Low-loss and nonlinear silicon-based integrated photonic circuits

Jean-Etienne Tremblay

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2021-234
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2021/EECS-2021-234.html

December 1, 2021

Copyright © 2021, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Low-loss and nonlinear silicon-based integrated photonic circuits

by

Jean-Étienne Tremblay

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ming C. Wu, Chair
Professor Eli Yablonovitch
Professor Holger Müller

Fall 2020

Low-loss and nonlinear silicon-based integrated photonic circuits

Copyright 2020
by
Jean-Étienne Tremblay

Abstract

Low-loss and nonlinear silicon-based integrated photonic circuits

by

Jean-Étienne Tremblay

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ming C. Wu, Chair

Photonic integrated circuits enable tight integration of optical components used as building blocks for much larger optical systems, similarly to electronic integrated circuits. In particular, silicon-based photonics can leverage the yield and scaling technology improvements that also benefit the electronics industry. Moreover, photonic integrated circuits can reduce the power consumption in communication and computation applications, and open the door to new applications in sensing and ranging.

The design of linear and nonlinear passive photonic components is key to the performance improvement of integrated photonics. In the first part of this dissertation, we present a platform for creating on-chip frequency combs using materials with a large third-order nonlinear index. Specifically, using chalcogenide glass waveguides, we produce a frequency comb spanning greater than an octave bandwidth using input optical pulses on the picojoule level. In the second part, we discuss fabrication processes and optical designs to improve the fiber-to-chip coupling loss, which is a limiting factor in large-scale optical switch applications. We present the design of an adiabatic evanescent coupler that is polarization diverse and has a bandwidth greater than 100 nm in the O band. Our design achieves better than 2 dB insertion loss and is compatible with different photonic fabrication processes.

To my parents

Contents

Contents ii

List of Figures iv

List of Tables vii

1 Introduction 1

2 On-chip optical frequency synthesis 2
- 2.1 Optical frequency combs . . . 2
- 2.2 Nonlinear Schrödinger equation . . . 4
- 2.3 Materials . . . 5
- 2.4 Supercontinuum generation . . . 6
- 2.5 Simulation . . . 7
- 2.6 Coherence . . . 9
- 2.7 Efficient octave-spanning supercontinuum in chalcogenide glass . . . 10

3 Fabrication techniques 18
- 3.1 Lithography . . . 18
- 3.2 Propagation loss . . . 25
- 3.3 Chalcogenide glasses . . . 26
- 3.4 Silicon nitride . . . 28
- 3.5 Oxide cladding . . . 30
- 3.6 Silicon . . . 31

4 Adiabatic component design 33
- 4.1 Adiabatic theorem and mode evolution . . . 33
- 4.2 Simulation of adiabatic devices . . . 35
- 4.3 Simple tapers . . . 35
- 4.4 Bent waveguides . . . 38
- 4.5 TE0-TE1 mode converter . . . 40
- 4.6 Spectral filters . . . 41
- 4.7 Optimization of nonlinear tapers . . . 44

5 Edge coupling **46**
5.1 Chip-to-chip edge coupling . . . 46
5.2 Chip-to-fiber edge coupling . . . 50
5.3 Substrate leakage . . . 54

6 Silicon photonics MEMS switch packaging **56**
6.1 4×4 silicon photonics MEMS switch . . . 57
6.2 Electrical packaging . . . 59
6.3 Optical packaging . . . 61
6.4 Alignment strategies . . . 63
6.5 Adhesives . . . 66

7 Evanescent coupling **68**
7.1 Literature review on evanescent couplers . . . 69
7.2 Platforms for intermediate chip . . . 70
7.3 Evanescent coupler design in silicon . . . 70
7.4 Evanescent coupler design in silicon nitride . . . 72
7.5 Splitting taper in silicon . . . 73
7.6 Oxidized taper evanescent coupler . . . 77
7.7 Evanescent coupler alignment . . . 85

8 Conclusion **86**

Bibliography **87**

A High voltage parallel MEMS drivers **95**
A.1 4×4 driver with microcontroller . . . 95
A.2 4×4 driver with FPGA . . . 97
A.3 16×16 driver with FPGA . . . 99

B FiconTEC FL300 procedures **101**
B.1 Components . . . 101
B.2 Programs . . . 105
B.3 Alignment . . . 107

List of Figures

2.1 Optical frequency comb representations . . . 3
2.2 Self-referenced frequency comb system . . . 3
2.3 On-chip frequency comb concept . . . 4
2.4 2nd order soliton propagation in a chalcogenide glass waveguide . . . 7
2.5 Cross section of the supercontinuum generation device . . . 11
2.6 Chalcogenide glass ring resonator measurement for determining propagation loss . . 12
2.7 Optical frequency domain reflectometry measurement of a 4 cm long waveguide . . . 12
2.8 $Ge_{23}Sb_7S_{70}$ material dispersion. . . . 13
2.9 Measured and simulated dispersion of the chalcogenide waveguide for the fundamental modes of the ring resonator . . . 14
2.10 Simulated dispersion of the chalcogenide waveguide as a function of width . . . 14
2.11 Propagation constant in the moving reference frame . . . 15
2.12 Simulated and measured supercontinuum generation . . . 16
2.13 Simulated coherence of the generated supercontinuum . . . 16

3.1 Lithography illumination . . . 22
3.2 Comparison of aerial image between conventional and annular illumination . . . 23
3.3 Standing wave patterns in photoresist. . . . 24
3.4 Propagation loss from sidewall roughness as a function of waveguide width . . . 26
3.5 Chalcogenide waveguide cross section after top cladding deposition . . . 28
3.6 Photonic damascene process . . . 29
3.7 Photonic damascene stress release . . . 29
3.8 Cleaning of silicon grating coupler etch residues in dilute HF . . . 32

4.1 Wide taper length sweep . . . 36
4.2 Narrow taper length sweep . . . 37
4.3 Minimum feature size of vertical and lateral double tapers . . . 38
4.4 Refractive index transformation of a waveguide with 10 μm bend radius . . . 39
4.5 Waveguide bend shapes . . . 40
4.6 TE0-TE1 mode converter . . . 41
4.7 Modal dispersion between the fundamental and second TE mode in two waveguides of different widths . . . 42
4.8 Adiabatic spectral splitter concept . . . 42

4.9 Simulation of a 1×3 spectral splitter . . . 43
4.10 Adiabatic spectral splitter performance . . . 43
4.11 Optimized taper shape for the TE0-TE1 converter . . . 45
4.12 Length sweep for the optimized TE0-TE1 converter . . . 45

5.1 Optical chip dicing process . . . 47
5.2 Optical facet in SiO_2 after trench etching . . . 48
5.3 Silicon nitride chip edge coupled to an indium phosphide laser chip . . . 49
5.4 Mechanical carrier for chip-to-chip alignment . . . 50
5.5 Numerical aperture definition . . . 51
5.6 Waveguide matched to fiber fabricated with two thin SiN layer . . . 52
5.7 Edge spot size converter between silicon taper and single mode fiber . . . 52
5.8 Edge spot size converter fabrication process . . . 53
5.9 Mode size converter array layout . . . 53
5.10 Effect of cladding index on mode shape and substrate leakage . . . 55

6.1 Crossbar switch architecture . . . 56
6.2 Epoxy dispensed on 4×4 switch . . . 57
6.3 Fiber glued to 4×4 switch . . . 58
6.4 4×4 switch packaged inside box . . . 58
6.5 4×4 switch optical response time . . . 59
6.6 Coupled inductor boost converter . . . 60
6.7 Voltage doubled boost converter . . . 61
6.8 Optical coupling methods . . . 61
6.9 FiconTEC FL300 . . . 63
6.10 Gripper designs . . . 65
6.11 Gripper alignment considerations . . . 65
6.12 Fabricated gripper designs . . . 66

7.1 Wafer-scale silicon photonics switch concept . . . 69
7.2 Maximum silicon thickness as a function of taper width, for 95 % efficiency . . . 71
7.3 3 step silicon taper design . . . 72
7.4 Silicon nitride taper design . . . 73
7.5 Silicon separate TE/TM taper design . . . 74
7.6 TE tip overlap loss for a taper minimum width of 200 nm . . . 75
7.7 Design process of splitting taper evanescent coupler . . . 75
7.8 E-field propagation along the splitting taper . . . 76
7.9 Taper length sweep for splitting taper design using eigenmode expansion . . . 76
7.10 Splitting taper evanescent coupler fabrication tolerance . . . 77
7.11 Silicon taper oxidation process . . . 78
7.12 Transition between full etch and partial etch region in the oxidized taper process . . . 79
7.13 Oxidation simulation using Sentaurus . . . 79

7.14 Oxidized silicon taper design space . . . 80
7.15 Taper length sweep of oxidized taper evanescent coupler . . . 81
7.16 Tolerance of oxidized tip evanescent to epoxy gap thickness . . . 81
7.17 Glass interposer leakage loss as a function of epoxy index . . . 82
7.18 Fabricated oxidized waveguide . . . 83
7.19 Cross section of oxidized silicon taper . . . 83
7.20 Fiber to silicon insertion loss as a function of wavelength for evanescent coupler with oxidized tip . . . 84
7.21 Potential improvements to evanescent coupler loss . . . 85
7.22 Interference fringes during evanescent coupler alignment . . . 85

A.1 Block schematic of 4×4 driver with microcontroller . . . 96
A.2 PCB of 4×4 driver with microcontroller . . . 96
A.3 Block schematic of 4×4 driver with FPGA . . . 98
A.4 PCB of 4×4 driver with FPGA . . . 98
A.5 Block schematic of 16×16 driver with FPGA . . . 99
A.6 PCB of 16×16 driver with FPGA . . . 100

B.1 Jog panel . . . 102
B.2 Camera panel . . . 102
B.3 Vision algorithm editor . . . 103
B.4 Shape editor . . . 103
B.5 DAQ panel . . . 104
B.6 Chucks and mounts . . . 105
B.7 Bottom view camera calibration tool . . . 106

List of Tables

2.1 Materials for Kerr nonlinear applications 6

4.1 Bending loss for narrow and wide waveguides of different shapes 40
4.2 TE0-TE1 mode converter dimensions 41

6.1 Adhesive selection 67

7.1 Breakdown of loss of individual components of separate taper evanescent coupler . . 77
7.2 Breakdown of loss of individual components of oxidized tip evanescent coupler . . . 84

A.1 High voltage driver PCB versions 95
A.2 Microcontroller commands for 4×4 switch 97

Acknowledgments

I would like to thank my great advisor, Professor Ming C. Wu, for all the guidance, support and discussions during my time at Berkeley. It has been a pleasure to work on all these exciting projects with him.

I want to acknowledge all my colleagues in the Wu lab, as well as my colleagues in the Yablonovitch and Chang-Hasnain group for all the extensive discussions and knowledge you shared with me: Andy, Phil, Jodi, Kevin, Jonas, Zain, and many others. A special thanks to Professor Eli Yablonovitch and Professor Holger Müller for serving on my dissertation committee.

None of this research would have been possible without the help of the Marvell Nanolab staff, who work hard to make sure this wonderful facility can run 24h per day, 7 days per week.

A special thanks to my friends who have made this journey manageable: Tristan, Derek and Antoine in California, and Yanick and Alexandre back in Montreal.

Thank you Camille, my wife, for all your encouragements and understanding.

Finally, thank you to my brother Pierre-Antoine, my sister Caroline, and my parents Sylvie and Laval, for your support and love.

