber alignment and fixing can be expensive, especially for a large number of photodiodes. In contrast, the method of subtractive photonics separates photodiode coupling from fiber alignment, greatly easing manufacturing constraints.

So far in our discussion of tradeoffs we have included bandwidth limitations due to transit time, capacitance, and diffusion current. Transit time can be minimized by minimizing the depletion region width, while capacitance can be minimized by minimizing the depletion region area. Both of these design techniques reduce responsivity. To mitigate the bandwidth limitation caused by diffusion current, the diffusion current must be cancelled, further reducing responsivity. To compensate for all of these reductions in responsivity and bandwidth, the ideal photodiode would consist entirely of a depletion region to eliminate diffusion current. The path of the carriers would be orthogonal to the light absorption, which would enable short transit times while maximizing responsivity. The increased length of the photodiode required for maximizing absorption would also increase capacitance, but the capacitance could be absorbed into a transmission line to increase the RC cutoff [102]. The bandwidth of this ideal photodiode, a waveguide-coupled traveling-wave photodetector, would be primarily limited by the mismatch between the optical group velocity and the electrical phase velocity². The ability to form this type of structure using waveguides in bulk CMOS is another advantage of subtractive photonics.

²Assuming the optical absorption is sufficiently high, the electrical attenuation can be neglected in the design [102].

In summary, subtractive photonics provides the following advantages for bulk CMOS photodetectors:

- On-chip waveguides can be used to direct light onto very small photodiodes independent of fiber alignment, minimizing capacitance while increasing manufacturing flexibility.
- On-chip waveguides can be used to form traveling wave photodiodes, maximizing absorption and bandwidth.

4.2 Tapering Waveguide-Photodiode Coupler

The possibility of traveling wave photodiodes motivates the use of tapered waveguide-photodiode couplers, as shown in Figure 4.5. These couplers have the additional advantage of providing adiabatic coupling to the silicon substrate, mitigating losses due to reflections at the glass-silicon interface.

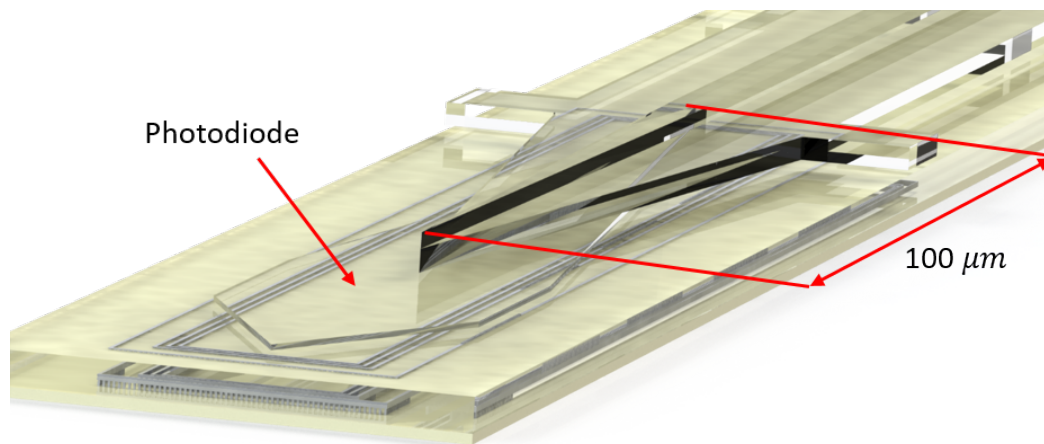


Figure 4.5: 3D model of the implemented tapering waveguide-photodiode couplers.

The photodiode active area is highlighted in Figure 4.6. Observe that the surrounding glass protects the photodiode interconnect from being etched. It is crucial that final checks before tapeout confirm that no interconnect metal is touching waveguide metal anywhere on the chip³, or disaster could result.

The ridge and slab of the waveguide are tapered down separately, as shown in Figure 4.7. The sharper slab taper has little effect on the mode until the end of the taper due to the confinement of the mode in the vicinity of the ridge.

³For example, by net tracing.

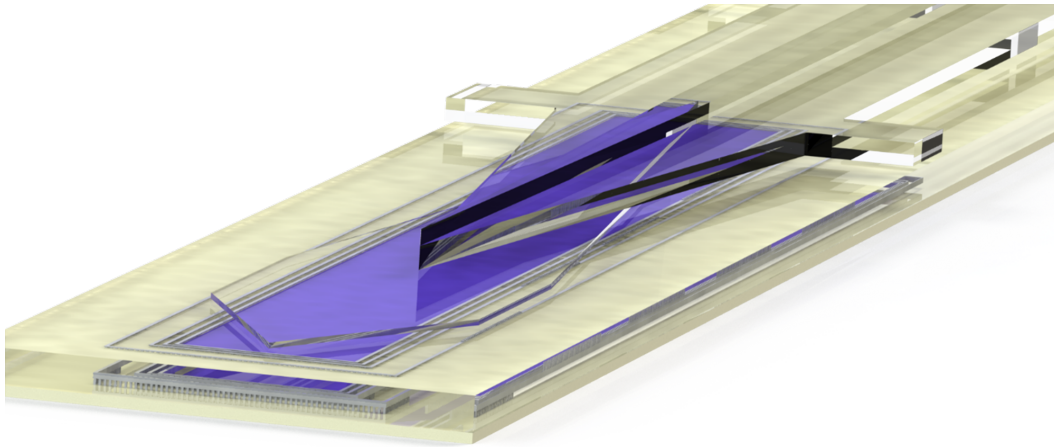


Figure 4.6: Location of the CMOS photodiode in the tapering coupler.

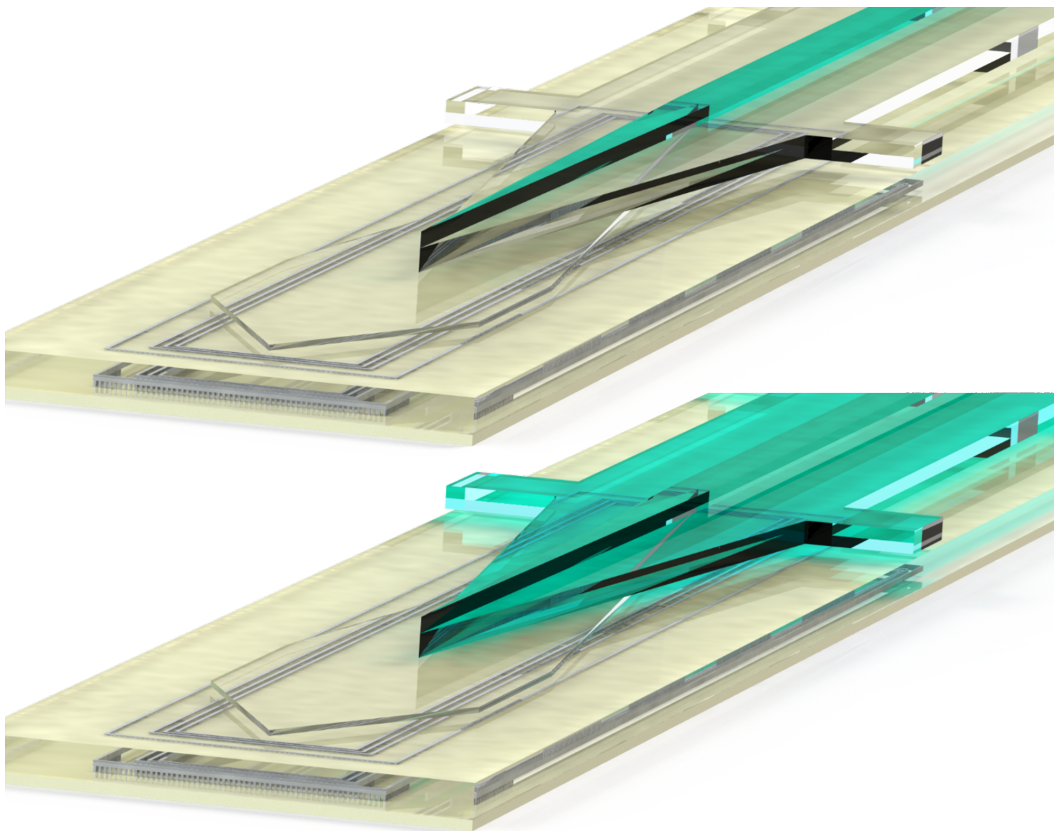


Figure 4.7: Tapers in the photodiode coupler. Top: Ridge taper. Bottom: Slab taper.

The most critical part of the coupler, the layer of glass connecting the waveguide to the photodiode's isolation glass, is shown in Figure 4.8. The thickness of this

layer is constrained by the waveguide geometry and photodiode interconnect. The waveguide used for the measured couplers is a V5M5-V4M4V3 waveguide. The photodiode uses M1 to provide contact to the silicon, and M2 to connect the M1 of each well to exterior pads. Since the M2 interconnect cannot be etched, the V2 layer must remain entirely glass to provide isolation from the etchant. Since the bottom layer of the waveguide is V3, this leaves only the M3 layer to provide an air gap between the photodiode and the waveguide, as well as serve as the coupling region between the two.

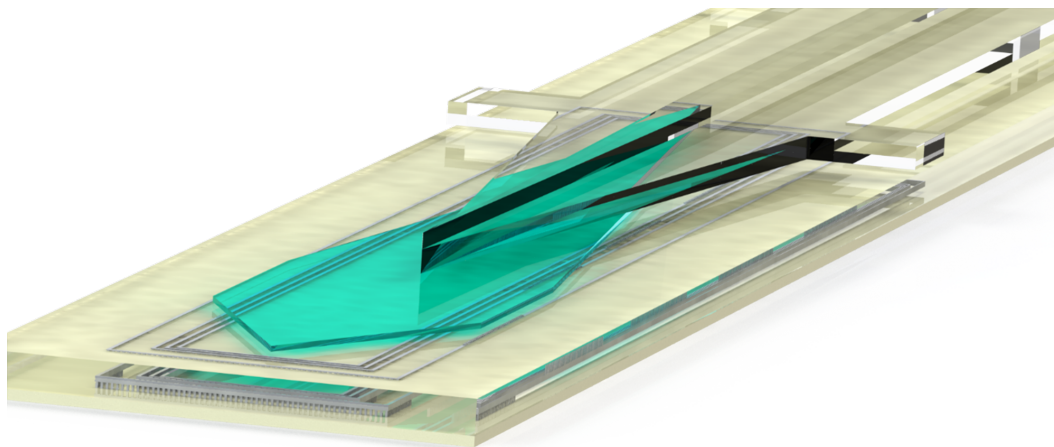


Figure 4.8: Coupling region between the waveguide and the photodiode isolation glass.

Small Photodiode

For testing the coupler, we fabricated a photodiode with an active area of $16 \mu\text{m} \times 150 \mu\text{m}$ (Fig. 4.9). The responsivity of the N+/PW junction was measured at 635 nm and 780 nm across a range of reverse biases, as shown in Figure 4.10. Note the alternating layers of low-k dielectric and silicon dioxide in the FDTD simulation, and that the responsivity could go much higher if the reverse bias was brought closer to breakdown.

In order to verify that light was coupling into the photodiode through the waveguide, we scanned the fiber across the waveguide facet and recorded the response, effectively imaging the facet using the coupled photodiode (Fig. 4.11). An image of the V5M5-V4M4V3 waveguide facet is shown in Figure 4.11, confirming that light was coupled through the waveguide, and demonstrating the importance of taking slab modes into account with this single mode waveguide.

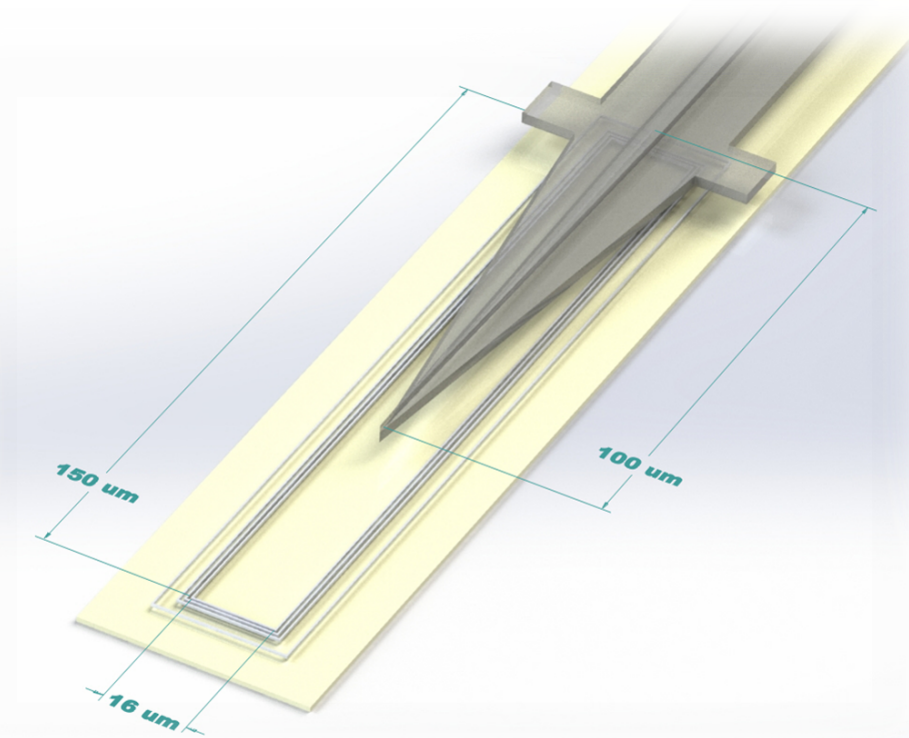


Figure 4.9: Dimensions of the small photodiode and tapering coupler.