Chapter 1

Introduction

More than twenty years ago, silicon photonics has emerged as a platform of choice for fabricating integrated photonic circuits. Thanks to the promising compatibility with CMOS fabrication process, massive investment and research has helped leveraging state of the art fabrication facilities to develop photonic devices with great scalability [1, 2].

Even though a huge library of optical components now exists in silicon photonics (e.g. modulators [3, 4] and photodetectors [5]), some devices such as lasers have proven difficult to fabricate in silicon. A promising more recent approach to provide light sources on silicon is through hybrid integration of other materials (such as III-V semiconductors [6]). Another area where silicon photonics can benefit from hybrid integration is nonlinear optics. Indeed, silicon has a large third-order nonlinear index, but suffers from significant two-photon absorption in the ubiquitous telecom bands [7]. In this thesis we explore the use of chalcogenide glass materials for nonlinear applications, in particular supercontinuum generation. Finally, the challenge of coupling light from silicon to other chips or to optical fibers has seen numerous solutions, such as the original inverted nanotaper [8] to more recent vertical grating couplers [9]. In this dissertation, we will also discuss about silicon photonics packaging, from edge coupling to evanescent coupling.

The first part of this thesis covers supercontinuum generation applications. In Chapter 2, we describe how optical frequency combs work, and how they can be generated on-chip. We optimize a chalcogenide glass waveguide for generating a supercontinuum spanning more than one octave, using the nonlinear Schrödinger equation. The details of the fabrication for the nonlinear waveguide are given in Chapter 3, as well as other useful fabrication techniques for nonlinear and low-loss applications. Chapter 4 covers the design of adiabatic components, which have many applications such as mode conversion, spot size converters, and polarization rotation. In particular, we present an adiabatic spectral splitter operating over an octave bandwidth. Chapter 5 discusses the packaging of optical frequency comb waveguides using chip-to-chip edge coupling.

The second part covers packaging designs and techniques for silicon photonics MEMS switch applications. In Chapter 6, we give the details of a fully packaged 4×4 optical switch, including automated alignment and electrical control. Finally, Chapter 7 is dedicated to several designs of evanescent couplers, which can be used to package large fiber arrays to wafer-scale silicon photonics switches in a simple process with low insertion loss.

Chapter 2

On-chip optical frequency synthesis

In this section we first present optical frequency combs and how they can be self-referenced to produce accurate absolute frequencies. We then study the basic theory of nonlinear optical propagation, in particular third-order effects which are present in most materials. Third-order nonlinearities are the source of phenomena such as self-phase modulation, Kerr focusing, four-wave mixing, and supercontinuum generation. We explain how noise affects supercontinuum operation and we show the results of a low power, octave-spanning, coherent supercontinuum generated in a chalcogenide glass waveguide.

2.1 Optical frequency combs

Optical frequency combs are simply an ensemble of optical frequencies with equal spacing f_r defined by

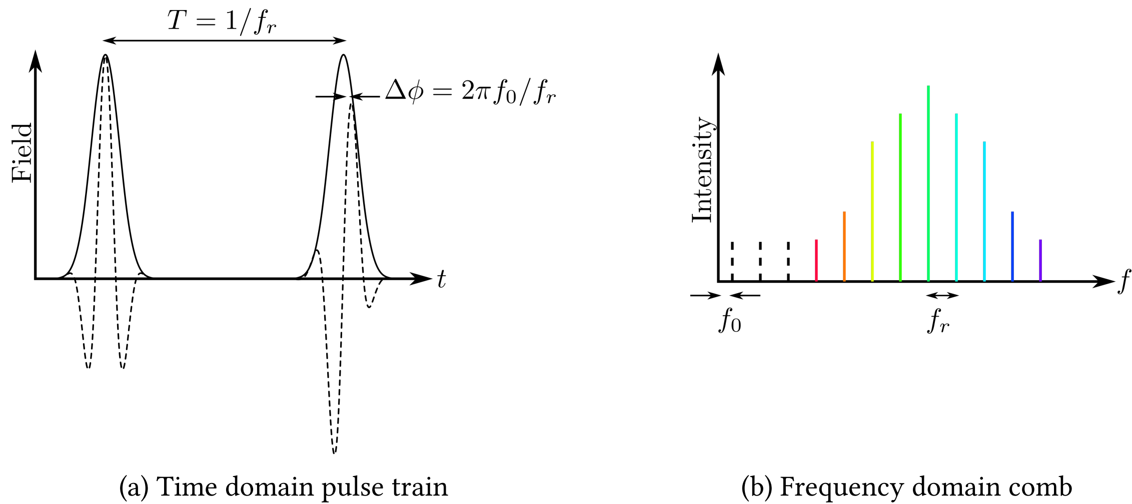
$$f_n = f_0 + nf_r \tag{2.1}$$

If all the optical frequencies are coherent in phase, the time domain representation of a frequency comb will be a series of short pulse with equal spacing of $T = 1/f_r$, the repetition period. The time and frequency domain representation of an optical frequency comb are shown in Figure 2.1. It should be noted that the frequency comb lines near 0 Hz do not necessarily physically exist in the optical frequency comb spectrum, and the minimum time domain pulse width is related to the optical bandwidth of the frequency comb spectrum. The frequency offset of the frequency comb, f_0, is related in the time domain to the carrier-envelope phase offset that is picked up in between each pulse by

$$\Delta\phi = 2\pi\frac{f_0}{f_r} \tag{2.2}$$

While each comb line frequency is in the optical domain (> 100 THz), the repetition rate f_r is typically in the microwave domain. The repetition rate can then easily be measured by reading the pulse train signal from a photodiode.

In some cases, it is required to precisely know the absolute frequency of each comb line, or to know the exact phase relationship between each pulse. Such applications include short (single-



(a) Time domain pulse train

(b) Frequency domain comb

Figure 2.1: Optical frequency comb representations

cycle) pulse generation, high accuracy spectroscopy, or the use of optical frequency standards to generate accurate clocks. The offset frequency f_0 cannot be directly measured like the repetition rate since it does not physically exist in the optical signal, however it can be measured through the help of a nonlinear interferometer [10, 11], which then creates a self-referenced frequency comb. This is represented in Figure 2.2.

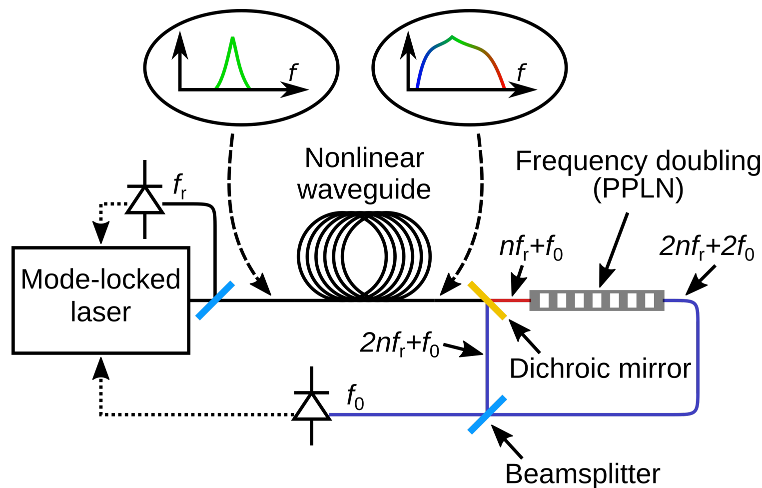


Figure 2.2: Self-referenced frequency comb system

In its simplest form, the self-referenced frequency comb is initiated from a mode-locked laser, which produces a series of optical pulses but with limited bandwidth. The optical bandwidth of the pulses is greatly increased in a nonlinear waveguide through a process known as *supercontinuum generation*, to create a frequency comb that has more than one octave bandwidth (we will now use the terms supercontinuum and frequency comb interchangeably). The supercontinuum

now contains the frequencies $f_0 + nf_r$ and $f_0 + 2nf_r$, which we refer to as the f and $2f$ comb lines. We then frequency double the f lines and mix the result with the $2f$ line to obtain the following

$$f_0 = 2(f_0 + nf_r) - (f_0 + 2nf_r) \tag{2.3}$$

It should be noted that other schemes do exist, for example $2f - 3f$ [12] when the frequency comb is of limited bandwidth, or $f - 3f$ [13] when a suitable material for frequency doubling is not available. When the optical power in the $2f$ or $3f$ line is not sufficient for second or third-harmonic generation, techniques such as amplification [14] or transfer lasers [15] can be used to increase the available signal.

In this work, we explore the use of chalcogenide glass to fabricate integrated waveguides for on-chip supercontinuum generation. This nonlinear waveguide chip is meant to be integrated in a larger self-referenced frequency comb system as shown in Figure 2.3.

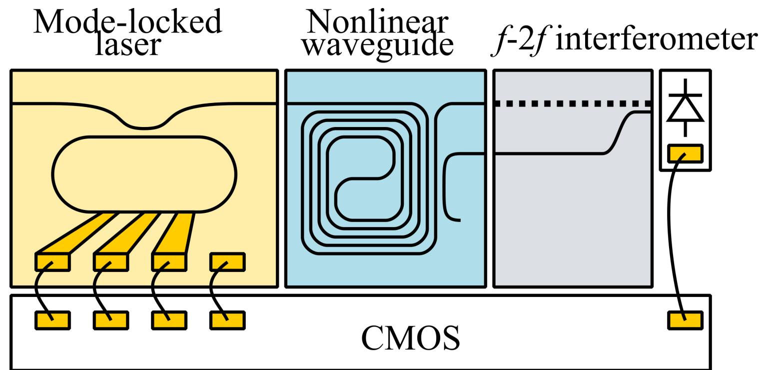


Figure 2.3: On-chip frequency comb concept

In the on-chip frequency comb device, the mode-locked laser is fabricated on indium phosphide [16], the $f - 2f$ interferometer is fabricated on lithium niobate [17], and the CMOS control electronics are fabricated using a conventional commercial process.

2.2 Nonlinear Schrödinger equation

Typically, optical nonlinearities are excited using pulsed light to achieved great optical intensities without requiring significant average power. The principal way of describing the third-order nonlinear propagation of light is by considering the slowly-varying envelope approximation to derive what is known as the nonlinear Schrödinger equation [18]:

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A - i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial T^2} + i\gamma\left(|A|^2 A + \frac{i}{\omega_0}\frac{\partial}{\partial T}\left(|A|^2 A\right) - T_R A\frac{\partial |A|^2}{\partial T}\right) \tag{2.4}$$

where $A(z, T)$ is the envelope of the propagating pulse, α is the linear propagation loss, β_2 is the dispersion, γ is the nonlinearity coefficient, ω_0 the carrier frequency and T_R the slope of the

Raman gain. A will typically be a complex number to account for phase. The Fourier transform in time of A represents the spectral content of the pulse, shifted by the optical carrier frequency. The left side of the equation represents the evolution of the pulse as it propagates in the z direction. The time T is centered around the pulse through a change of variables removing the group delay.

$$T = t - \frac{z}{v_g} \tag{2.5}$$

The right side is divided in linear terms and nonlinear terms. The linear terms accounts for loss and dispersion. If desired, the dispersion can be expanded in more terms, or treated in the frequency domain to account for more complicated dispersion relations.

The nonlinear terms are multiplied by γ, the nonlinearity coefficient. It is calculated from

$$\gamma = \frac{n_2 \omega_0}{c A_{\text{eff}}} = \frac{2\pi n_2}{\lambda A_{\text{eff}}} \tag{2.6}$$

where n_2 is the nonlinear index, w_0 the optical carrier frequency and A_{eff} is the effective nonlinear mode area. The first nonlinear term in the nonlinear Schrödinger equation represents the nonlinear phase shift (Kerr effect). The second term comes from the inclusion of the first derivative in the slowly varying envelope approximation and is responsible for self-steepening of the pulse. The last term accounts for the nonlinear Raman response (T_{R} is the slope of the Raman gain) and is responsible for self-frequency shift of the pulse.

It should be noted that, in the slowly varying approximation, the high frequency terms accounting for third harmonic generation were dropped, since they are usually not of interest for supercontinuum generation. Nonetheless, the nonlinear Schrödinger equation can account for many third order nonlinear effects, such as the Kerr effect, four-wave mixing, self-phase modulation, modulation instability, and so-on.

2.3 Materials

The main characteristic used to compare materials for Kerr nonlinear applications is the nonlinear index n_2. However, many materials also have two-photon absorption β_{TPA} which is a 3rd order nonlinear effect as well. The ratio of the nonlinear index and nonlinear loss is included in a figure of merit which is commonly cited in the literature:

$$FOM_{\text{TPA}} = \frac{n_2}{\lambda \beta_{\text{TPA}}} \tag{2.7}$$

A figure of merit larger than unity is advantageous for supercontinuum generation. For silicon photonics, common nonlinear material candidates are silicon and silicon nitride (Si_3N_4). Another class of promising materials are chalcogenide glasses [19]. In particular, we are interested in the specific composition $Ge_{23}Sb_7S_{70}$. These three materials are compared in Table 2.1. We also include SiO_2 as a reference, since it is well studied in fiber systems. The table also summarizes other

Material	Loss (dB/cm)	Nonlin. loss (cm/GW)	Index	n_2 (10^{-20} m^2/W)	A_{eff} (μm^2)	γ (1/Wm)	FOM
$Ge_{23}Sb_7S_{70}$ [20, 21]	0.5	0.01	2.2	93	1	7	6
Si_3N_4 [22]	0.1	~ 0	2.0	24	1	1	$\gg 1$
Silicon [23]	1.0	0.8	3.47	400	0.1	160	0.32
SiO_2 [18]	0.01	~ 0	1.45	2.7	10	0.01	$\gg 1$

Table 2.1: Materials for Kerr nonlinear applications

properties such as typical waveguide loss, refractive index, waveguide mode area and typical nonlinear parameter γ.

Even though silicon has a very large nonlinear index, the large two-photon absorption prohibits its use for supercontinuum generation in the C band. It has however successfully been used in the mid-infrared [24]. Silicon nitride and chalcogenide glasses are both good candidates for on-chip supercontinuum generation. However, the larger nonlinear index of $Ge_{23}Sb_7S_{70}$ means that the nonlinear interaction occurs on a shorter length and the device footprint will be minimized.

2.4 Supercontinuum generation

A *supercontinuum* is generally defined as beam of light which underwent several nonlinear processes to get massively spectrally broadened. One example of supercontinuum generation would be by pumping a photonic crystal fiber with the short pulses from a mode-locked laser. The spectral broadening can be on the order of one octave or more. Supercontinuum sources can be used for applications requiring broadband light, such as optical coherence tomography, imaging, and optical communication.

A related concept is the *optical frequency comb*. A frequency comb contains several, evenly spaced frequencies which have a well-defined phase relationship. This corresponds to a regular pulse train in the time domain. A frequency comb can be nonlinearly broadened to become a supercontinuum. On the other hand, a supercontinuum which consists of discrete frequency lines could be called a frequency comb.

The formation of supercontinuum can be understood by studying the nonlinear Schrödinger equation. If we simply include the dispersion and first nonlinear term in (2.4), and assuming $\beta_2 < 0$, we get solutions of the form

$$A = \sqrt{\frac{|\beta_2|}{T_0^2\gamma}}\,\text{sech}\left(\frac{T}{T_0}\right)\exp\left(i\frac{|\beta_2|}{2T_0^2}z\right) \tag{2.8}$$

This solution is called the fundamental *soliton*. The literature defines the dispersion and nonlinear characteristic lengths as following:

$$L_{\text{D}} = \frac{T_0^2}{|\beta_2|} \tag{2.9}$$

$$L_{\text{NL}} = \frac{1}{\gamma P_0} \tag{2.10}$$

where P_0 is the pulse peak power. For the fundamental soliton, we notice that $P_0 = |\beta_2|/T_0^2\gamma$ and hence $L_{\text{D}} = L_{\text{NL}}$. The soliton propagates unperturbed, or in other words the nonlinearity and dispersion interact on the same length scale.

There also exists higher order solitons which are solutions of the nonlinear Schrödinger equation. These solutions are characterized by the soliton number N:

$$N^2 = \frac{L_{\text{D}}}{L_{\text{NL}}} \tag{2.11}$$

Instead of propagating unperturbed like the fundamental soliton, the higher order solitons change shape periodically with a soliton period

$$z_0 = \frac{\pi}{2} L_{\text{D}} \tag{2.12}$$

which is independent of the soliton order. The simulated propagation of a soliton with $N = 2$ in a chalcogenide glass waveguide is shown in Figure 2.4.

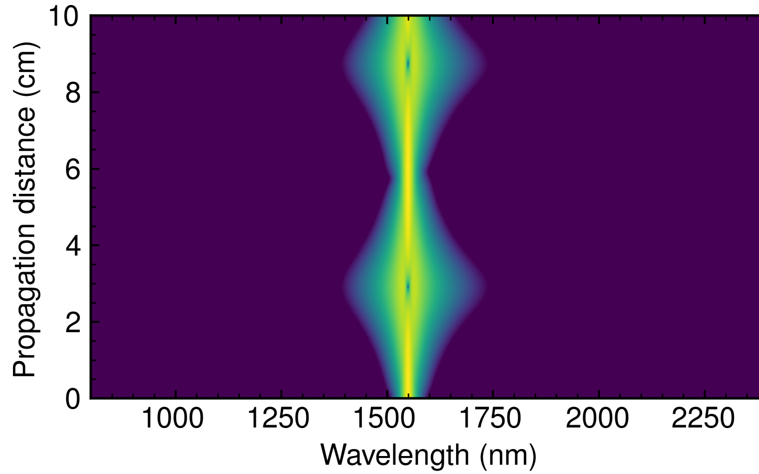


Figure 2.4: 2nd order soliton propagation in a chalcogenide glass waveguide. The soliton period is about 6 cm.

2.5 Simulation

The simulation of supercontinuum generation relies on the discretization of equation (2.4). The field amplitude A is propagated in discrete z steps. At each step, two contributions are calculated.

The linear term coming from loss and dispersion is calculated in the Fourier domain at each z step. This enabled direct multiplication of $\hat{A}(\omega) = \mathcal{F}(A)$ by the dispersion $\beta(\omega)$. The result is converted back in the time domain with an inverse Fourier transform. In other words, we replace the dispersion term in (2.4) with the more general

$$-\frac{\beta_2}{2}\frac{\partial^2 A}{\partial T^2} \rightarrow \mathcal{F}^{-1}\left(\beta(\omega)\hat{A}(\omega)\right) \tag{2.13}$$

assuming $\beta(\omega)$ is expanded as

$$\beta(\omega) = \frac{\beta_2}{2}(\omega-\omega_0)^2 + \frac{\beta_3}{6}(\omega-\omega_0)^3 + \cdots \tag{2.14}$$

The group delay per unit length $\frac{1}{v_g} = \beta_1$ is taken care of by the transformation of variable $T = t - \frac{z}{v_g}$. This will conveniently keep the pulse centered in the simulation window during propagation.

The nonlinear terms are computed directly in the time domain. The time derivatives can be calculated either by finite difference, or by multiplying by ω in the Fourier domain. There are several algorithms available for the z stepping. If we split the linear ($\hat{D}$) and nonlinear part ($\hat{N}$) of the nonlinear Schrödinger equation:

$$\frac{\partial A}{\partial z} = (\hat{D} + \hat{N})A \tag{2.15}$$

then the solution is

$$A(z+h) = \exp h(\hat{D} + \hat{N})A \tag{2.16}$$

Since $\hat{D}$ and $\hat{N}$ do not commute, the following would only be an approximation to the solution:

$$A(z+h) \approx \exp\left(h\hat{D}\right)\exp\left(h\hat{N}(z)\right)A(z) \tag{2.17}$$

As mentioned previously, the linear (dispersion) operator is evaluated in the Fourier domain and we can write this simple method more explicitly:

$$A(z+h) \approx \mathcal{F}^{-1}\exp\left(h\hat{D}(i\omega)\right)\mathcal{F}\exp\left(h\hat{N}(z)\right)A(z) \tag{2.18}$$

Moreover, since $\hat{N}$ is a function of z, this method can only be first order accurate. Several other schemes with higher accuracy have been developed. In particular, in this thesis we use the 4th order Runge-Kutta method in the interaction picture [25], which has 4th order accuracy while using a minimal number of Fourier transforms.

2.6 Coherence

During supercontinuum generation, not all nonlinear processes keep a known and predictable phase relationship between the frequencies that constitute the pulses. In particular, certain phenomena such as modulation instability amplify noise [26, 27].

The specific derivation of modulation instability gain can be found in [18]. Here, we are interested in understanding under which condition the coherence of the supercontinuum is preserved.

The first condition for modulation instability is anomalous dispersion ($\beta_2 < 0$), where the peak modulation instability gain is

$$g_{\mathrm{MI}} = 2\gamma P_0 \tag{2.19}$$

with P_0 the pulse peak power. The other condition for modulation instability to occur is that its gain is larger than the linear propagation loss α. Usually it is desirable to have waveguides with low propagation loss for maximum efficiency, so the modulation instability will not be prevented by the propagation loss.

On the other hand, the supercontinuum broadening accelerates around the soliton fission length z_{opt}, which is the approximate optimal length of the supercontinuum generating waveguide. Indeed, during the periodic soliton beating, there is a point where the pulse length is minimal and the bandwidth is maximal. This length is empirically related to the dispersion length by [28]

$$z_{\mathrm{opt}} \approx \frac{L_{\mathrm{D}}}{N} \approx \sqrt{\frac{T_0^2}{|\beta_2|\gamma P_0}} \tag{2.20}$$

In Figure 2.4, this corresponds to 3 cm of propagation where the pulse bandwidth is maximal.

From this, we can understand that shorter pulses require less peak power to generate a large bandwidth supercontinuum with the same propagation length. This will lead to smaller modulation instability gain and increased coherence.

We can compare the soliton fission length and the modulation instability characteristic length, which is proportional to the inverse of the modulation instability gain.

$$L_{\mathrm{MI}} \propto \frac{1}{\gamma P_0} \tag{2.21}$$

which corresponds to L_{NL} defined previously. In practice, we can estimate $L_{\mathrm{MI}} \approx 16 L_{\mathrm{NL}}$ [28] which corresponds to sufficient gain to amplify the quantum shot noise level to the typical pulse power level.

On the other hand, since the soliton fission length goes as L_{D}/N, a high degree of coherence is ensured when

$$16 L_{\mathrm{NL}} > \frac{L_{\mathrm{D}}}{N} \tag{2.22}$$

or $N < 16$. In other words, a low soliton number is required for greater coherence. For a given average power, reducing the pulse length yields a lower soliton number for the pulse, which once again confirms the intuition that a supercontinuum created from shorter pulses will be more coherent.