The responsivity shown in Figure 4.10 includes loss due to fiber-waveguide coupling, as well as the loss due to the coupler itself. We can take all of the losses associated with the waveguide into account by measuring the response of the photodiode due to a vertically coupled lensed fiber at 780 nm, as shown in Figure 4.13. At zero bias, the vertically coupled photodiode exhibits a responsivity of 8.5 mA/W, while the waveguide coupled photodiode exhibits a responsivity of 0.95 mA/W, corresponding to a loss of 9.5 dB. An FDTD simulation of the waveguide coupler, including the full low-k dielectric stackup and silicon substrate, and assuming excitation of the fundamental mode, yields a coupling efficiency of 30% at 780 nm (Fig. 4.14), or 5.2 dB loss. This leaves 4.3 dB loss for the fiber coupling efficiency, which may be compared against the simulated coupling efficiency shown in Figure 3.42, which posits 1.5 dB loss for ideal coupling and no more than 3 dB loss for a fiber up to 10 μm from the waveguide facet. This 2.8-1.3 dB excess loss may be ascribed to imperfect fiber alignment in the plane of the waveguide facet, since the simulated fiber-waveguide and waveguide-photodiode coupling efficiencies assumed excitation of only the fundamental mode. The higher order confined modes of the V5M5-

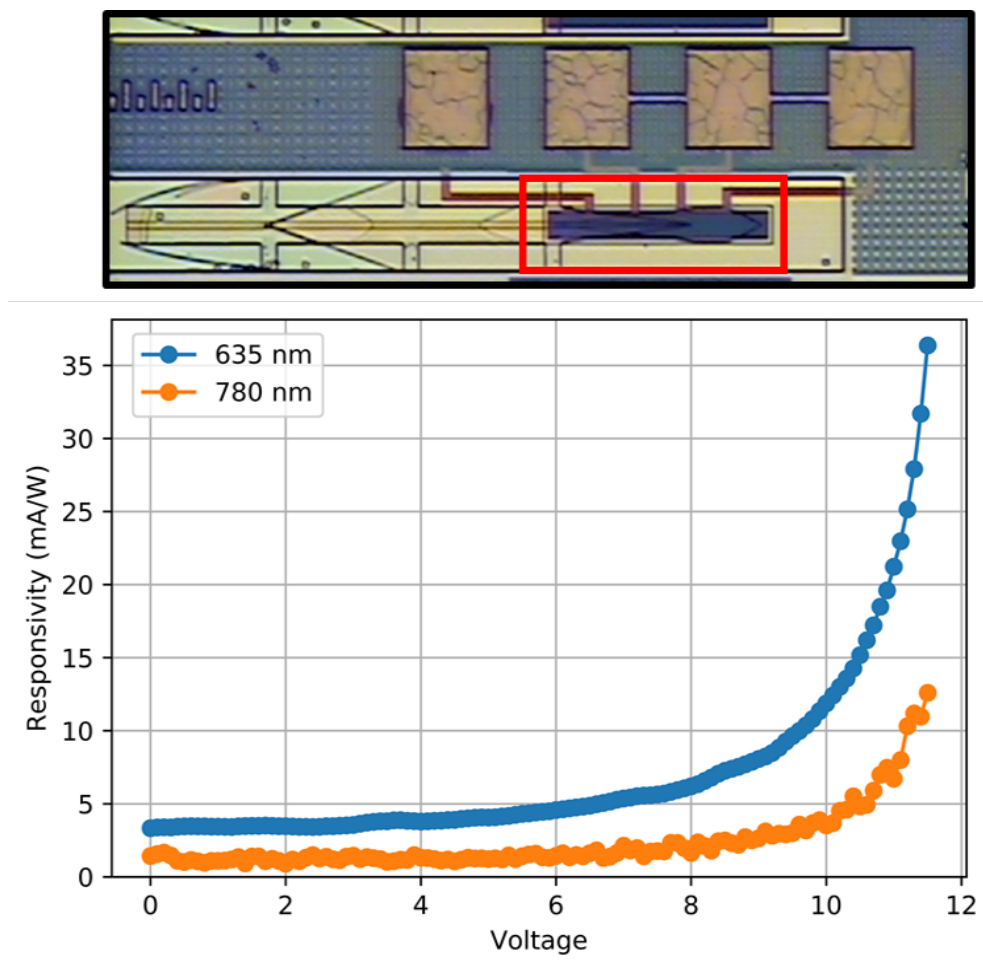


Figure 4.10: Measurements of the small photodiode. Top: Die photo with the small photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.

V4M4V3 waveguide are slab modes, as indicated by Figure 4.12, so any power coupling into these modes would be poorly confined and more likely to scatter at changes in the waveguide geometry, such as at supports or at the taper, reducing coupling efficiency.

Large Photodiode

In order to determine the impact of photodiode width on the waveguide-photodiode coupler, we also fabricated a $50 \mu\text{m} \times 150 \mu\text{m}$ photodiode, with the same coupler used with the small photodiode (Fig. 4.15). Similar to the small photodiode, responsivity was measured at the wavelengths of 635 nm and 780 nm (Fig. 4.16) and the waveguide facet was imaged (Fig. 4.17).

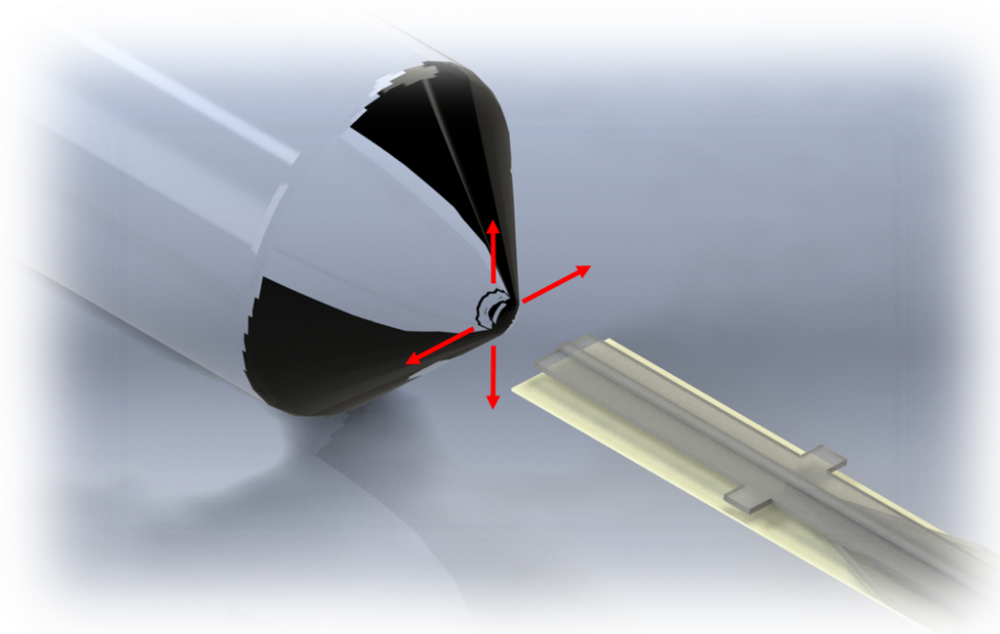


Figure 4.11: Setup for imaging the waveguide facets by scanning a lensed fiber.

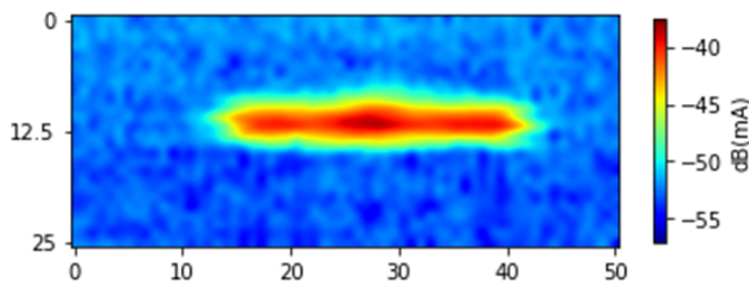


Figure 4.12: V5M5-V4M4V3 waveguide facet imaged using the small photodiode.

The larger responsivity due to the $50\ \mu\text{m}$ width can be seen not only in Figure 4.16, but also in a comparison of Figures 4.11 and 4.17.

In Figure 4.18, the zero bias responsivity of the vertically coupled photodiode is $14.1\ \text{mA/W}$, while the responsivity of the waveguide coupled photodiode is $6.6\ \text{mA/W}$, corresponding to a total coupling loss of $3.3\ \text{dB}$. The FDTD simulated coupling efficiency is 34% , or a loss of $4.7\ \text{dB}$. This discrepancy between simulation and measurement may again be explained by the assumption of excitation of only the fundamental mode in the simulation. If higher order slab modes are excited during fiber-waveguide coupling, these modes may scatter more than the fundamental mode, and the large area of this photodiode may collect that light, increasing the responsivity

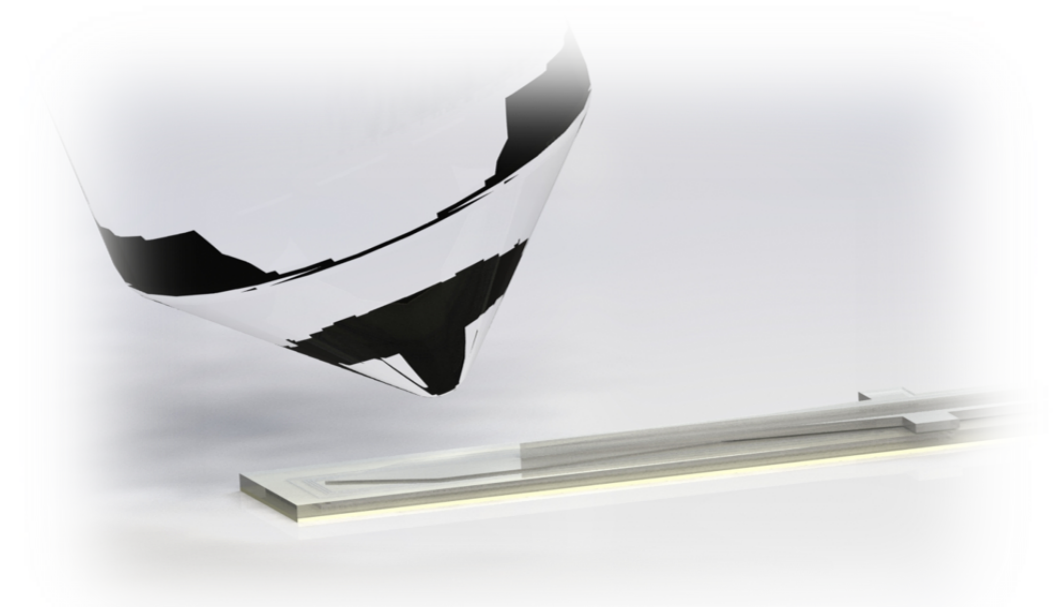


Figure 4.13: Setup for measuring the responsivity of the small photodiode independent of the waveguide.

to higher than that due to the fundamental mode alone.

Segmented Photodiode

In order to determine how the geometry of the tapering waveguide coupler might be optimized, we fabricated a segmented photodiode to measure the relative absorption of light along the length of the taper, as shown in Figure 4.19. In this photodiode, each segment consisted of an N+/PW junction $18\ \mu\text{m}$ wide and $23\ \mu\text{m}$ long.

The responsivity is shown in Figure 4.20 for wavelengths of 635 nm and 780 nm. The peaks in the center, at segments 5 and 6, indicate the end of the waveguide taper. The peak at segment 1 is most likely due to imperfect fiber coupling - the excitation of higher order slab modes or radiation modes that travel below the ridge of the waveguide and are easily absorbed by segment 1. The conclusion to be drawn from Figure 4.20 is that the very smallest widths of the taper are causing disproportionate coupling to the substrate, a fact which should be taken into account in the design of any traveling wave photodiodes.