We can use these conditions to compare materials having a large nonlinear figure of merit such as GeSbS and SiN defined in Table 2.1. For a given (low) soliton number, we want to minimize the total propagation loss:

$$FOM_{\text{lin}} = \alpha z_{\text{opt}} \sim \alpha L_{\text{NL}} = \frac{\alpha}{\gamma P_0} \tag{2.23}$$

This is not a fundamental material property, but it can be useful to compare two waveguide platforms under the same input pulse conditions, when the two-photon absorption figure of merit comparison is not meaningful. For example, with $P_0 = 100$ W, assuming the waveguide dispersion is designed correctly, for GeSbS the linear figure of merit is 0.07, and for silicon nitride it is 0.1. In this case, in addition to having a smaller footprint compared to silicon nitride, the chalcogenide glass waveguide will produce a slightly more efficient supercontinuum given the same input pulse.

In simulation, the coherence properties of the supercontinuum are modelled by adding quantum limited noise with energy variance of one half photon per time bin and random phase, which corresponds to adding one half photon per mode in the frequency domain [27].

$$\sigma_{\delta A}^2 = \frac{h\nu}{2\delta t} \tag{2.24}$$

When converted to a spectral power density, this noise corresponds to the classical shot noise level,

$$S_{\delta P} = \bar{P} h\nu \tag{2.25}$$

which means that the quantum noise simulates the general case of shot noise.

As the pulse propagates in the simulation, the noise will affect the pulse via nonlinear interactions. The coherence of the output pulse is then evaluated with the modulus of the first-order coherence [26]:

$$\left| g_{12}^{(1)}(\lambda) \right| = \left| \frac{\left\langle E_1^*(\lambda) E_2(\lambda) \right\rangle}{\sqrt{\left\langle |E_1(\lambda)|^2 \right\rangle \left\langle |E_2(\lambda)|^2 \right\rangle}} \right| \tag{2.26}$$

The coherence is greater when $\left| g_{12}^{(1)} \right|$ is closer to 1, and the supercontinuum is incoherent when $\left| g_{12}^{(1)} \right| = 0$.

2.7 Efficient octave-spanning supercontinuum in chalcogenide glass

In this section we summarize the design of chalcogenide glass waveguides for supercontinuum generation, which is described in details in [29].

The devices are fabricated using the chalcogenide glass fabrication process described in Chapter 3. Additionally, in order to couple light efficiently without inducing damage to the facet, we

use silicon nitride spot size converters on the edge of the chip [30]. We have successfully fabricated the silicon nitride spot size converters in both the photonic damascene process, as well as using a regular deposition and etch process with plasma-enhanced chemical vapor deposition (PECVD) silicon nitride. A cross section of the device is shown in Figure 2.5. The top cladding is deposited using PECVD, which shows typical voids starting from the inside corners of the rib waveguide.

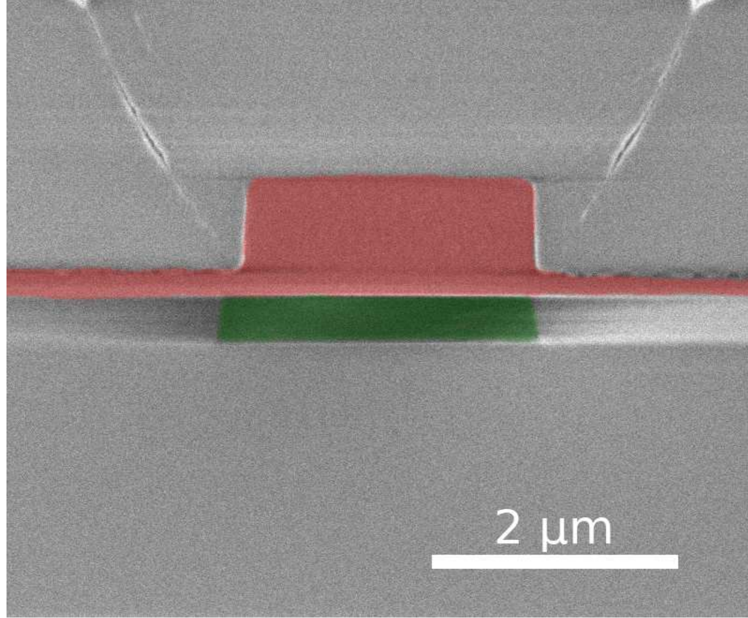


Figure 2.5: Cross section of the supercontinuum generation device. Red: chalcogenide glass waveguide. Green: silicon nitride layer for edge coupling.

The propagation loss of this device is 0.56 dB/cm, measured using both ring resonators (Figure 2.6) and optical frequency domain reflectometry (Figure 2.7). In the case of the ring resonator measurement, the propagation loss is related to the quality factor with the following [31]:

$$\alpha_{\text{lin}} = \frac{2\pi n_g}{Q_{\text{int}}\lambda_0} \tag{2.27}$$

The linear propagation loss is related to the dB propagation loss by

$$\alpha_{\text{dB}} = 10\log_{10}(\text{e})\alpha_{\text{lin}} \approx 4.34\alpha_{\text{lin}} \tag{2.28}$$

The ring resonance gives us the *loaded* quality factor

$$Q_{\text{loaded}} = \frac{\lambda_0}{\text{FWHM}} \tag{2.29}$$

from which we can determine the *intrinsic* quality factor with [31]

$$Q_{\text{int}} = \frac{2Q_{\text{loaded}}}{1 + \sqrt{\text{ER}^{-1}}} \tag{2.30}$$

where ER is the extinction ratio of the resonance.

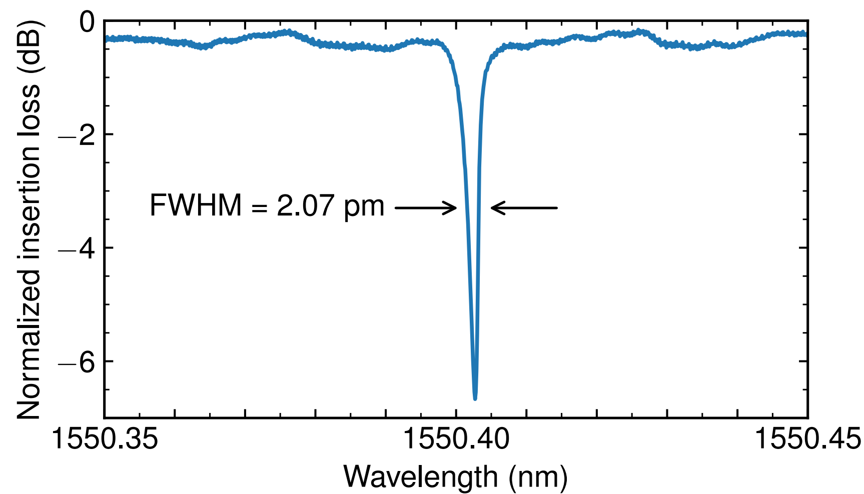


Figure 2.6: Chalcogenide glass ring resonator measurement for determining propagation loss

In the case of optical frequency domain measurement, the instrument measures the backscattered signal from the waveguide as a function of distance. Since the emitted backscattered signal is proportional to the attenuated signal, and undergoes the same loss on the way back, the slope of the OFDR signal is twice the propagation loss in the waveguide.

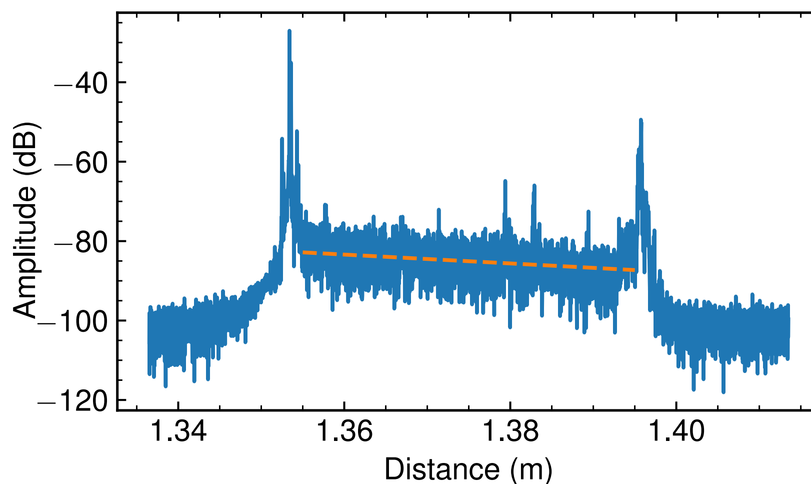


Figure 2.7: Optical frequency domain reflectometry measurement of a 4 cm long waveguide. The slope is related to the propagation loss.

An important parameter to engineer is the waveguide dispersion. We use the material refractive index from [20] which is reproduced in Figure 2.8. We confirm it is adequate to model dispersion by comparing with experimental data from ring resonators. The ring resonators cross section (e.g. Figure 2.5) is simulated in Lumerical MODE and the dispersion is extracted.

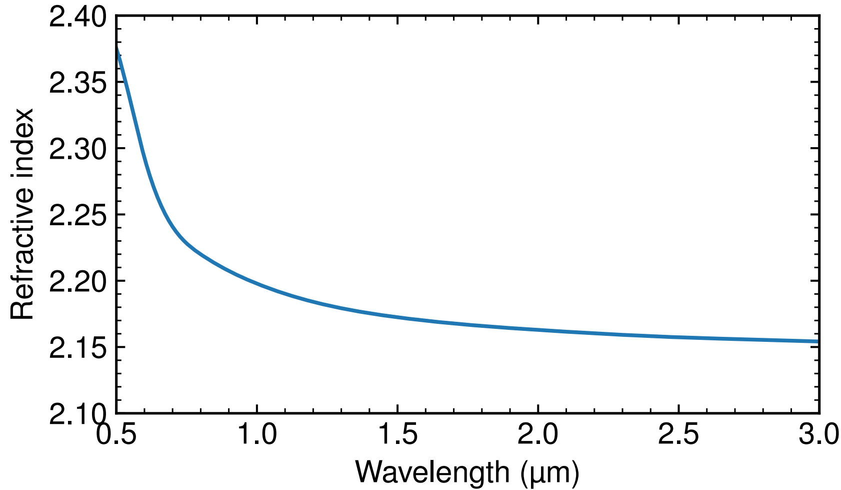


Figure 2.8: Ge_{23} Sb_7 S_{70} material dispersion.

Then, the experimental dispersion is determined by the derivative of the group index n_g, which is found from the ring resonator mode spacing:

$$n_g = \frac{\lambda^2}{\Delta\lambda L} \tag{2.31}$$

where $L = 2\pi R$ is the ring resonator circumference. We calculate one group index per free spectral range then use a low order polynomial fit to reduce the noise. Then, the dispersion is given by the derivative of the group index. The simulated and experimental dispersions are shown on Figure 2.9 for the fundamental quasi-TE and quasi-TM modes in the ring resonator, with good agreement for both modes.

This agreement of the dispersion gives confidence to use the dispersion model over the whole wavelength range. The dispersion from 1 µm to 3 µm for different waveguide widths is shown in Figure 2.10. The peak dispersion is located around 1.6 µm. We notice that changing the waveguide width mostly influences the red side dispersion.

One important feature of supercontinuum is the presence of *dispersive waves*. These are specific wavelength peaks in the supercontinuum which are generated from the resonance between the solitons and other wavelengths. Dispersive waves are useful in a context of self-referenced frequency combs where we would like to concentrate the supercontinuum energy at f and $2f$.