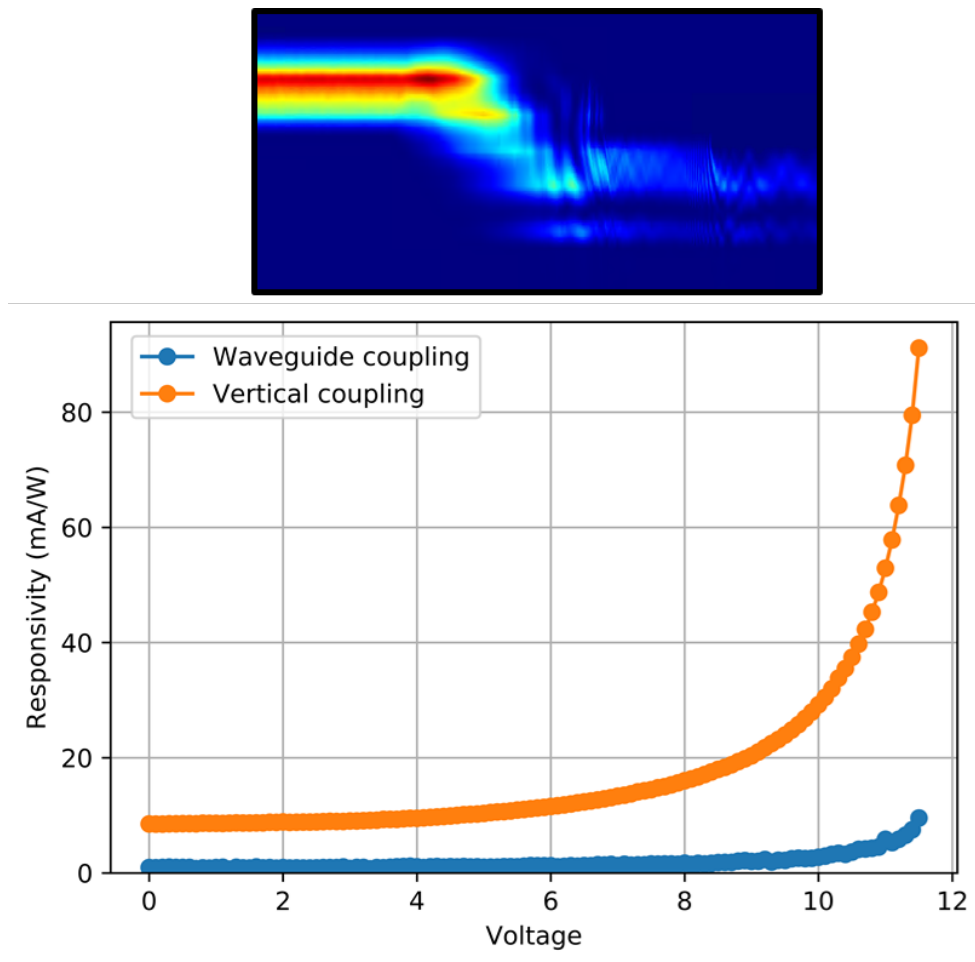


Figure 4.14: Determining coupling loss for the small photodiode. Top: FDTD simulation of coupling into the small photodiode. Bottom: Comparison of waveguide coupling and vertical coupling.

4.3 Reflecting Waveguide-Photodiode Coupler

As mentioned above, one of the advantages of subtractive photonics is that it decouples fiber alignment from photodiode coupling, easing the implementation of very small photodiodes. The ideal method of coupling to very small photodiodes would be to terminate a horizontal waveguide with a reflecting surface oriented at an angle to reflect light into the CMOS photodiode. Of course, such angled surfaces are not possible to fabricate in an unmodified bulk CMOS process, and the best we can do is approximate such a structure by using a number of layers. As shown in Figure 4.21, such a reflector will consist of a staircase type of structure for, roughly speaking, "bending" the waveguide downwards. For the structure shown in Figure 4.21, the layers were initially aligned in a staircase pattern to match as closely as possible an

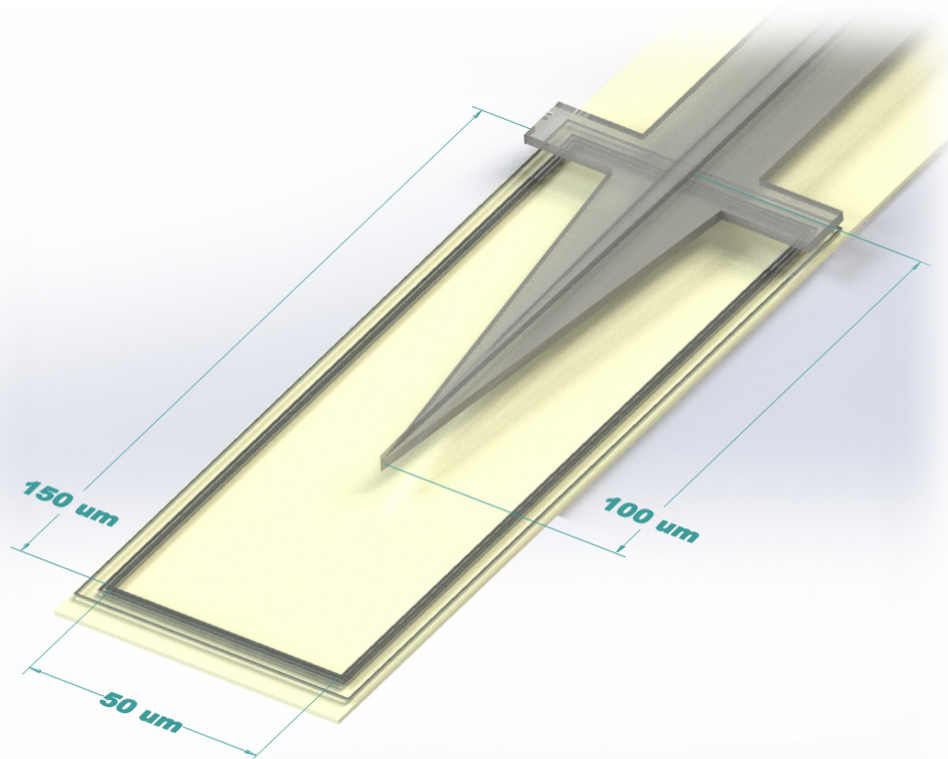


Figure 4.15: Dimensions of the large photodiode and tapering coupler.

angled reflector with a smooth, flat surface. The endpoints of the layers were then optimized using particle swarm optimization in a commercial FDTD tool. It was found that the addition of a second reflector moderately improved performance, but the addition of a third reflector improved performance very little. The structure shown in Figure 4.21 has a transmission of 54% into the silicon substrate at a wavelength of 780 nm.

The transmission of the implemented reflecting coupler was measured with respect to a standalone reference photodiode of the same dimensions. As can be seen in the die photo of Figure 4.21, the insertion loss includes fiber-waveguide coupling loss and the loss of a 100 μm radius bend. It is assumed that the loss of the waveguide itself, including reflections from supports, is negligible. The insertion loss was measured to be 9.6 dB at a wavelength of 635 nm, and 10.7 dB at a wavelength of 780 nm. As mentioned above, Figure 3.42 shows that the fiber coupling loss varies from 1.5 dB for ideal coupling up to 3 dB for a fiber 10 μm away from the waveguide facet. Taking the simulated 2.7 dB coupling loss into account, this leaves 5-6.5 dB due to

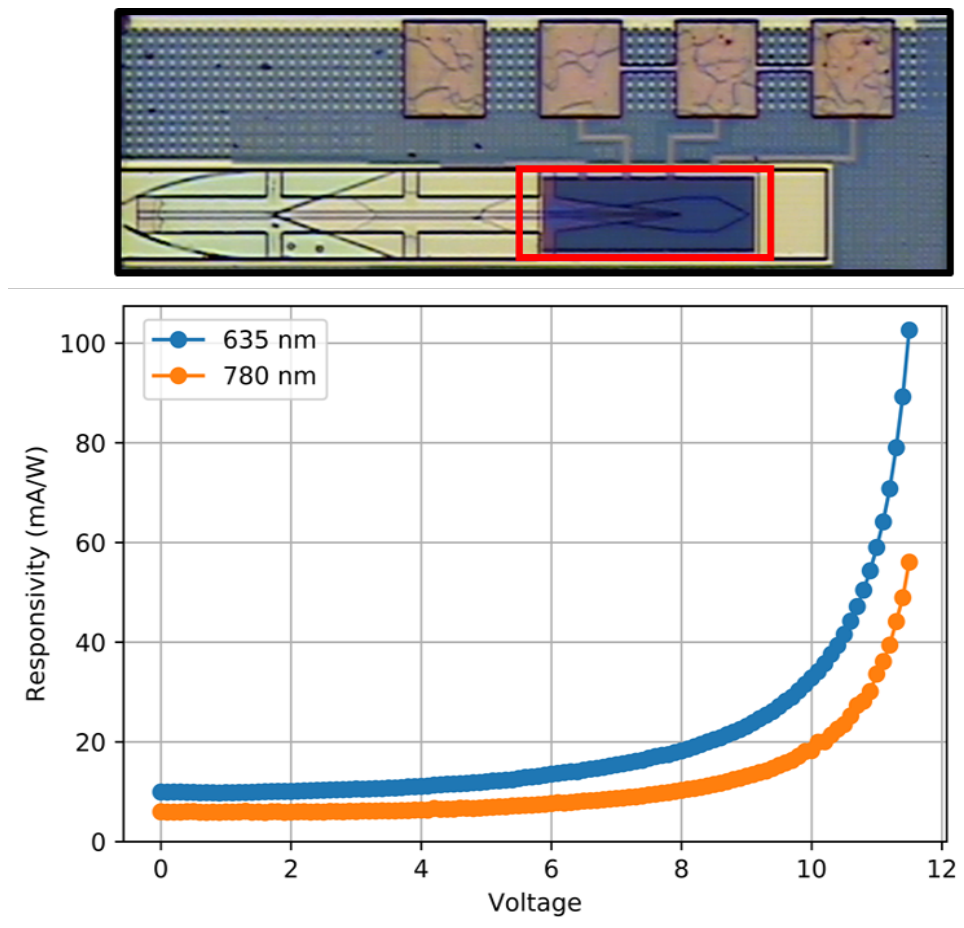


Figure 4.16: Measurements of the large photodiode. Top: Die photo with the large photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.

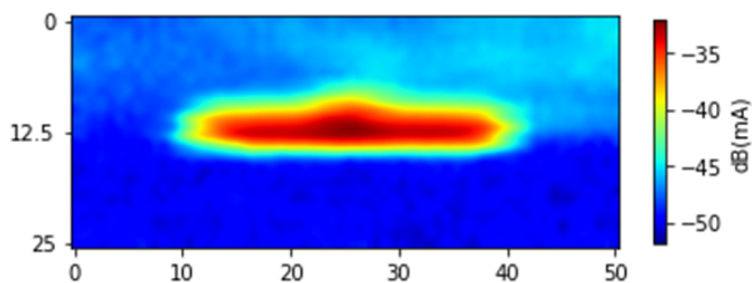


Figure 4.17: V5M5-V4M4V3 waveguide facet imaged using the large photodiode.

bend loss, which is low for a standard V5M5-V4M4V3 bend (Figure 3.44). It may be that, as with the tapering waveguide couplers, imperfect fiber coupling has resulted in the excitation of slab modes, improving the performance over the expected single mode behavior.

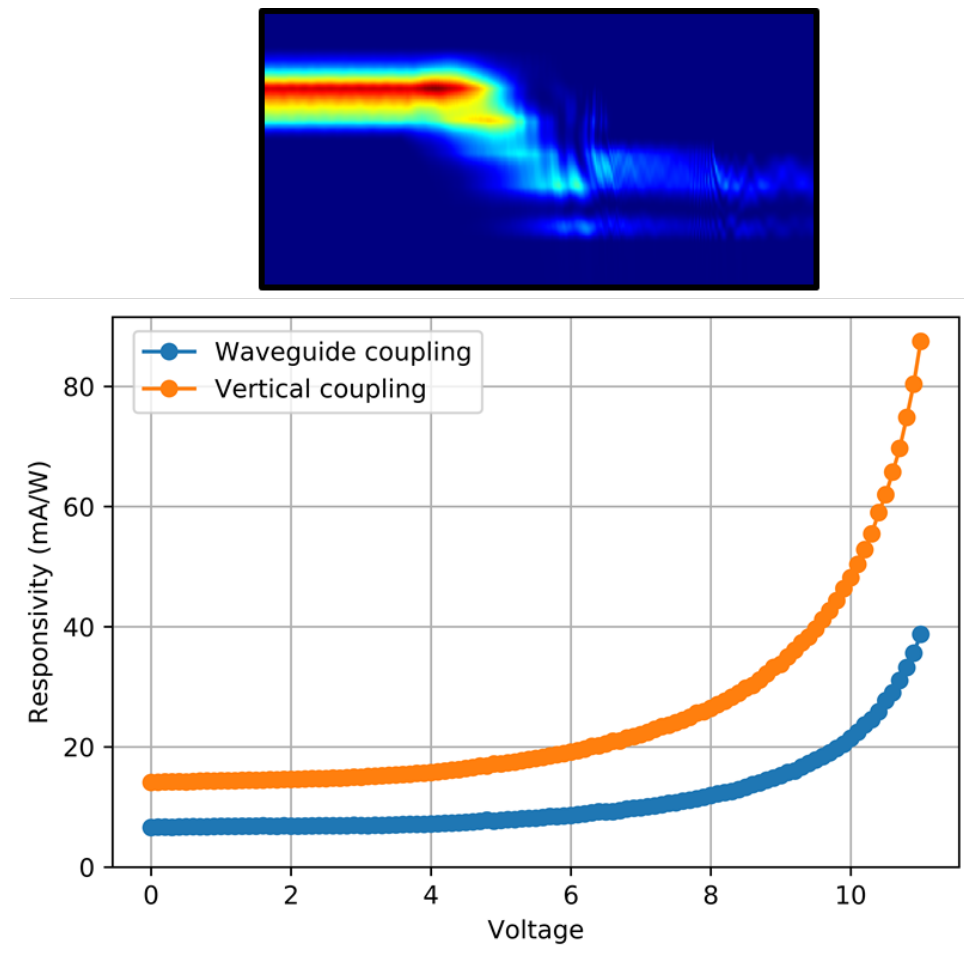


Figure 4.18: Determining coupling loss for the large photodiode. Top: FDTD simulation of coupling into the large photodiode. Bottom: Comparison of waveguide coupling and vertical coupling.

Although the implemented reflecting coupler has superior transmission to the tapering couplers described above, it may be possible to optimize the tapering couplers and reduce the gap between them. The main advantage of the reflecting coupler is the greatly reduced area compared to the tapering coupler, and the main disadvantage is that very small photodiodes may not be able to match the absorption and bandwidth of traveling wave photodiodes.

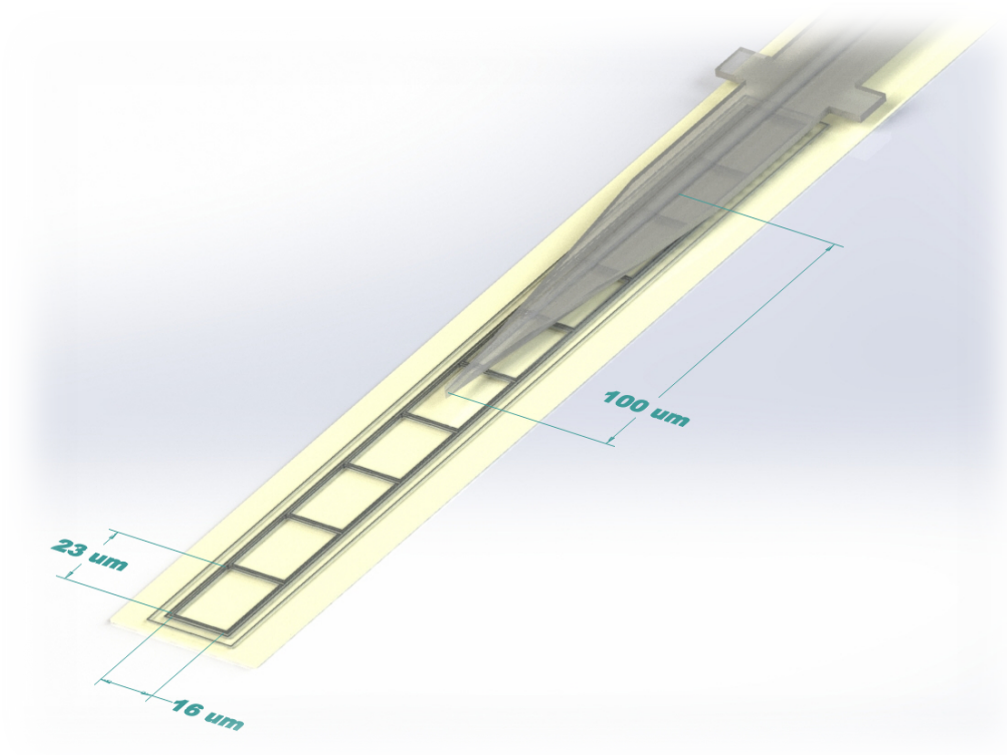


Figure 4.19: Dimensions of the segmented photodiode and tapering coupler.

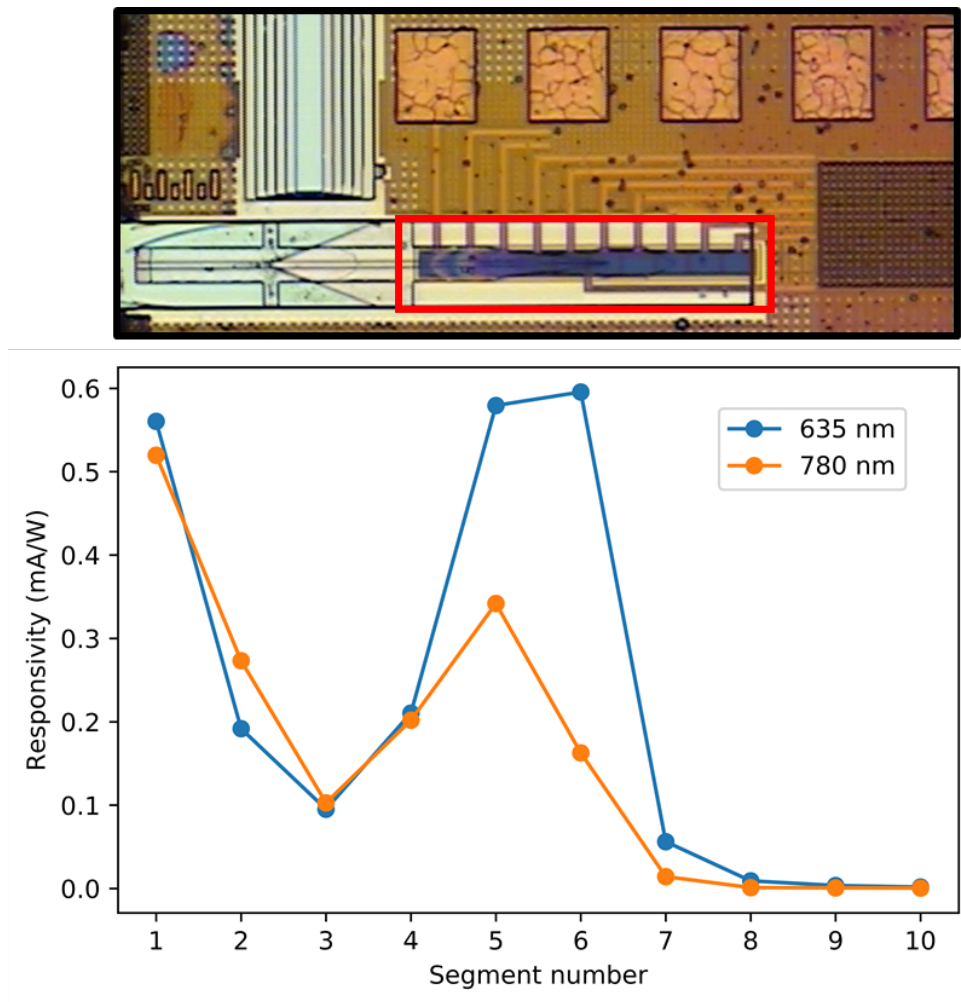


Figure 4.20: Measurements of the segmented photodiode. Top: Die photo with the segmented photodiode highlighted. Bottom: Responsivity at the wavelengths of 635 nm and 780 nm.

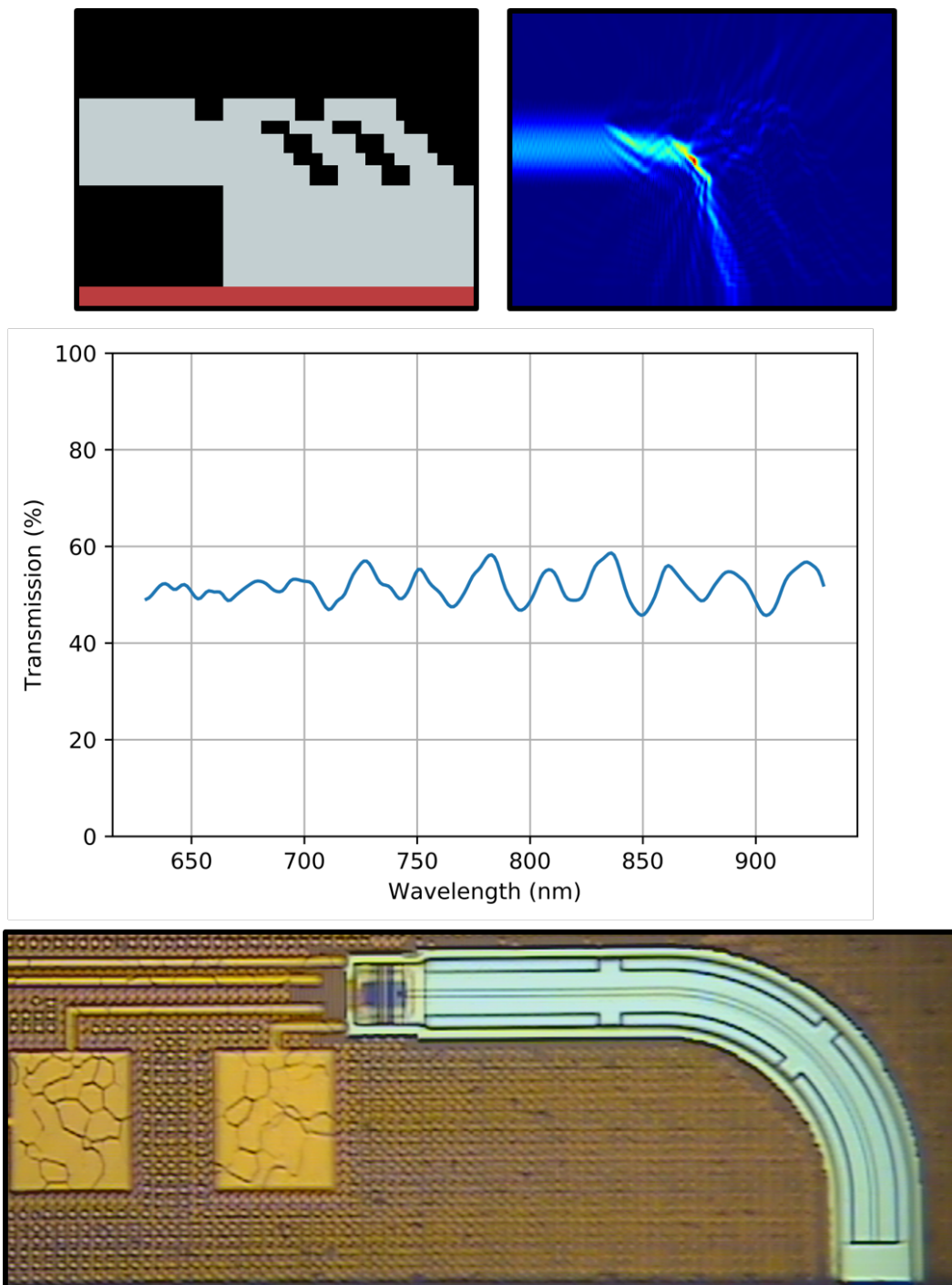


Figure 4.21: Reflecting coupler simulations. Top: FDTD simulation setup and cross section. Middle: Transmission of the reflecting coupler across wavelengths. Bottom: Die photo of the reflecting coupler and connected waveguide.

OPTICAL REFERENCE DISTRIBUTION

Note: The work described in this chapter was performed in collaboration with Matan Gal-Katziri and Armina Khakpour.

An interesting application of the method of subtractive photonics is optical reference distribution. The concept was first considered in the context of integrated circuits in [21], in which the author considered the use of waveguides integrated in a VLSI chip for distributing the clock of a digital system. A similar application for analog circuits would be distributing the reference for a phased array. Figure 5.1 shows an example of optical (red) reference distribution on a phased array, complemented by electrical (blue) reference distribution.

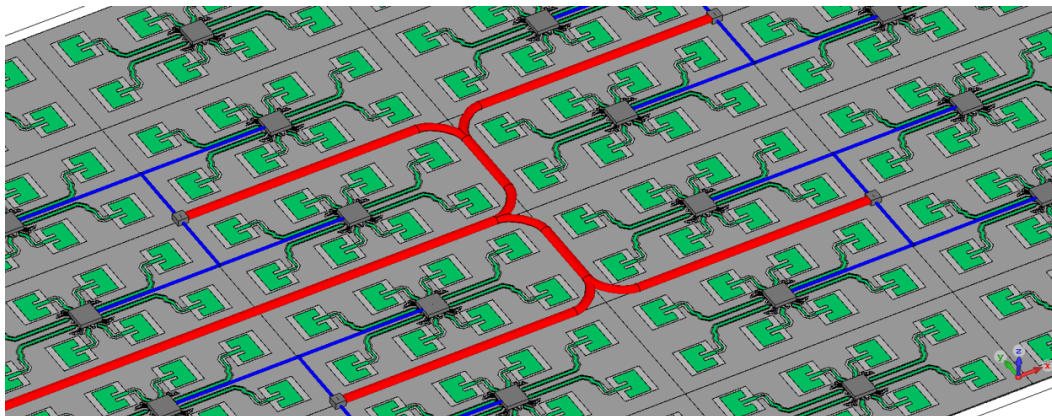


Figure 5.1: An example of an RF phased array with both optical (red) and electrical (blue) reference distribution.