The dispersive wave location cannot be easily determined from the dispersion curve itself (the second derivative of the propagation constant), since it requires a phase matching condition. For

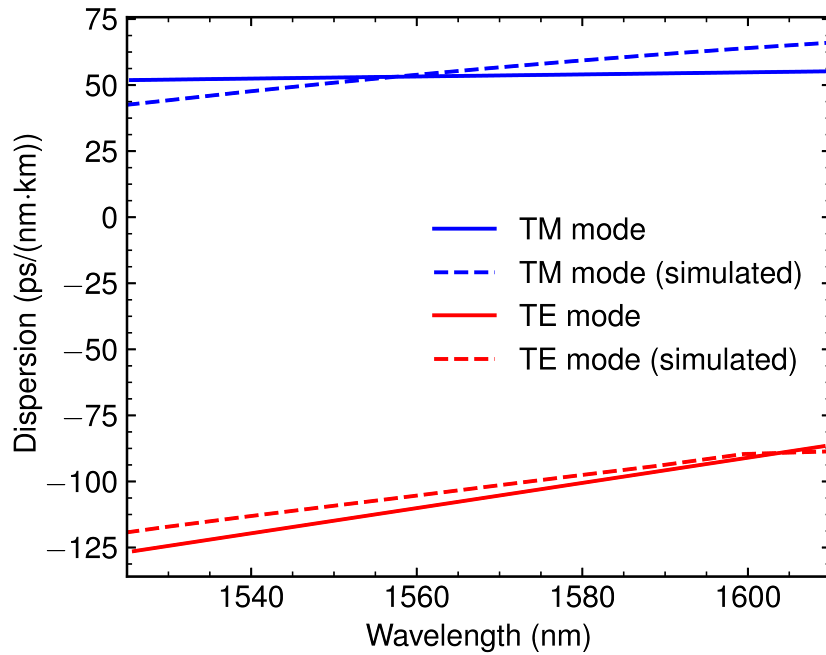


Figure 2.9: Measured and simulated dispersion of the chalcogenide waveguide for the fundamental modes of the ring resonator

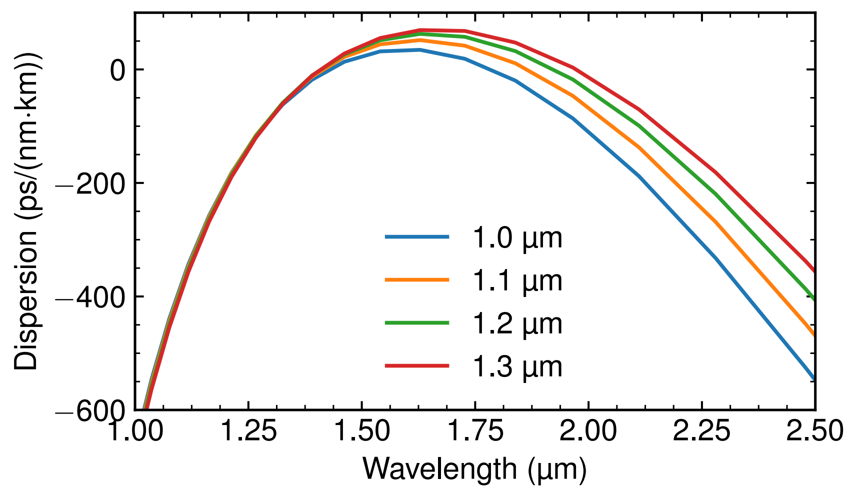


Figure 2.10: Simulated dispersion of the chalcogenide waveguide as a function of width

low soliton numbers, the dispersive wave is located at the following resonance condition [32]:

$$\beta(\omega_R) - \beta_1(\omega_R) = \frac{\gamma(\omega_S)P_S}{2} \tag{2.32}$$

The left hand side represents the propagation constant in the moving frame. In this case, (with a peak power of 77 W in the waveguide) the right hand side equals to 165 m^{-1}. In Figure 2.11, which shows the propagation constant for a 1.2 µm wide waveguide, we can find the dispersive wave location by finding the points where the propagation constant equals the right hand side. In this example, this happens at wavelengths of 1.1 µm and 2.9 µm. It should be noted that, for large soliton numbers, there will be a blue shift of the dispersive wave, and a more detailed analysis is required to get the precise dispersive wave location [32].

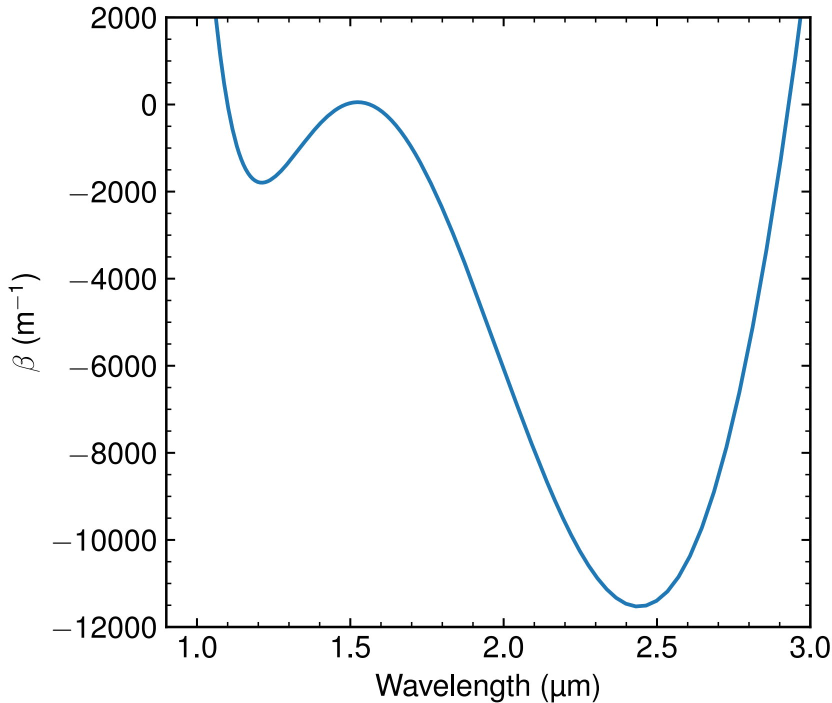


Figure 2.11: Propagation constant in the moving reference frame

In the supercontinuum simulation on Figure 2.12, the dispersive waves appear near the predicted wavelengths of 1.1 µm and 2.9 µm. The supercontinuum was measured experimentally in the fabricated chalcogenide waveguides. The pump laser is a Calmar FPL-01TP femtosecond fiber laser. The pulse width is 240 fs, measured with a Femtochrome FR-103HS autocorrelator. The repetition rate is 25 MHz. Accounting for the 7 dB insertion loss between the fiber and the waveguide, the peak power inside the waveguide is calculated to be 77 W, and the pulse energy is 26 pJ. For these parameters, we estimate the soliton number to be $N = 11$. The waveguide length is 2 cm, which is close to the calculated optimum length $z_{opt} = 2.05$ cm. The supercontinuum

spectrum is measured using an OceanOptics NIRQUEST512-2.5 and shown on Figure 2.12. The blue dispersive wave as well as the general shape near the pump wavelength are closely matched experimentally. Unfortunately, the spectrometer range could not extend far enough to detect the red dispersive wave.

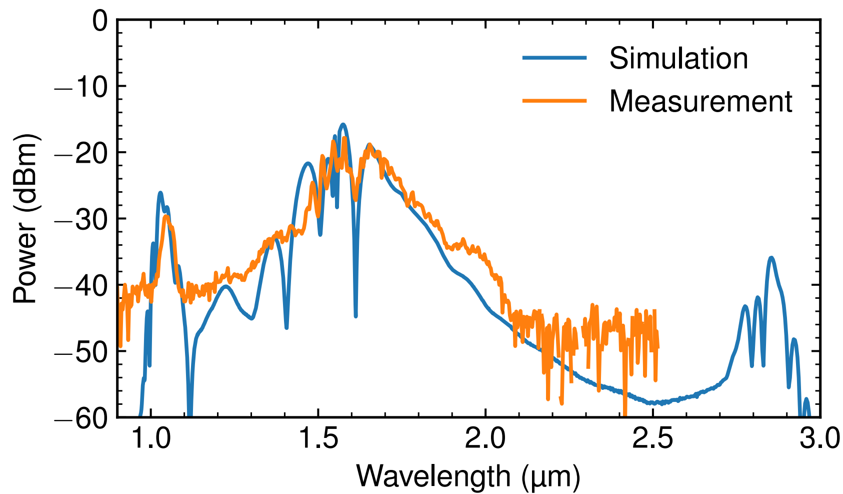


Figure 2.12: Simulated and measured supercontinuum generation

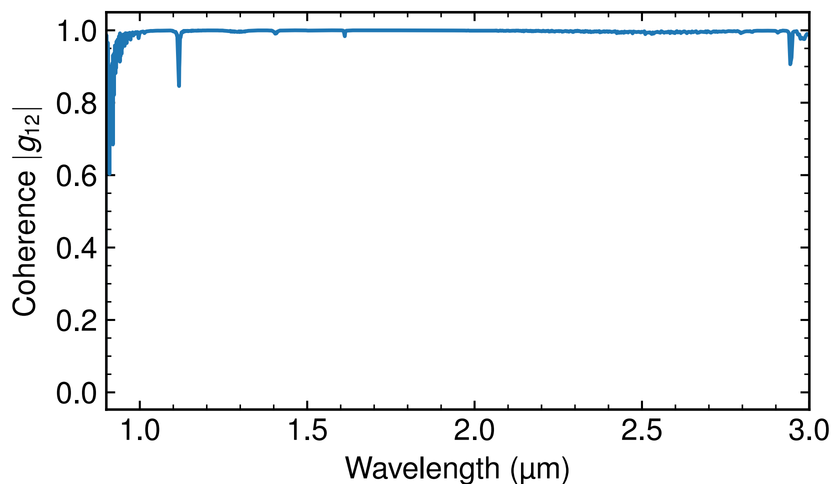


Figure 2.13: Simulated coherence of the generated supercontinuum, showing near-unity from 1.0 μm to 3.0 μm

Since the soliton number is estimated to be less than 16, we expect the supercontinuum coherence to be good (close to unity). To simulate the coherence, we simulate multiple propagations

of the same pulse with random noise and average the result according to equation (2.26). The results are plotted on Figure 2.13. Indeed, the simulation predicts near-unity coherent from 1.0 μm to beyond 3.0 μm.

The supercontinuum coherence is a more difficult parameter to measure. It can be confirmed indirectly, by measuring the $f-2f$ beat note in a nonlinear interferometer. However, a more direct measurement of the pulse to pulse coherence can be done using a regular interferometer and an optical spectrum analyzer [33]. This measurement is difficult to perform with low repetition rate sources (e.g. 25 MHz fiber lasers) because the interferometer length needs to equal the pulse-to-pulse spacing, which can be several meters long.

As mentioned previously, other schemes can be used to measure the f_0 offset frequency. In future work, the dispersion of the waveguide could be further optimized to use the alternate schemes and to improve the efficiency of the nonlinear interferometer. This could enable direct on-chip self-referenced frequency combs, without the use of external amplifiers [14]. Some examples of improvements include:

- In the current waveguide, there is another dispersive wave near 3 μm. The dispersion can be optimized to have the two dispersive waves at exactly 1.0 μm and 3.0 μm and use the $f-3f$ scheme.
- The long wavelength could be shifted towards 2.0 μm, to improve the efficiency of the $f-2f$ scheme.
- The short wavelength could be shifted toward 0.775 μm, to use the pump itself as the f frequency and improve the efficiency of the frequency doubling.

Chapter 3

Fabrication techniques

In this chapter, we describe some relevant techniques for fabricating integrated photonic waveguides. All the processes described here are intended for silicon substrates. They were developed in or adapted for the Marvell Nanofabrication laboratory at UC Berkeley. We first describe some lithography techniques relevant to integrated photonics, then we discuss about edge roughness and propagation loss, and finally we cover deposition, etching and other processes specific to chalcogenide glass, silicon nitride, and silicon waveguides. While most of the material in this chapter is a review of the literature, some emphasis will be made on specific process improvements that we developed.