We demonstrated a proof-of-concept of an optically synchronized phased array using an optical fiber coupled directly to a CMOS photodiode in bulk CMOS. The 7 GHz reference detected by the photodiode is amplified by a transimpedance amplifier (TIA) chain, multiplied to 28 GHz by a PLL, and distributed using on-chip transmission lines to eight transmitter channels (Fig. 5.2).

The bulk CMOS photodiode used the N+/PW junction to reduce diffusion current and enable avalanche gain, as described in Chapter 4. The photodiode used three fingers to reduce series resistance, and had an area of about $45 \mu\text{m} \times 45 \mu\text{m}$ to

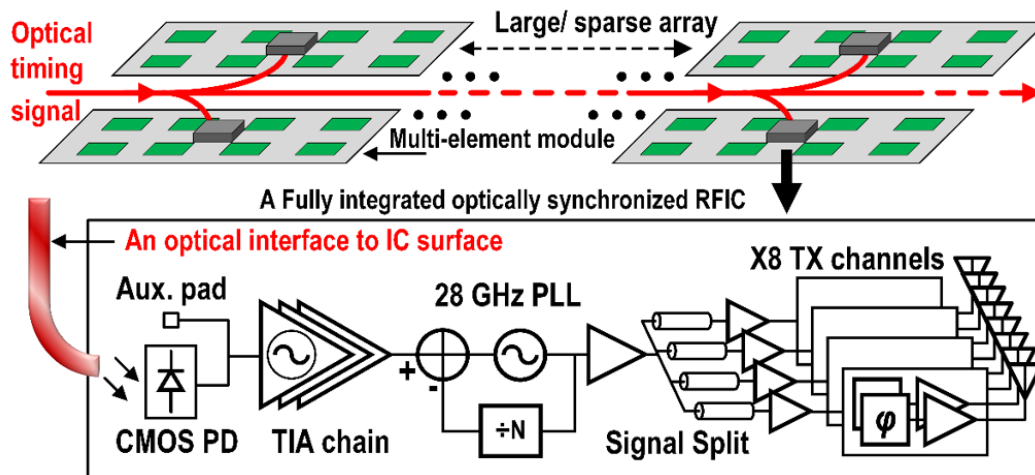


Figure 5.2: Overview of the optically synchronized phased array. An optical timing signal (red) is distributed to subarrays, each of which receive the reference, amplify it, and distribute it to eight TX channels.

provide tolerance for the fiber alignment. The responsivity of the photodiode with bias is shown in Figure 5.3, reaching a high of 9.3 A/W near breakdown.

The photodiode capacitance was estimated to be near 500 fF, which results in an RC limited bandwidth of 6.4 GHz with a load of 50 Ω . Since the reference is at 7 GHz, the capacitance of the photodiode was resonated out using an on-chip inductor, which also provides DC bias to the photodiode (Fig. 5.4). The reference from the photodiode is sensed by a differential pair, which injection locks the LC oscillator it is in parallel with. The differential reference output is then AC coupled to an amplifier chain divided down to 3.5 GHz before being multiplied to 28 GHz by the PLL.

The fiber was attached to the board using the setup shown in Figure 5.5. Note that the RFIC is mounted on the upper board, on the same side as the eight antennas, while a lower board secures a pipe that the fiber passes through. The pipe and the accompanying mount are necessary because the fiber could not be fixed to the chip due to a mistake in the chip layout that left only a few samples available for measurement. Consequently, the fiber was epoxied to the lower board only. A hole was opened in the upper board to enable the fiber to pass through it, and this hole was then filled with index-matching gel to substitute for the index matching that would have been provided by epoxy, had the fiber been epoxied directly to the chip. As a result, this setup enabled fibers to be removed from an RFIC and reattached as needed.

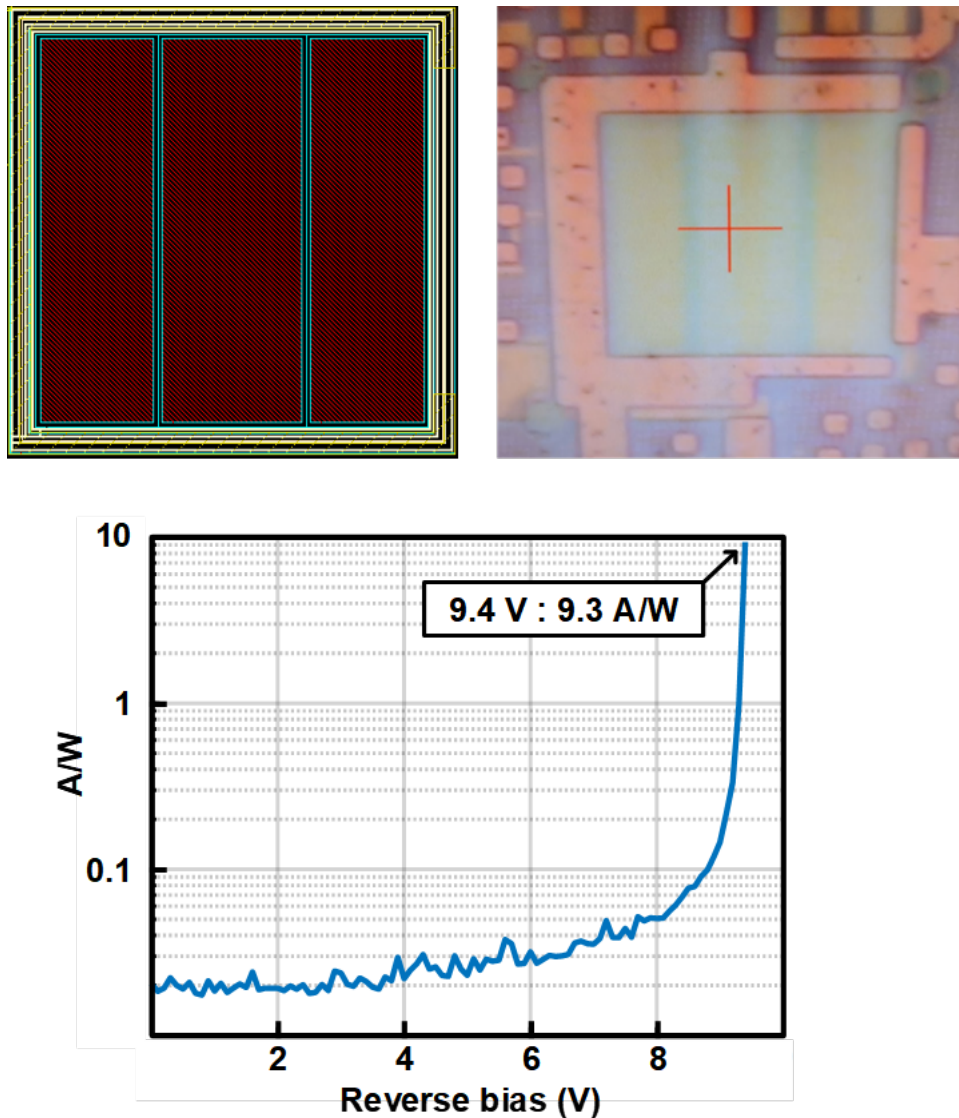


Figure 5.3: The 65 nm bulk CMOS photodiode. Top: Layout and die photo. Bottom: Measurement of the photodiode responsivity with bias.

The full measurement setup is shown in Figure 5.6. A 780 nm laser is externally modulated at 7 GHz using an RF signal generator. The optical signal is then passed through an SOA with about 9 dBm of power at the output. The optical reference signal arrives at the RFIC after traveling through 25 m of optical fiber. The phased array module is mounted on a turn table that enables measurement of the beam pattern and is controlled by a local PC. The RFIC is also programmed from the PC using a standard serial interface. The probe is a horn antenna mounted 40 cm from the phased array module, and it is surrounded by RF absorbing foam to prevent reflections. The measured beam patterns of the optically synchronized phased array

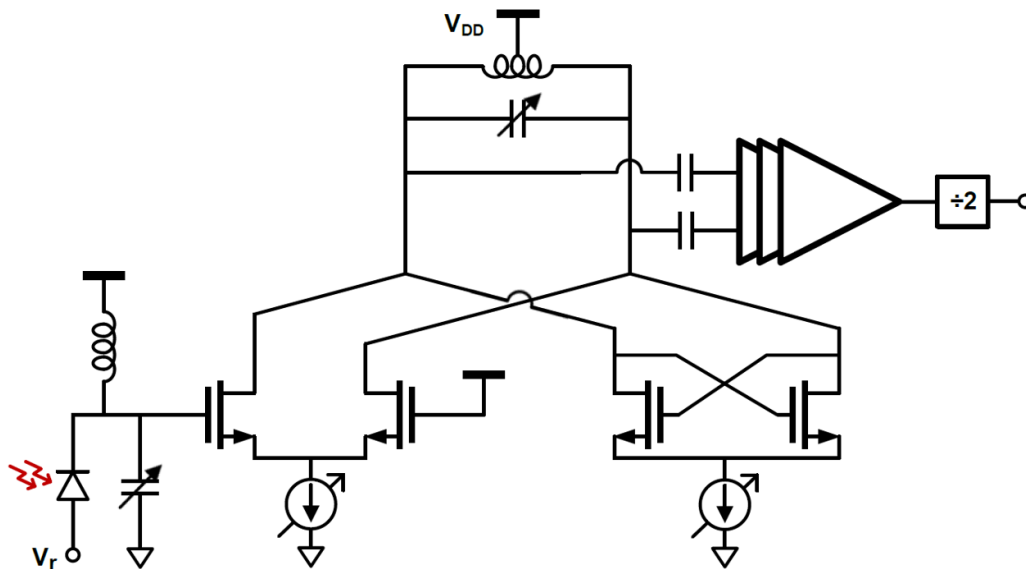


Figure 5.4: The injection locked TIA. Note that V_r provides reverse bias to the photodiode.

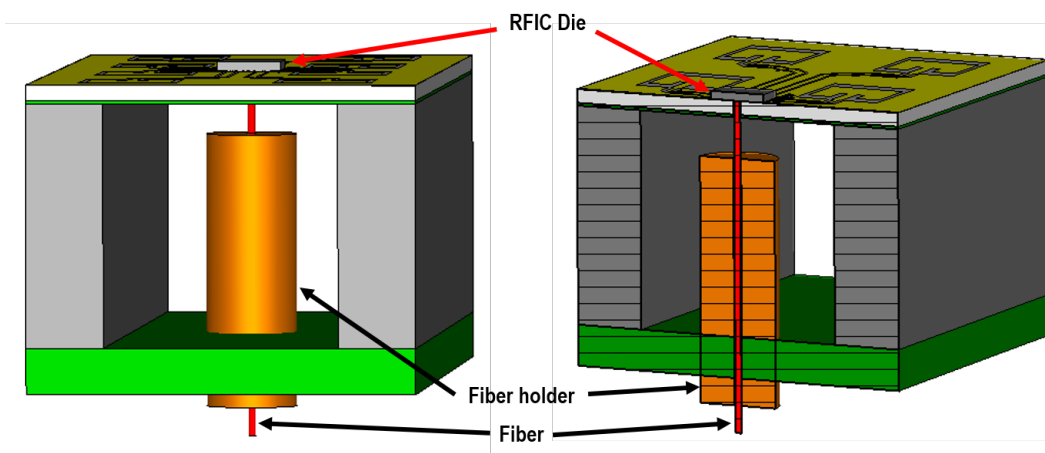


Figure 5.5: Two views of the board setup used to couple the fiber to the RFIC. The copper pipe is held in place by the lower board, and does not touch the upper board. The fiber passes through the pipe and then through a hole cut in the upper board that is filled with index matching gel. The fiber is then fixed to the lower board and the pipe.

module are shown in Figure 5.7, demonstrating beam steering at 15° , 30° , and 45° .

Another measurement setup was created to verify the synchronization of two phased array modules (Fig. 5.8). The two modules were placed in parallel with 5 m of fiber providing the optical reference to each module. With one element activated on each

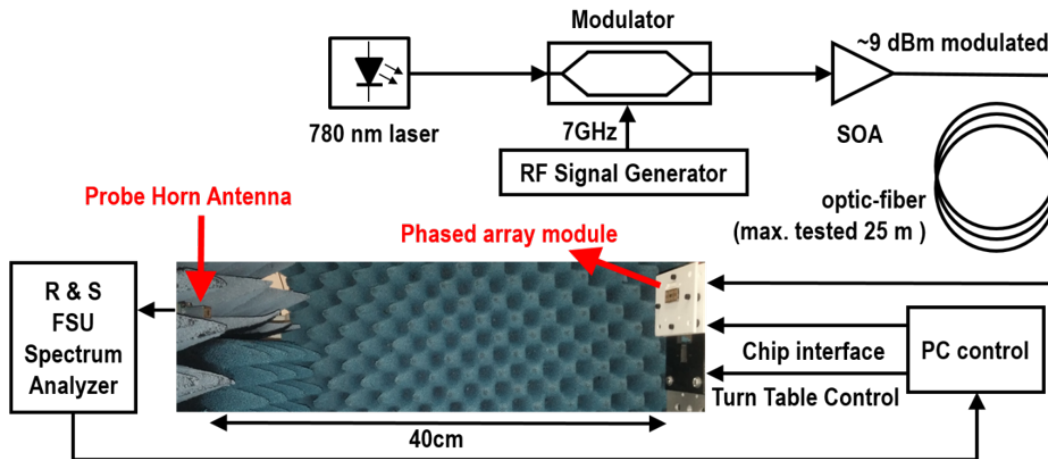


Figure 5.6: Beam pattern measurement setup.

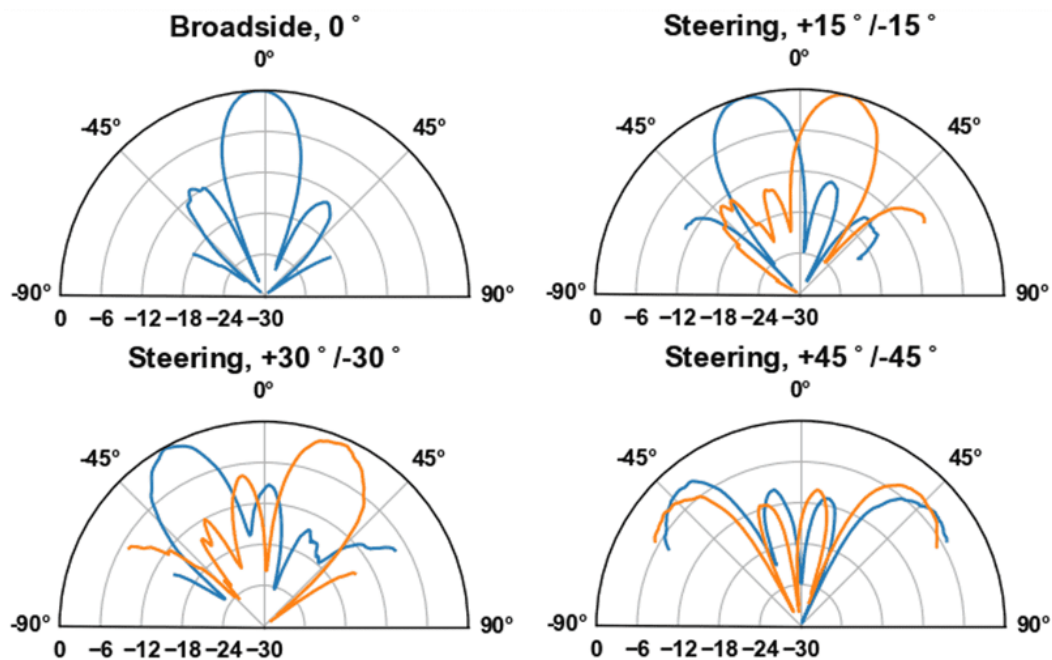


Figure 5.7: Measured beam patterns of the phased array module using optical reference distribution.

module, the phase of one module was swept 360° . The coherent addition of the two beams is shown in Fig. 5.8, and closely follows the expected sinusoid.

The difficulty of fixing the fiber and the distribution of the 28 GHz reference on the chip using transmission lines both indicate the value of subtractive photonics. With optical waveguides on chip, the reference could be distributed directly to each transmitter channel. And with edge couplers the fiber could be attached in the plane

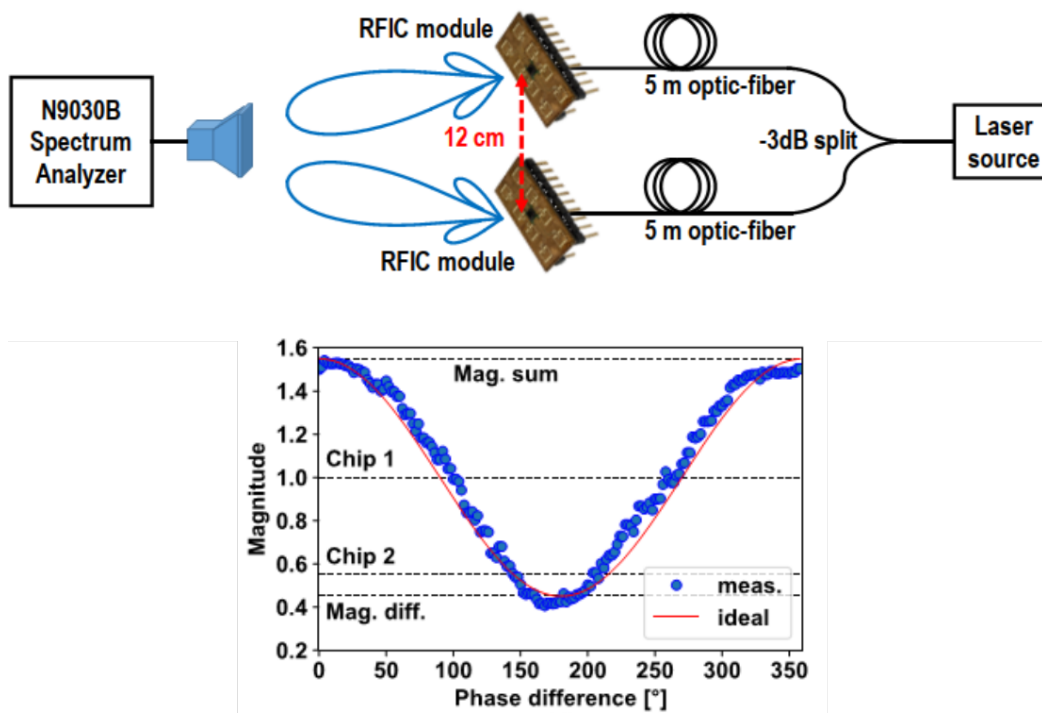


Figure 5.8: Demonstration of synchronization of two phased array modules.

of the chip, resting on the board and making fixing the fiber significantly easier. For these reasons, subtractive photonics is poised to significantly improve the design of systems using optical reference distribution.

*Chapter 6***CONCLUSION****6.1 Overview of Thesis Work**

The goal of this thesis has been to explore the possibilities of integrating photonics into bulk CMOS. The introduction showed that a number of options have been previously explored in the literature. Most of these options require using the FEOL, either by etching the substrate underneath polysilicon gate waveguides, or by modifying the process. The former option results in waveguides with losses of 55 dB/cm, or 38 dB/cm for polysilicon waveguides using subwavelength gratings. For the latter option, the method using process modification that has been published the most is deep trench isolation. Losses as low as 10.5 dB/cm have been achieved using this method. However, the process modification creates a formidable barrier to market adoption, since the cost of modifying the process must be born for every process that is to have photonics integrated.

The other option for integrating photonics into bulk CMOS is to use the BEOL. This is highly appealing since it decouples the area and process steps used by electronics and photonics, while preserving the flexibility and low parasitics of monolithic integration. Most approaches in the BEOL category require the deposition of waveguides, modulators, and photodetectors above the electronic interconnect, using either polysilicon waveguides with doping for modulation using the plasma dispersion effect, or using silicon nitride waveguides with deposited electro-optic polymer for modulation. The main downside to both of these approaches is the cost required for the deposition of the photonic components.

The method of subtractive photonics provides a convenient BEOL alternative. By forming the waveguides using the BEOL dielectric, the BEOL processing in electronics processes is put to double use: once for electronic interconnect, and again for the patterning of photonic devices. The method of subtractive photonics only requires etching the electronics chips in Aluminum Etch Type A to remove the metal, and then etching them in a 1:1 solution of hydrogen peroxide and EDTA to remove any remaining barrier layers. This simple wet etch is much less expensive than all the deposition steps required in process-modifying FEOL or BEOL approaches, and yields waveguides with much lower loss than those waveguides fabricated using the

zero-change method.

We have shown that although the method of subtractive photonics opens up a large and previously untapped design space, it is also subject to considerable constraints. A large number of DRC violations in initial experiments led to the adoption of optical structures using metal-via overlap on all non-optical vertical surfaces. The tediousness of implementing this overlap rule on all twelve layers led to the adoption of scripting all structures, which had the additional benefit of parameterizing the structures. The particular 180nm RF process used in this thesis forbids large contiguous areas of metal on the top via layer, which was circumvented with waffling. However, the waffling constrains the design space since all waffling must have only metal above and below, or else it cannot be removed by etching. The waffling itself can be circumvented by switching from the RF 180nm process to the standard 180nm process, at the cost of RF device models and the ultra thick top metal layer, which is useful for inductors and large waveguides.

The first waveguide considered in this platform is the M6-V5M5V4 waveguide, which uses the passivation layer associated with the ultra thick top metal layer for forming the ridge. This waveguide was used in several tapeouts, and the best measured upper bound on waveguide loss was found to be 4.1 dB/cm at 1550 nm, 4.8 dB/cm at 780 nm, and 7.8 dB/cm at 635 nm. The measured loss at 1550 nm represents an improvement of several orders of magnitude over the best previously published result of 38 dB/cm for waveguide loss in unmodified bulk CMOS. The measured losses at 780 nm and 635 nm represent, to the best of our knowledge, the first measurements of waveguide loss at visible and near-visible wavelengths in unmodified bulk CMOS.

An attempt was made to fabricate modulators using mechanical movement for the modulation mechanism, a type of MEMS device. It was shown that after post-processing the modulators were bent upward, most likely due to intrinsic compressive stress in the back-end dielectrics, and the ridges of the waveguides peeled off with the waffling, rendering these modulators inoperable.

An attempt was also made to create a coherent receivers using MMI splitters and directional couplers, as well as using an MMI 90° hybrid. It was shown that the ridges of the V5M5-V4M4 waveguides were removed during the course of post-processing. The similarities of the V5M5-V4M4 waveguide to geometries that have been proven to survive post-processing, such as the V5M5-V4M4V3 or V5M5V4-M4V3 geometries, made it difficult to ascertain the exact cause of the ridge removal. Since the V5M5-V4M4 waveguide was chosen due to the inadequacy

of the V5M5-V4M4V3 geometry as a single-mode waveguide, this problem of ridge removal will need to be studied in future research.

Two different types of waveguide-photodiode couplers were demonstrated, a tapering coupler and a reflecting coupler. The tapering coupler holds promise for improving the responsivity and bandwidth of CMOS photodiodes through traveling wave designs, while the reflecting coupler holds promise for improving the bandwidth of CMOS photodiodes through a reduction in capacitance, with the photodiode's compactness additionally benefiting the design of large systems.

All of these results together constitute the initial investigation into the method of subtractive photonics in bulk CMOS. The upper bound on the loss of the M6-V5M5V4 waveguide is promising, but more work needs to be done in the following areas:

- Measuring the waveguide loss using the cutback method
- Measuring bend loss for standard, Euler, and single-sided slab bends
- Measuring edge coupling loss
- Studying the cause of ridge removal for the V5M5-V4M4 waveguide
- Finding an optimal waveguide geometry for single-mode behavior
- Designing a high speed modulator using deposited materials.

Given its largely unexplored nature, the large photonic design space, and the monolithic integrated electronics, subtractive photonics in bulk CMOS offers tremendous opportunity for future research.

6.2 Future Outlook

Future Devices

This thesis has demonstrated waveguides with significantly lower loss in unmodified bulk CMOS than previous publications. Waveguide coupled CMOS photodiodes have also been demonstrated for detecting light in visible and near-visible wavelengths. In the literature, Schottky photodiodes for detecting infrared light have been demonstrated with responsivities up to several mA/W [72], [103] using back side illumination. Front side illumination (FSI) results in lower responsivities, on the order of $\mu\text{A}/\text{W}$ [104], but it may be possible to use integrated waveguides to achieve higher FSI responsivities using novel waveguide-photodiode couplers. The

combination of waveguides and photodiodes implies that optical signals can be propagated with low loss across the chip and then detected to convert the optical signals into electrical signals. Clearly the platform is still missing a device to convert electrical signals into optical signals - a modulator.

An attempt at creating a fully integrated MEMS modulator was discussed in Section 3.6. Surveying the possibilities for alternatives in [38], it is clear that the best options for high speed modulation in the context of subtractive photonics are deposited electro-optic polymers or transparent conducting oxides. Electro-optic polymers offer the particular advantage for subtractive photonics that the material can be applied using spin-on techniques to fill channels and potentially create channel waveguides. Historically electro-optic polymers have suffered from poor reliability, but recent improvements in thermal stability have made polymers an attractive option [105]. Transparent conducting oxides offer the possibility of ultra-compact and high speed modulators, but since these materials are not normally deposited in liquid form, any modulators in the subtractive photonics platform will require careful design.

Advanced Nodes

Most of the work presented in this thesis applies the method of subtractive photonics to a 180 nm RF process. This node combines low-cost area with rapid turnaround to make experimenting easier, and the RF optimization provides an ultra thick top metal layer that is useful for fabricating low-loss multimode waveguides. However, subtractive photonics is a general method that is intended to be applicable to any process with metal interconnect, so it is worthwhile to consider the effects of moving to more advanced nodes.