3.1 Lithography

Lithography is one of the most critical fabrication steps. The choice of technology could be determined by the desired feature size, or photoresist selectivity constraints might call for thick photoresists which only work with certain technologies. In this section, we concentrate on wafer-scale lithography processes and we detail the use of 248 nm DUV stepper, I-line stepper and contact aligner for fabricating optical devices. We also mention other technologies available such as maskless aligners, and more advanced DUV steppers available at commercial foundries. Finally, we present considerations for simulating the lithography process.

DUV stepper

The Nanolab has an ASML 5500/300 DUV stepper (248 nm) which can expose 6 inch wafers with field sizes larger than 2 cm by 2 cm. The automated coating tracks can spin photoresists with thicknesses ranging from 0.36 µm to 4.7 µm.

The photoresist thickness for a particular application is chosen based on selectivity, etching depth and resolution requirements. Generally, 0.43 µm thick resist is used for etching GeSbS, 0.87 µm for patterning trenches for silicon nitride deposition, and 4.7 µm for masking the etching of deep trenches before dicing.

The use of a DUV stepper has several advantages for fabricating optical devices:

- Resolution of 250 nm and smaller, which is sufficient to fabricate many directional couplers, ring resonators and nanotapers
- Very fast for exposing large scale devices such as long spiraling waveguides
- Good field stitching and layer to layer alignment capability (typically 50 nm of alignment error)
- Good reproducibility and exposure uniformity
- Quality masks with nm scale address size can print circular feature with high fidelity

The last point is particularly important when fabricating low loss waveguide spirals and ring resonators.

For 220 nm SOI silicon photonics, a thinner resist can be used during etching. In this case we use 0.36 µm UV210 photoresist which has a 4:1 selectivity in the lam8 etcher. In many cases, a bottom anti-reflection coating (BARC) layer will be required when dealing with topography and fine resolution on silicon layers. Indeed, thin film interference effects become very important in this case since the photoresist thickness will vary over the topography. To insure uniform exposure over all the wafer, the BARC coating will prevent these thin film effects. The BARC layer is approximately 60 nm and replaces the hexamethydisilazane (HMDS) priming step. The BARC layer will not be removed by the photoresist developer, so it needs to be etched before the main layer etch. Fortunately, several dry etching chemistries will remove the BARC layer. For example, when etching silicon, the oxide breakthrough step is extended in duration to also etch away the BARC layer. The oxide breakthrough step is designed to etch away the native silicon oxide before the main etch, and uses a fluorine based chemistry (e.g. CF_4). The anti-reflection layer we use (AR3-600) is sensitive to humidity, so the wafer is baked at 190 °C for 60 s.

I-line stepper

The Nanolab also has a GCA-8500 i-line stepper. The resolution is 0.7 µm and the alignment accuracy is 0.3 µm. This stepper is compatible with the thick AZ P4620 photoresist used for backend processes. It is possible to include the GCA alignment marks on the first ASML lithography layer, such that both steppers can be used during the same process (mix and match process). This has the advantage of enabling optimal resolution for the critical layers, while opening up the photoresist options for the backend layers. There is a slight rotation ($-0.47°$) between the Nanolab ASML and GCA stepper stages, which can be compensated in the ASML job file. The GCA-last mix and match lithography process is as follow:

1. Expose and etch the ASML PM mark layer, including the wafer rotation (ASML reticle)
2. Expose and etch the GCA alignment mark layer (ASML reticle), including the wafer rotation, aligned to PM marks

3. Expose and process any further ASML layers, including wafer rotation, aligned to PM marks
4. Expose and process the GCA layers, aligned to GCA marks

Contact aligner

In some cases, coarse features are acceptable but thicker photoresist might be required. This is the case for example while etching deep trenches for releasing dies. In contact lithography, a photomask is first aligned to the substrate, then put into direct contact with the photoresist before being exposed. The mask can be as large as the wafer, exposing all the dies in one step. The same alignment fiducials from the DUV lithography step can be reused by including a negative image on the mask. The light source is typically a mercury-vapor lamp emitting both G-line and I-line wavelengths. The resolution can be as good as 1 µm, however the alignment is a manual process and highly dependent on the operator and specific alignment marks. An alignment accuracy of 2 µm can be achievable with a well-designed process.

For the purpose of fabricating optical facets for edge coupling, we deposit a 8 µm thick photoresist (AZ P4620) that can withstand etching tens of microns of oxide and hundreds of microns of silicon. The resist is exposed with approximately 300 mJ and hard baked at 90 °C for 30 minutes.

Compared to using an I-line stepper, the contact aligner is faster since it exposes the whole wafer at once. Both lithography methods have similar resolution. However, the contact between the photomask and the photoresist might leave residues on the photomask which needs to be cleaned often, and the alignment accuracy is more dependent on the operator's skills.

Maskless aligner

Maskless alignment tools use an array of UV lasers and light modulators to scan the wafer. This defines a custom pattern in a manner similar to e-beam lithography, but with conventional UV photoresists such as i-line sensitive resists. The resolution and alignment accuracy is similar to i-line steppers, but it is more convenient for chip-scale lithography since the exposure time is proportional to the area. Unfortunately, since the exposure is rasterized, this technique is not suitable for exposing low-loss photonic components which have bent or tapered structures.

Foundries

Contrary to electronic devices, the scaling of optical devices is limited by the wavelength of operation. For many optical components, especially operating in the infrared and in low-index contrast systems such as doped glass or silicon nitride, the resolution limit of 248 nm DUV steppers is sufficient. However, there are certain components, especially in high index contrast platforms, which require a feature size smaller than 250 nm.

Some features which might benefit from the better resolution provided by foundries include:

1. Various gratings: grating couplers, Bragg gratings, and especially subwavelength gratings

2. Inverse nanotapers for coupling to fibers
3. Components using inverse design [34]

Lithography simulation and optimization

Reviewing the principles of lithography is useful to understand the effects of different fabrication parameters on the resolution. The simulation is divided in 3 steps. Although more advanced models exist today, we will give an overview of the simple models from [35] since they can easily be implemented and require a limited number of parameters.

The first step is the aerial image formation. In terms of Fourier optics, this step computes the image of the mask considering the source illumination pattern, and the objective lens pupil. Assuming a normalized incoherent source $S(\chi)$, a pupil function $P(f_x)$, and a mask transparency function $T(f_x)$, the intensity a the wafer plane is written as [35]

$$I(x) = \int S(\chi) |\mathcal{F} [T(f_x - \chi) P(f_x)]|^2 \, \mathrm{d}\chi \tag{3.1}$$

In this representation, the mask is shifted instead of the source to account for source points that hit the mask at an angle.

The aerial image can be readily computed using libraries for computing the image in microscopy systems, such as Microlith [36].

Once the aerial image is calculated, the image in the photoresist is calculated. This can be done in several levels of approximation, depending on the wavelength, photoresist thickness, numerical aperture and feature size:

- The light intensity can be assumed to be constant as a function of depth in the photoresist
- The thin film interference effects and the defocus effects can be added by multiplying the aerial image by the depth dependent interference pattern and adding a defocus phase term in the pupil function
- A scalar or vector wave equation model can be used to propagate the aerial image in the photoresist, for example using the FDTD method. This will take into account any kind of topography. Both light polarizations can be taken into account.

A simple model linking the light intensity to the photoresist image is called the Dill model, which has three parameters, A, B, and C [35].

The light absorption in the photoresist is determined by

$$\frac{\partial I}{\partial z} = -(Am + B)I \tag{3.2}$$

where B is a constant absorption coefficient, and A linked the absorption to the activated concentration m. The activated concentration depends on the light intensity as

$$\frac{\partial m}{\partial t} = -CIM \tag{3.3}$$

where C links the activated concentration to the light intensity. The photoresist image might need to be determined iteratively since the absorption will depend on the absorbed energy. The Dill's model parameters are often cited in the photoresist datasheet.

Positive DUV resists are commonly chemically amplified. The amplification takes place during the post exposure bake. Additional diffusion of the activated species occurs during the bake, which helps in removing the standing wave patterns in the photoresist, but might also affect the linewidth. Because of this, the post exposure bake time and temperature must be well controlled. The amplification and diffusion processes are modelled using additional parameters, but since the post exposure bake recipe we used is fixed (130 °C for 90 s) we will not study this in more detail.

The development rate r can be calculated from different models [35]. Once the development rate is known, the development is modelled using the eikonal equation (3.4).

$$|\nabla u| = \frac{1}{r} \tag{3.4}$$

This problem is similar to geometric optics propagation in a material with a refractive index gradient. The solution represents the shortest time to travel to a certain point. The final resist profile is given by the contour line for the given development time. Several algorithms and libraries exist to solve this equation, such as [37].

During lithography, several parameters can be tweaked to improve the fabrication tolerance (process window). In particular, using a higher numerical aperture will increase the resolution, at the expense of decreased depth of focus. A higher numerical aperture will also increase the proximity effects, which would need to be modelled using more advanced methods to quantify properly.

When smaller features need to be resolved, the use of annular illumination (instead of conventional partially coherent illumination) might also be desirable (Figure 3.1).

(a) Straight illumination

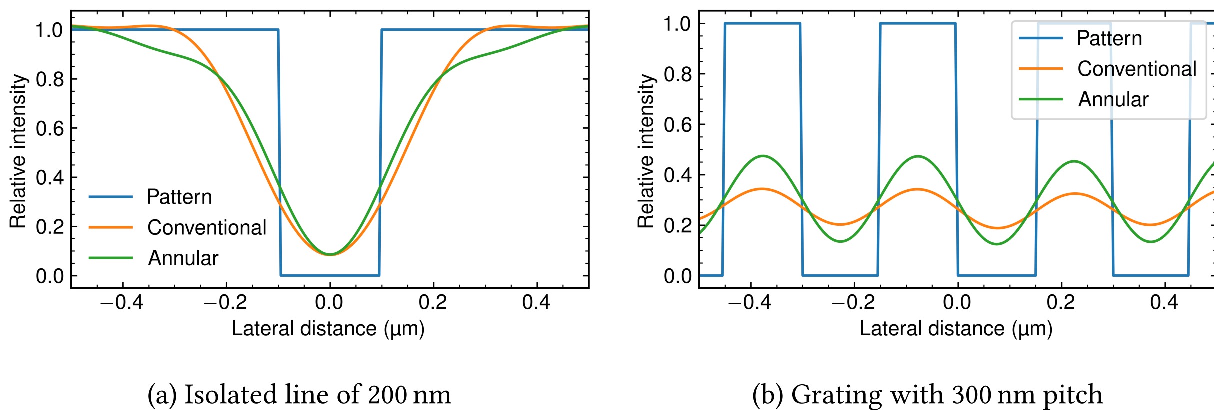
(b) Angled illumination

Figure 3.1: Lithography illumination

Indeed, when the numerical aperture is not sufficient to resolve the first diffraction order, angled illumination will tilt the diffraction pattern to let pass the first diffraction order (at the expense of the first negative diffraction order). This is similar to dark field illumination in microscopy, which is used to highlight the edges of the sample. The annular illumination parameters

are expressed as a fraction of the numerical aperture, both for the inner and outer limit of the illumination ring.

An example of the difference in aerial image between conventional and annular illumination for a 200 nm line and for a 300 nm pitch grating is given on Figure 3.2. The improvement in linewidth is slight for the isolated line, but the contrast is greatly improved in the case of the grating with small pitch.