Nodes newer than 65 nm have continued to use copper dual damascene interconnect [75], although cobalt and ruthenium have been introduced at the local interconnect level, and are expected to be introduced to intermediate interconnect as well [106]. Fortunately, there are a variety of etchants that can be used to remove cobalt and ruthenium without affecting glass [82], so the etching techniques used in this thesis continue to be applicable. Low-k dielectrics have not advanced as quickly as the industry hoped, with lightly porous organosilicate glass appearing at the 65 nm node and moderately porous organosilicate glass used in following nodes, with some of the more advanced nodes using air gaps on some layers [75]. Highly porous (> 30%) organosilicate glass has not been implemented due to mechanical

and chemical difficulties [75], [106]. The overall low rate of change with regards to the materials used in BEOL interconnect for CMOS is an advantage for designers using the method of subtractive photonics, as the loss results obtained in this thesis can likely be assumed to apply to more advanced nodes. However, designers must keep in mind the reduction in mechanical strength and adhesion that accompanies increasing porosity in low-k dielectrics.

Although the materials used in BEOL interconnect have not changed much, other aspects have. The continual advancement of lithography needed for patterning ever finer metal pitches in more advanced nodes benefits subtractive photonics through a reduction in sidewall roughness and the ability to pattern very small optical features. For example, narrower gaps in directional couplers enable the lengths to be decreased, and subwavelength structures can be used to create a variety of devices [107]. The other advancement is the increasing number of layers in BEOL interconnect, expected to increase to 19 layers in the future [106]. The local and intermediate interconnect layers are very thin and must be combined together for forming structures with sufficient mechanical strength. The global interconnect layers are thicker and can be used in a manner similar to the waveguides shown in this thesis. In all, the larger number of layers greatly increases the design space and provides flexibility to designers.

BIBLIOGRAPHY

- [1] “1947: Invention of the point-contact transistor.” Accessed: 2023-04-03. (), [Online]. Available: <https://www.computerhistory.org/siliconengine/invention-of-the-point-contact-transistor/>.
- [2] “1948: Conception of the junction transistor.” Accessed: 2023-04-03. (), [Online]. Available: <https://www.computerhistory.org/siliconengine/conception-of-the-junction-transistor/>.
- [3] D. Gabor, “A new microscopic principle,” *Nature*, vol. 161, no. 4098, pp. 777–778, 1948. DOI: 10.1038/161777a0.
- [4] “1958: All semiconductor "solid circuit" is demonstrated.” Accessed: 2023-04-03. (), [Online]. Available: <https://www.computerhistory.org/siliconengine/all-semiconductor-solid-circuit-is-demonstrated/>.
- [5] “1959: Practical monolithic integrated circuit concept patented.” Accessed: 2023-04-03. (), [Online]. Available: <https://www.computerhistory.org/siliconengine/practical-monolithic-integrated-circuit-concept-patented/>.
- [6] “1960: Metal oxide semiconductor (mos) transistor demonstrated.” Accessed: 2023-04-03. (), [Online]. Available: <https://www.computerhistory.org/siliconengine/metal-oxide-semiconductor-mos-transistor-demonstrated/>.
- [7] T. Maiman, “Stimulated optical radiation in ruby,” *Nature*, vol. 4736, pp. 493–494, 1960.
- [8] G. Smith, “The early laser years at hughes aircraft company,” *IEEE Journal of Quantum Electronics*, vol. 20, no. 6, pp. 577–584, 1984. DOI: 10.1109/JQE.1984.1072445.
- [9] R. N. Hall, G. E. Fenner, J. Kingsley, T. Soltys, and R. Carlson, “Coherent light emission from gaas junctions,” *Physical Review Letters*, vol. 9, no. 9, p. 366, 1962.
- [10] M. I. Nathan, W. P. Dumke, G. Burns, F. H. Dill Jr, and G. Lasher, “Stimulated emission of radiation from gaas p-n junctions,” *Applied Physics Letters*, vol. 1, no. 3, pp. 62–64, 1962.
- [11] T. M. Quist, R. H. Rediker, R. Keyes, *et al.*, “Semiconductor maser of gaas,” *Applied Physics Letters*, vol. 1, no. 4, pp. 91–92, 1962.
- [12] A. Yariv and R. Leite, “Dielectric-waveguide mode of light propagation in p-n junctions,” *Applied Physics Letters*, vol. 2, no. 3, pp. 55–57, 1963.

- [13] K. C. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," in *Proceedings of the Institution of Electrical Engineers, IET*, vol. 113, 1966, pp. 1151–1158.
- [14] S. E. Miller, "Integrated optics: An introduction," *The Bell System Technical Journal*, vol. 48, no. 7, pp. 2059–2069, 1969. DOI: 10.1002/j.1538-7305.1969.tb01165.x.
- [15] R. L. Van Tuyl, "The Early Days of GaAs ICs," in *2010 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, 2010, pp. 1–4. DOI: 10.1109/CSICS.2010.5619700.
- [16] C. A. Mead, "Schottky barrier gate field effect transistor," *Proceedings of the IEEE*, pp. 307–308, 1966.
- [17] "Explore Intel's history." Accessed: 2023-04-03. (), [Online]. Available: <https://timeline.intel.com/>.
- [18] R. Noyce and M. Hoff, "A history of microprocessor development at intel," *IEEE Micro*, vol. 1, no. 01, pp. 8–21, 1981.
- [19] A. Yariv, "Active integrated optics," in *Fundamental and applied laser physics. Proceedings of the symposium held at Esfahan, Iran, August 29 to September 5, 1971*, M. S. Feld, A. Javan, and N. Kurnit, Eds., Wiley, 1972.
- [20] "How a Small Class at Caltech Helped Launch a Computer Revolution." Accessed: 2023-04-03. (), [Online]. Available: <https://www.caltech.edu/about/news/how-a-small-class-at-caltech-helped-launch-a-computer-revolution>.
- [21] J. Goodman, F. Leonberger, S.-Y. Kung, and R. Athale, "Optical interconnections for vlsi systems," *Proceedings of the IEEE*, vol. 72, no. 7, pp. 850–866, 1984. DOI: 10.1109/PROC.1984.12943.
- [22] R. Soref and J. Lorenzo, "Single-crystal silicon: A new material for 1.3 and 1.6 μm integrated-optical components," *Electronics Letters*, vol. 21, no. 21, pp. 953–954, 1985.
- [23] R. Soref and J. Lorenzo, "All-silicon active and passive guided-wave components for $\lambda = 1.3$ and 1.6 μm ," *IEEE Journal of Quantum Electronics*, vol. 22, no. 6, pp. 873–879, 1986. DOI: 10.1109/JQE.1986.1073057.
- [24] R. Soref and B. Bennett, "Electrooptical effects in silicon," *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, pp. 123–129, 1987. DOI: 10.1109/JQE.1987.1073206.
- [25] "Company info." Accessed: 2023-08-04. (), [Online]. Available: https://www.tsmc.com/english/aboutTSMC/company_profile.
- [26] L. Liao, "Low loss polysilicon waveguides for silicon photonics," M.S. thesis, Massachusetts Institute of Technology, 1997. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/10230>.

- [27] M. R. Black, “Losses in polysilicon waveguides,” M.S. thesis, Massachusetts Institute of Technology, 1995. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/11209>.
- [28] J. Foresi, M. Black, A. Agarwal, and L. Kimerling, “Losses in polycrystalline silicon waveguides,” *Applied Physics Letters*, vol. 68, no. 15, pp. 2052–2054, 1996.
- [29] A. M. Agarwal, L. Liao, J. S. Foresi, M. R. Black, X. Duan, and L. C. Kimerling, “Low-loss polycrystalline silicon waveguides for silicon photonics,” *Journal of Applied Physics*, vol. 80, no. 11, pp. 6120–6123, Dec. 1996. DOI: 10.1063/1.363686. [Online]. Available: <https://doi.org/10.1063/1.363686>.
- [30] D. Edelstein, J. Heidenreich, R. Goldblatt, *et al.*, “Full copper wiring in a sub-0.25 μm CMOS ULSI technology,” in *International Electron Devices Meeting. IEDM Technical Digest*, 1997, pp. 773–776. DOI: 10.1109/IEDM.1997.650496.
- [31] C. S. Yoo, *Semiconductor Manufacturing Technology*. WORLD SCIENTIFIC, Mar. 2008. DOI: 10.1142/6123. [Online]. Available: <https://doi.org/10.1142/6123>.
- [32] “0.18-micron technology.” Accessed: 2023-08-04. (), [Online]. Available: https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_018micron.
- [33] “65nm technology.” Accessed: 2023-08-04. (), [Online]. Available: https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_65nm.
- [34] “Luxtera raises \$7 million round.” Accessed: 2023-03-28. (), [Online]. Available: <https://www.thestreet.com/technology/luxtera-raises-7-million-round-10007995>.
- [35] Wikipedia contributors. “Luxtera — Wikipedia, the free encyclopedia.” [Online; accessed 29-March-2023]. (2020), [Online]. Available: <https://en.wikipedia.org/w/index.php?title=Luxtera&oldid=993209404>.
- [36] A. Liu, R. Jones, L. Liao, *et al.*, “A high-speed silicon optical modulator based on a metal–oxide–semiconductor capacitor,” *Nature*, vol. 427, no. 6975, pp. 615–618, 2004.
- [37] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. Thomson, “Silicon optical modulators,” *Nature Photonics*, vol. 4, no. 8, pp. 518–526, 2010.
- [38] G. Sinatkas, T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, “Electro-optic modulation in integrated photonics,” *Journal of Applied Physics*, vol. 130, no. 1, p. 010901, 2021.

- [39] L. Kimerling, D. Ahn, A. Apsel, *et al.*, “Electronic-photonic integrated circuits on the CMOS platform,” in *Silicon photonics*, SPIE, vol. 6125, 2006, pp. 6–15.
- [40] J. Fedeli, R. Orobtcchouk, C. Seassal, and L. Vivien, “Integration issues of a photonic layer on top of a CMOS circuit,” in *Silicon Photonics*, SPIE, vol. 6125, 2006, pp. 97–111.
- [41] J. S. Orcutt, A. Khilo, M. A. Popović, *et al.*, “Demonstration of an Electronic Photonic Integrated Circuit in a Commercial Scaled Bulk CMOS Process,” in *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies*, Optica Publishing Group, 2008, CTuBB3. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=CLEO-2008-CTuBB3>.
- [42] C. W. Holzwarth, J. S. Orcutt, H. Li, *et al.*, “Localized Substrate Removal Technique Enabling Strong Confinement Microphotronics in Bulk Si CMOS Processes,” in *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies*, Optica Publishing Group, 2008, CThKK5. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=CLEO-2008-CThKK5>.
- [43] V. Stojanović, R. J. Ram, M. Popović, *et al.*, “Monolithic silicon-photonic platforms in state-of-the-art CMOS SOI processes,” *Optics Express*, vol. 26, no. 10, pp. 13 106–13 121, 2018.
- [44] I. A. Young, E. Mohammed, J. T. S. Liao, *et al.*, “Optical i/o technology for tera-scale computing,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 1, pp. 235–248, 2010. DOI: 10.1109/JSSC.2009.2034444.
- [45] H.-C. Ji, K. H. Ha, K. W. Na, *et al.*, “Bulk silicon photonic wire for one-chip integrated optical interconnection,” in *7th IEEE International Conference on Group IV Photonics*, 2010, pp. 96–98. DOI: 10.1109/GROUP4.2010.5643414.
- [46] N. Sherwood-Droz, A. Gondarenko, and M. Lipson, “Oxidized silicon-on-insulator (oxsoi) from bulk silicon: A new photonic platform,” *Optics Express*, vol. 18, no. 6, pp. 5785–5790, 2010.
- [47] J. S. Orcutt, A. Khilo, C. W. Holzwarth, *et al.*, “Nanophotonic integration in state-of-the-art CMOS foundries,” *Optics Express*, vol. 19, no. 3, pp. 2335–2346, Jan. 2011. DOI: 10.1364/OE.19.002335. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-19-3-2335>.
- [48] J. S. Orcutt, S. D. Tang, S. Kramer, *et al.*, “Low-loss polysilicon waveguides fabricated in an emulated high-volume electronics process,” *Optics Express*, vol. 20, no. 7, pp. 7243–7254, 2012.
- [49] R. Meade, O. Tehar-Zahav, Z. Sternberg, *et al.*, “Integration of silicon photonics in a bulk CMOS memory flow,” in *2013 Optical Interconnects Conference*, 2013, pp. 114–115. DOI: 10.1109/OIC.2013.6552950.

- [50] H. Byun, J. Bok, K. Cho, *et al.*, “Bulk-si photonics technology for DRAM interface [invited],” *Photonics Research*, vol. 2, no. 3, A25, Apr. 2014. DOI: 10.1364/prj.2.000a25. [Online]. Available: <https://doi.org/10.1364/prj.2.000a25>.
- [51] R. Meade, J. S. Orcutt, K. Mehta, *et al.*, “Integration of silicon photonics in bulk CMOS,” in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1–2. DOI: 10.1109/VLSIT.2014.6894427.
- [52] C. Sun, M. Georgas, J. Orcutt, *et al.*, “A Monolithically-Integrated Chip-to-Chip Optical Link in Bulk CMOS,” *IEEE Journal of Solid-State Circuits*, vol. 50, no. 4, pp. 828–844, 2015. DOI: 10.1109/JSSC.2014.2382101.
- [53] A. H. Atabaki, S. Moazeni, F. Pavanello, *et al.*, “Integrating photonics with silicon nanoelectronics for the next generation of systems on a chip,” *Nature*, vol. 556, no. 7701, pp. 349–354, Apr. 2018. DOI: 10.1038/s41586-018-0028-z. [Online]. Available: <https://doi.org/10.1038/s41586-018-0028-z>.
- [54] P. Chaisakul, D. Marris-Morini, J. Frigerio, *et al.*, “Integrated germanium optical interconnects on silicon substrates,” *Nature Photonics*, vol. 8, no. 6, pp. 482–488, 2014.
- [55] Y.-J. Hung, C.-T. Tang, T.-H. Chen, T.-H. Yen, M.-J. Tsai, and S.-L. Lee, “Low-loss polysilicon subwavelength grating waveguides and narrowband Bragg reflectors in bulk CMOS,” *Optics Express*, vol. 28, no. 6, pp. 7786–7798, 2020.
- [56] M. Smit, K. Williams, and J. Van Der Tol, “Past, present, and future of inp-based photonic integration,” *APL Photonics*, vol. 4, no. 5, 2019.
- [57] Accessed: 2023-08-5. (), [Online]. Available: <https://www.inpho.io/products>.
- [58] M. Streshinsky, R. Ding, Y. Liu, *et al.*, “The road to affordable, large-scale silicon photonics,” *Optics and Photonics News*, vol. 24, no. 9, p. 32, Sep. 2013. DOI: 10.1364/opn.24.9.000032. [Online]. Available: <https://doi.org/10.1364/opn.24.9.000032>.
- [59] T. Pinguet, S. Denton, S. Gloeckner, *et al.*, “High-volume manufacturing platform for silicon photonics,” *Proceedings of the IEEE*, vol. 106, no. 12, pp. 2281–2290, 2018. DOI: 10.1109/JPROC.2018.2859198.
- [60] P. De Dobbelaere, A. Dahl, A. Mekis, *et al.*, “Advanced silicon photonics technology platform leveraging a semiconductor supply chain,” in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 34.1.1–34.1.4. DOI: 10.1109/IEDM.2017.8268493.

- [61] A. Elsherbini, S. Liff, J. Swan, K. Jun, S. Tiagaraj, and G. Pasdast, “Hybrid bonding interconnect for advanced heterogeneously integrated processors,” in *2021 IEEE 71st Electronic Components and Technology Conference (ECTC)*, 2021, pp. 1014–1019. DOI: 10.1109/ECTC32696.2021.00166.
- [62] V. Stojanović, K. T. Settaluri, S. Lin, *et al.*, “High-density 3d electronic-photonic integration,” in *2015 Fourth Berkeley Symposium on Energy Efficient Electronic Systems (E3S)*, 2015, pp. 1–2. DOI: 10.1109/E3S.2015.7336803.
- [63] M. Rakowski, C. Meagher, K. Nummy, *et al.*, “45nm CMOS-silicon photonics monolithic technology (45CLO) for next-generation, low power and high speed optical interconnects,” in *Optical Fiber Communication Conference*, Optica Publishing Group, 2020, T3H–3.
- [64] V. M. Stojanovic, “Bulk CMOS photonic/electronic integration (Conference Presentation),” in *Silicon Photonics XIV*, SPIE, vol. 10923, 2019, p. 1 092 302.
- [65] “SiGe BiCMOS and Silicon Photonics Technologies.” Accessed: 2023-08-05. (), [Online]. Available: <https://www.ihp-microelectronics.com/services/research-and-prototyping-service/mpw-prototyping-service/sigec-bicmos-technologies>.
- [66] “Scalable production from a few dozens to several thousands of ics.” Accessed: 2023-08-05. (), [Online]. Available: <https://www.ihp-solutions.com/index.php?id=34&L=0#production>.
- [67] R. J. Baker, *CMOS: Circuit design, layout, and simulation*. John Wiley & Sons, 2019.
- [68] Y. H. D. Lee and M. Lipson, “Back-End Deposited Silicon Photonics for Monolithic Integration on CMOS,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 2, pp. 8 200 207–8 200 207, 2013. DOI: 10.1109/JSTQE.2012.2209865.
- [69] J. D. Bradley, C. C. Evans, J. T. Choy, *et al.*, “Submicrometer-wide amorphous and polycrystalline anatase tio2 waveguides for microphotonic devices,” *Optics Express*, vol. 20, no. 21, pp. 23 821–23 831, 2012.
- [70] I. Hegeman, M. Dijkstra, F. Segerink, W. Lee, and S. Garcia-Blanco, “Development of low-loss tio2 waveguides,” *Optics Express*, vol. 28, no. 5, pp. 5982–5990, 2020.
- [71] M.-J. Lee and W.-Y. Choi, “Performance Optimization and Improvement of Silicon Avalanche Photodetectors in Standard CMOS Technology,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 2, pp. 1–13, 2018. DOI: 10.1109/JSTQE.2017.2754359.

- [72] W. Diels, M. Steyaert, and F. Tavernier, "Schottky Photodiodes in Bulk CMOS for High-Speed 1310/1550 nm Optical Receivers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 6, pp. 1–8, 2018. DOI: 10.1109/JSTQE.2018.2827667.
- [73] R. Fatemi, C. Ives, A. Khachaturian, and A. Hajimiri, "Subtractive photonics," *Optics Express*, vol. 29, no. 2, pp. 877–893, 2021. DOI: 10.1364/OE.410139,
- [74] S. Shekhar, "Silicon photonics: A brief tutorial," *IEEE Solid-State Circuits Magazine*, vol. 13, no. 3, pp. 22–32, 2021. DOI: 10.1109/MSSC.2021.3088966.
- [75] M. Rudan, R. Brunetti, and S. Reggiani, Eds., *Springer Handbook of Semiconductor Devices*. Springer International Publishing, 2023. DOI: 10.1007/978-3-030-79827-7. [Online]. Available: <https://doi.org/10.1007/978-3-030-79827-7>.
- [76] Y. Lian, *Semiconductor Microchips and Fabrication*. Wiley, Oct. 2022. DOI: 10.1002/9781119867814. [Online]. Available: <https://doi.org/10.1002/9781119867814>.
- [77] S. P. Murarka, "Multilevel interconnections for ULSI and GSI era," *Materials science and engineering: R: Reports*, vol. 19, no. 3-4, pp. 87–151, 1997.
- [78] "Aluminum etchants." Accessed: 2023-04-11. (), [Online]. Available: <https://transene.com/aluminum/>.
- [79] Accessed: 2023-04-11. (), [Online]. Available: <https://www.cnfusers.cornell.edu/sites/default/files/Area-Resources/Etchant%5C%20Datasheets.pdf>.
- [80] "Tantalum wet etching for photolithography." Accessed: 2019-08-15. (), [Online]. Available: <https://www.finishing.com/56/48.shtml>.
- [81] "Etching solution for tantalum/tantalum nitride." Accessed: 2019-08-15. (), [Online]. Available: <https://www.finishing.com/178/71.shtml>.
- [82] "Etchant/metal compatibility chart." Accessed: 2023-05-18. (), [Online]. Available: <https://transene.com/etch-compatibility/>.
- [83] A. Yariv, *Optical electronics in modern communications*. Oxford University Press, USA, 2007, vol. 1.
- [84] M. Cherchi, S. Ylinen, M. Harjanne, M. Kapulainen, and T. Aalto, "Dramatic size reduction of waveguide bends on a micron-scale silicon photonic platform," *Optics Express*, vol. 21, no. 15, pp. 17 814–17 823, 2013.
- [85] F. Ladouceur and E. Labeye, "A new general approach to optical waveguide path design," *Journal of Lightwave Technology*, vol. 13, no. 3, pp. 481–492, 1995. DOI: 10.1109/50.372446.

- [86] W. Chang, S. Jang, C. Yu, S. Sun, and M. Liang, "A manufacturable and reliable low-k inter-metal dielectric using fluorinated oxide (fsg)," in *Proceedings of the IEEE 1999 International Interconnect Technology Conference (Cat. No.99EX247)*, 1999, pp. 131–133. DOI: 10.1109/IITC.1999.787099.
- [87] C. Pai, A. Velaga, W. Lindenberger, *et al.*, "A manufacturable embedded fluorinated SiO₂ for advanced 0.25 μ m CMOS VLSI multilevel interconnect applications," in *Proceedings of the IEEE 1998 International Interconnect Technology Conference (Cat. No.98EX102)*, 1998, pp. 39–41. DOI: 10.1109/IITC.1998.704745.
- [88] C. Chiang, M. Chen, Z. Wu, *et al.*, "TDDB reliability improvement in Cu damascene by using a bilayer-structured PECVD SiC dielectric barrier," in *Proceedings of the IEEE 2002 International Interconnect Technology Conference (Cat. No.02EX519)*, 2002, pp. 200–202. DOI: 10.1109/IITC.2002.1014933.
- [89] H. Philipp, "Silicon dioxide (SiO₂)(glass)," in *Handbook of optical constants of solids*, Elsevier, 1997, pp. 749–763.
- [90] N. Quack, H. Sattari, A. Y. Takabayashi, *et al.*, "Mems-enabled silicon photonic integrated devices and circuits," *IEEE Journal of Quantum Electronics*, vol. 56, no. 1, pp. 1–10, 2020. DOI: 10.1109/JQE.2019.2946841.
- [91] M. Bao, *Analysis and design principles of MEMS devices*. Elsevier, 2005.
- [92] A. Evans and J. Hutchinson, "The thermomechanical integrity of thin films and multilayers," *Acta Metallurgica et Materialia*, vol. 43, no. 7, pp. 2507–2530, 1995.
- [93] J.-G. Su, H.-M. Hsu, S.-C. Wong, C.-Y. Chang, T.-Y. Huang, and J. Y.-C. Sun, "Improving the RF performance of 0.18 μ m CMOS with deep n-well implantation," *IEEE Electron Device Letters*, vol. 22, no. 10, pp. 481–483, 2001. DOI: 10.1109/55.954918.
- [94] Y.-M. Sheu, S.-J. Yang, C.-C. Wang, *et al.*, "Modeling mechanical stress effect on dopant diffusion in scaled mosfets," *IEEE Transactions on Electron Devices*, vol. 52, no. 1, pp. 30–38, 2005. DOI: 10.1109/TED.2004.841286.
- [95] S. M. Sze, Y. Li, and K. K. Ng, *Physics of semiconductor devices*. John Wiley & Sons, 2021.
- [96] H. Zimmermann and H. Zimmermann, *Integrated silicon optoelectronics*. Springer, 2010, vol. 148.
- [97] F. Tavernier and M. Steyaert, *High-speed optical receivers with integrated photodiode in nanoscale CMOS*. Springer Science & Business Media, 2011.
- [98] S. Donati, *Photodetectors: Devices, circuits and applications*. Prentice Hall, 2000.

- [99] S. Radovanovic, A.-J. Annema, and B. Nauta, "A 3-Gb/s optical detector in standard CMOS for 850-nm optical communication," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 8, pp. 1706–1717, 2005. DOI: 10.1109/JSSC.2005.852030.
- [100] C. Hermans and M. Steyaert, "A high-speed 850-nm optical receiver front-end in 0.18- μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 7, pp. 1606–1614, 2006. DOI: 10.1109/JSSC.2006.873855.
- [101] "Silicon ionization coefficients." Accessed: 2023-07-18. (), [Online]. Available: https://impact-ionisation.group.shef.ac.uk/ionisation_coeff/Si/.
- [102] K. Giboney, M. Rodwell, and J. Bowers, "Traveling-wave photodetectors," *IEEE Photonics Technology Letters*, vol. 4, no. 12, pp. 1363–1365, 1992. DOI: 10.1109/68.180577.
- [103] W. Diels, M. Steyaert, and F. Tavernier, "1310/1550 nm Optical Receivers With Schottky Photodiode in Bulk CMOS," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 7, pp. 1776–1784, 2020. DOI: 10.1109/JSSC.2020.2991517.
- [104] W. Diels, M. Steyaert, and F. Tavernier, "Modelling, design and characterization of Schottky diodes in 28nm bulk CMOS for 850/1310/1550nm fully integrated optical receivers," in *2017 47th European Solid-State Device Research Conference (ESSDERC)*, 2017, pp. 224–227. DOI: 10.1109/ESSDERC.2017.8066632.
- [105] C. Eschenbaum, A. Mertens, C. Füllner, *et al.*, "Thermally stable silicon-organic hybrid (soh) mach-zehnder modulator for 140 gbd pam4 transmission with sub-1 v drive signals," in *2022 European Conference on Optical Communication (ECOC)*, 2022, pp. 1–4.
- [106] International Roadmap for Devices and Systems, "More Moore 2022 Update," International Roadmap for Devices and Systems, Tech. Rep., 2022.
- [107] R. Halir, A. Ortega-Moñux, D. Benedikovic, *et al.*, "Subwavelength-grating metamaterial structures for silicon photonic devices," *Proceedings of the IEEE*, vol. 106, no. 12, pp. 2144–2157, 2018. DOI: 10.1109/JPROC.2018.2851614.