(a) Isolated line of 200 nm

(b) Grating with 300 nm pitch

Figure 3.2: Comparison of aerial image between conventional and annular illumination

We have used annular illumination to improve the process window of silicon nanotapers down to 200 nm wide on an ASML DUV stapper with a wavelength of 248 nm, with the following parameters:

- $NA = 0.63$
- $\sigma_{\text{out}} = 0.81$
- $\sigma_{\text{in}} = 0.51$

It should be noted, however, that annular illumination will enhance some features, such as the examples given above, at the expense of other features. For example, gratings of specific pitches might not resolve properly, and the tolerance on some dimensions larger than the critical dimension might degrade. Finally, an increase in numerical aperture will also decrease the depth of focus and increase the proximity effect.

Another method to improve the process window of small features is to deposit a bottom anti-reflective coating (BARC) layer underneath the photoresist. This layer readily absorbs UV light while having a refractive index engineered to minimize Fresnel reflections. This will prevent standing wave patterns in the photoresist and make the dose to clear the photoresist less dependent on the exact thickness. The BARC layer has two advantages:

- The photoresist thickness control can be looser while achieving the same critical dimension tolerance

- The topography will have reduced influence on the dimensions, i.e. the features at high and low topographies will resolve at similar doses.

The standing wave pattern is described using [35]

$$E(z) = E_0 \frac{\tau_{12} \left(\exp(-2i\pi n_2 z/\lambda) + \rho_{23}\tau_{\mathrm{D}}^2 \exp(2i\pi n_2 z/\lambda)\right)}{1 + \rho_{12}\rho_{23}\tau_{\mathrm{D}}^2} \tag{3.5}$$

where

$$\begin{aligned} \rho_{ij} &= \frac{n_i - n_j}{n_i + n_j} \\ \tau_{ij} &= \frac{2n_i}{n_i + n_j} \\ \tau_{\mathrm{D}} &= \exp\{-ik_2 D\} \\ k_j &= \frac{2\pi n_j}{\lambda} \end{aligned}$$

and n_j is the complex index of refraction in layer j. When dealing with more than one intermediate layer, libraries using the transfer matrix method can be used [38]

Figure 3.3 gives an example of the standing wave pattern on a Si substrate, on a thick SiO_2 layer and on a 60 nm BARC layer. On a silicon substrate, the standing wave pattern is significant, but can be greatly reduced with the BARC layer. On a SiO_2 layer (e.g. with a thick bottom cladding), the standing wave pattern is not significant enough to justify the use of BARC.

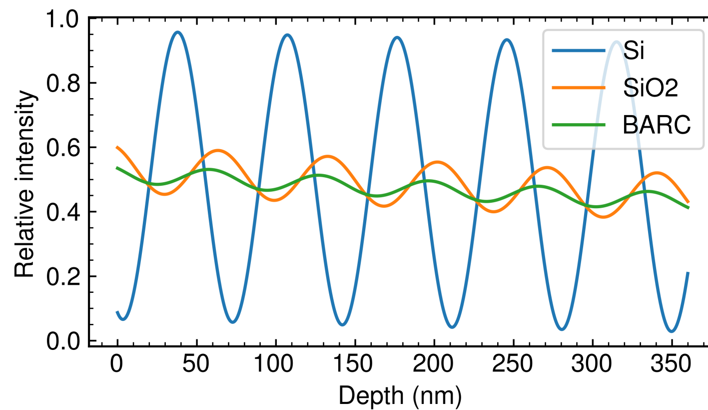


Figure 3.3: Standing wave patterns in photoresist.

3.2 Propagation loss

For most photonic devices, the propagation loss is an important performance parameter. The waveguide propagation losses come mainly from three factors

- Material absorption
- Scattering due to surface roughness
- Leakage, especially to a higher index contrast substrate.

The material absorption will be discussed in the next sections for the specific materials, and leakage will be discuss in the next chapters.

Payne and Laycey [39] developed a theoretical model for propagation loss due to scattering by rough surfaces

$$\alpha \leq \frac{\sigma^2}{k_0 d^4 n_1}\kappa \tag{3.6}$$

where σ is the roughness amplitude, k_0 is the light wavenumber, d is the waveguide half-width, n_1 is the waveguide refractive index, and κ is a constant that depends on the roughness statistics. This represents an upper bound and was developed for a slab waveguide. It is nonetheless useful in the context of 3D waveguides of arbitrary shape: it predicts that the scattering loss depends on the square of the roughness amplitude, and on the 4th power of the waveguide width. Consequently, one of the most effective ways to reduce the propagation loss is by design, with wider waveguides. This might be possible up to a certain point, where the waveguide becomes multimode and other properties (such as minimum bending radius) start to degrade. The second most important element to consider when reducing the propagation loss is the amplitude of the roughness. In many cases extra processing steps can be performed to improve the surface of the waveguide.

The Payne and Lacey model is less accurate for two dimensional waveguides and for large index contrasts [40]. In this case, numerical methods can be used to estimate the propagation loss for different waveguide widths. One such method is to model the sidewall roughness as a lossy material with the imaginary part of the refractive index calibrated to a measurement [41].

For wide waveguides, the propagation loss should decrease as determined by the Payne and Lacey models. For very narrow waveguides, since the waveguide confinement decreases the loss is expected to decrease and to be limited instead by the cladding material loss. We simulate a 220 nm silicon strip waveguide with SiO_2 cladding for different widths on Figure 3.4. The sidewalls are modelled using a 10 nm thick region with the imaginary part of the refractive index calibrated for 5 dB/cm loss at 400 nm for the TE mode. This is a typical value that was measured on devices fabricated at UC Berkeley. The propagation loss peaks around 200 nm waveguide width, where there is maximum interaction with the sidewalls. Moreover, the simulation predicts that the propagation loss is lower in general for the TM mode because the fields interact more with the top and bottom surfaces, and less with the sidewalls compared to the TE polarization.

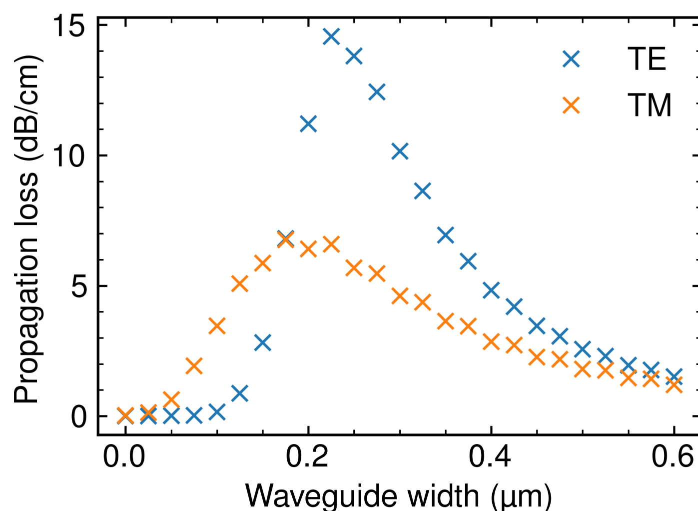


Figure 3.4: Propagation loss from sidewall roughness as a function of waveguide width

Not to be neglected is the roughness of the top and bottom surface. Some deposition methods (e.g. LPCVD SiO_2) produce a surface with measurable roughness which can be improved using chemical-mechanical polishing (CMP).

Some methods to reduce the sidewall roughness of silicon waveguides include cycles of oxidation and oxide removal [42] and photoresist reflow [43, 44].

3.3 Chalcogenide glasses

Chalcogenide glasses form a family of compounds which contain oxide-like elements (sulfur, selenide, tellurium). Many of them are optically transparent, and often exhibit nonlinear optical properties. They are used as the core of integrated waveguides in applications where transparency in the mid-infrared is required, or when high nonlinearity is desirable.

Some common examples of chalcogenide glasses used for integrated optics include As_2S_3 [45], As_2Se_3 [46] and GeAsSe [47]. We note that many of these compounds contain arsenic, a material which might not be desirable to manipulate in some laboratory settings. Many chalcogenide glass compounds also suffer from photodarkening or photobleaching effects [48].

Chalcogenide glasses such as $Ge_{23}Sb_7S_{70}$ are readily available in bulk form. They are easy to evaporate (with glass transition temperatures around 250 °C) to create thin films on wafer substrates. Both thermal evaporation and electron beam evaporation have been reported in the literature to evaporate chalcogenide glasses. In this work we only explored thermal evaporation since the extra heating power from an electron beam evaporator was not required.

Several etching techniques can be used with $Ge_{23}Sb_7S_{70}$. Isotropic wet etching (especially in alkaline solutions) has been observed. The film also etches very fast in a XeF2 silicon etching

chamber. It behaves similarly to SiO_2 under plasma etching. Etching using fluorine-based as well as chlorine-based chemistries has been demonstrated [49].

Deposition

For depositing the chalcogenide glass, we chose thermal evaporation for its simplicity. We used the NRC thermal evaporator in the Marvell Nanofabrication laboratory which did not have material restriction for $Ge_{23}Sb_7S_{70}$. The oldest tool in the cleanroom was modified with increased source to wafer distance for greater uniformity (4%), and electrodes with a better clamping mechanism were installed to improve the reproducibility. The film thickness is monitored using a crystal monitor. The tooling factor was calibrated on a known substrate measured using SEM. The evaporator base pressure is typically around 2×10^{-6} Torr or better.

Since a too high evaporation rate might affect the quality and stoichiometry of the film, we chose a deposition rate of approximately 1.5 nm/s. The film composition was verified using energy-dispersive X-ray spectroscopy (EDX) and matches the bulk composition within 1% mass composition. After depositing the chalcogenide film, a thin PECVD SiO_2 layer (usually 10 nm) is deposited to mask the layer from the alkaline photoresist developer which can etch the chalcogenide glass. The PECVD layer is deposited at a lower temperature of 250 °C to prevent melting of crystallization of the film.

Etching

The Nanolab etcher ptherm is a reactive ion etcher (RIE) chamber with no material restriction and a large variety of etching gasses. A combination of 80 sccm of CHF_3 and 4 sccm of O_2 at a power of 150 W will yield smooth and vertical sidewalls with an etching rate of 50 nm/min. This is similar to a standard recipe for etching SiO_2. The etched floor is smooth provided that the chamber is conditioned with a dummy wafer coated with a blank GeSbS film for at least one minute prior to etching the sample. Unfortunately the etching non-uniformity is around 10% in this chamber, which lowers the yield on the fabricated devices.

After etching, the photoresist is removed using a combination of O_2 plasma ashing at low power (100 W) for a few minutes and a long soak in 1165 photoresist remover (20 minutes). The use of a more powerful plasma asher such as the matrix tool was found to damage the film. Then, the PECVD SiO_2 cladding is deposited again at 250 °C to prevent damage to the film.

A cross section of the completed waveguide is shown in Figure 3.5. The rib waveguide has nearly vertical sidewalls. Some imperfections from the fabrication process can be observed: the etching loading created a nonuniform slab thickness near the waveguide, and the PECVD deposition being not perfectly conformal creates "rabbit ear" shaped voids in the top cladding.

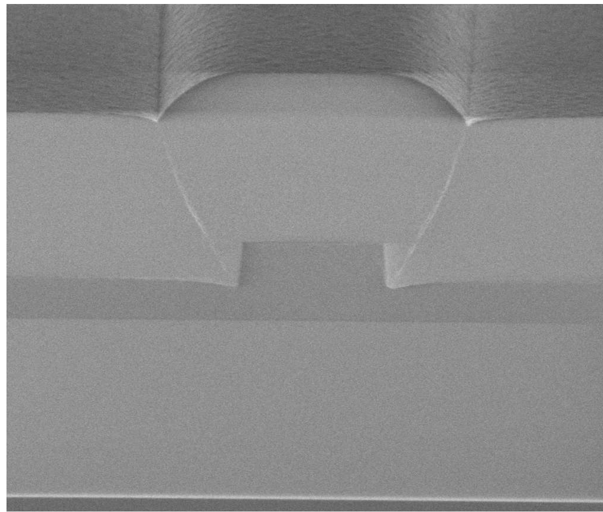


Figure 3.5: Chalcogenide waveguide cross section after top cladding deposition

3.4 Silicon nitride

Silicon nitride is a material commonly used in CMOS fabrication. The fabrication techniques and limitations are well-known. In integrated optics, silicon nitride is often used as a core material for low propagation loss waveguides. One of the unique challenge in integrated optics processing compared to CMOS is the difficulty to fabricate thick (>300 nm) layers because of the high tensile stress of the deposited films.

Silicon nitride can be deposited using either plasma-enhanced chemical vapor deposition, or low-pressure chemical vapor deposition. The former has a high deposition rate, and is deposited at lower temperature. It is often used in backend processes where the thermal budget is constrained. The latter deposition method requires elevated temperature (>800 °C) but gives a higher quality film. LPCVD silicon nitride also has the desirable property of being conformal: it can fill trenches without voids.

Depending on the gas ratio during deposition, the resulting silicon nitride can be either stoichiometric Si_3N_4 or low stress. The low-stress variant is silicon-rich and has a slightly higher refractive index.

Photonic damascene process

A clever way to deal with the silicon nitride film stress is to deposit the waveguide inside a preformed trench [50], a method which is called the photonic damascene process. The preform is made of silicon dioxide and is etched using a mask which is the negative of the waveguide. In our process, we use an amorphous silicon hardmask for etching the silicon dioxide. After depositing the silicon nitride film, the excess is removed using chemical-mechanical polishing (CMP). The process is summarized in Figure 3.6. Silicon nitride is a good candidate material for the photonic damascene process since its deposition is conformal. The geometry of the preform

helps in releasing the stress by converting the tensile stress into bending stress (Figure 3.7), and any remaining stress is removed during the CMP step.

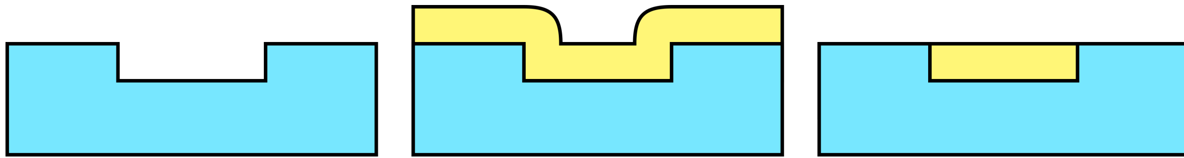


Figure 3.6: Photonic damascene process: preform etch, silicon nitride deposition and silicon nitride polishing

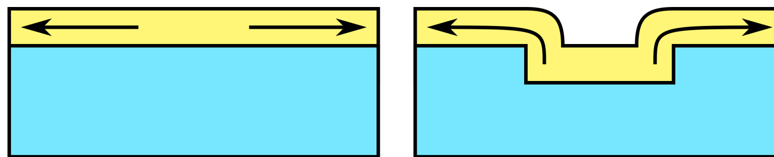


Figure 3.7: Photonic damascene stress release

We used the photonic damascene process in a hybrid silicon-nitride/chalcogenide glass process, where the chalcogenide glass is deposited after the silicon nitride waveguide is formed.

Silicon nitride deposition and annealing

While the photonic damascene process can easily deal with the issue of silicon nitride film stress, it is also possible with careful technique to deposit thick plain films. For example, splitting the deposition in several steps or rotating the wafer during the deposition has been shown to yield crack-free silicon nitride films [51, 52, 53]. In this case, the silicon nitride film can be dry etched post-deposition using fluorine-based chemistries such as CF_4/O_2 mixes.

The LPCVD deposited silicon nitride films still contain hydrogen atoms from the ammonia (NH_3), the deposition gas used to carry the nitrogen atoms. This hydrogen forms $Si-H$ and $N-H$ bonds that greatly contribute to the silicon nitride optical loss. A well-known technique to drive out the hydrogen is to anneal the film at high temperature for several hours [54]. For example, we use 4 h at 1200 °C, or longer times at 1150 °C [55].

Chemical-mechanical polishing of silicon nitride

In the chemical-mechanical polishing process, several parameters can be adjusted to improve uniformity or polishing rate:

- Table rotation speed
- Arm rotation speed

- Arm pressure
- Wafer back pressure
- Slurry type and flow rate
- Pad type

A good starting point for wafer uniformity is to use a table rotation speed similar to the arm rotation speed [56]. When both are equal, the relative velocity of the points on the wafer to the table is constant across the wafer. A slight difference in rotation speed can be used to adjust a nonuniform rate between the center and the edge of the wafers.

Additionally, the wafer back pressure can be used to compensate for nonuniformity. On some machines the back pressure is applied in the middle of the wafer to compensate for edge effects on the pressure distribution.

3.5 Oxide cladding

Silicon dioxide (SiO_2) is often used as a top and bottom cladding for dielectric waveguides, because of its good optical quality and reproducible fabrication processes. It can be deposited or grown in several ways, from fastest rate to slowest rate:

- Plasma-enhanced chemical vapor deposition (PECVD)
- Low-pressure chemical vapor deposition (LTO LPCVD)
- Thermal oxidation

PECVD oxide will have varying quality depending on the deposition temperature, but can be grown at a rate around 67 nm/min. LPCVD has a slower deposition rate around 13 nm/min, but in this case several wafers can be processed in batch inside a furnace.

In all cases, a post-deposition annealing might be desirable to remove hydrogen impurities and densify the oxide. More information on the annealing process can be found in [55].

The cladding thickness varies with the waveguide confinement. Chalcogenide and silicon nitride waveguides would typically require 2 μm to 4 μm of top and bottom cladding, while low-index waveguides with modes matched to single mode fibers need at least 20 μm total cladding. For very thick claddings, film stress induces wafer bowing which becomes an issue for some fabrication steps, in particular lithography. The relation between film stress and wafer bow is given by the Stoney equation (3.7).

$$r = \frac{E_s t_s^2}{(1 - \nu_s) 6 \sigma_f t_f} \tag{3.7}$$

One method to mitigate wafer bowing is to deposit a similar film on the back side of the wafer. LPCVD does this inherently, but PECVD requires a second deposition step with the wafer upside

down. Even with balanced stress between the top and bottom films, the wafer still undergoes expansion or contraction which might not be insignificant. The wafer expansion is given by the biaxial strain formula (3.8).

$$\epsilon = \frac{\sigma}{E}(1 - \nu) \tag{3.8}$$

With a 10 μm, 100 MPa PECVD SiO_2 film on both sides of a 6 inch wafer, the wafer expansion is approximately 1.9 μm which can cause alignment errors in automated lithography tools such as the ASML 5500/300 DUV stepper. In this case, the alignment program parameters need to include a compensation for the change in wafer size.

3.6 Silicon

Silicon is one of the most widely used materials in integrated photonics, hence the name "silicon photonics". It has several properties that are useful for creating complex optical systems:

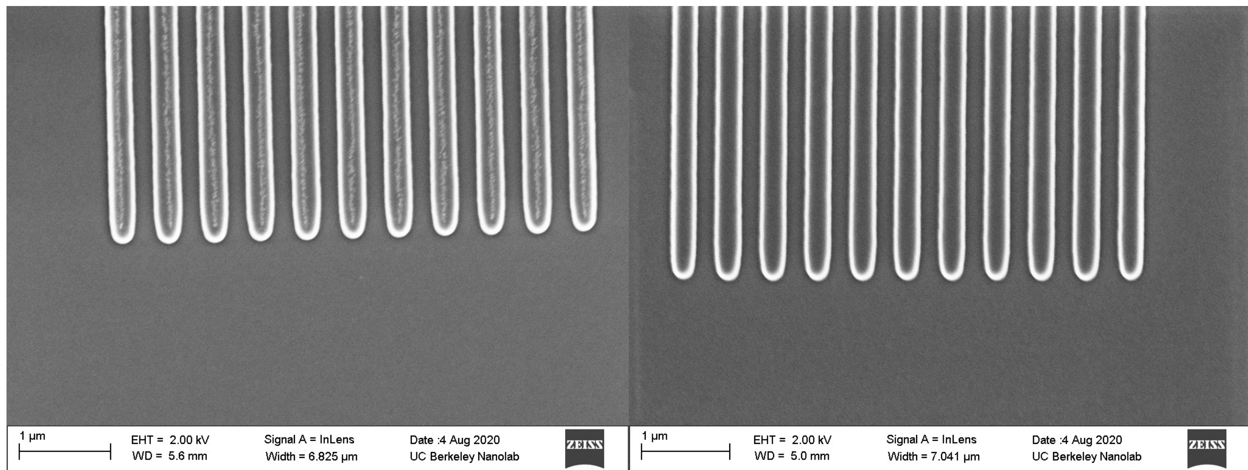
- Thermal stability
- High refractive index
- Can modulate refractive index and absorption electrically
- Low absorption in telecom wavelengths
- Mature fabrication processes

The last point might be the most important in determining the success of silicon photonics. The advanced fabrication in state of the art foundries can be easily leveraged to create devices with high yield and tight fabrication tolerances, which can also be directly integrated with electronics [57, 58].

A simple process for silicon photonics is to start with a silicon-on-insulator (SOI) wafer. A SOI wafer consists of a thin silicon device layer on top of a buried oxide (BOX) layer, on a thicker silicon substrate. These wafers are available commercially and the silicon device layer is of high quality since it is fabricated by bonding and cutting from another crystalline silicon wafer.

The silicon layer is then etched using one of several methods. Wet etchants include KOH which etches the $\langle 100 \rangle$ planes with great selectivity over $\langle 111 \rangle$ planes. Plasma etching methods are preferred for photonics to produce structures with vertical sidewalls. In the Marvell Nanofabrication Laboratory, the silicon etcher uses a mixture of HBr and Cl_2 for good selectivity to SiO_2 and etching rate.

The HBr etching chemistry produces polymers which cover the sidewalls for greater etching anisotropy. These polymers are a mixture of Si, C and Br [59]. However, these polymers are hard to remove, especially the ones which deposit on the photoresist sidewalls. Several wet cleaning methods have been found to work in this case:



(a) Before HF clean (b) After HF clean

Figure 3.8: Cleaning of silicon grating coupler etch residues in dilute HF

- Dilute HF dip or vapor HF [60]
- RCA clean, especially SC-1 [61]
- Hydroxylamine-based solvents, such as EKC 265 [62]

For simplicity, and because we do not have thin gate oxides which require high selectivity, we clean the HBr residues using 20 s of 100 : 1 dilute HF mixture, followed by rinsing in a quick dump rinse tank (QDR) and a spin rinse dryer (SRD). During cleaning of the full silicon etch step where the bottom oxide layer is exposed, only a few nanometers of the oxide will be etched by the cleaning process. The effect of cleaning is shown in Figure 3.8.

Chapter 4

Adiabatic component design

4.1 Adiabatic theorem and mode evolution

Adiabatic evolution is a well-known concept in quantum mechanics. In short, if a system is perturbed slowly enough and is initially in an eigenstate, it will remain in this eigenstate. This can be used for example to change the spin of an electron. It can also be applied to waveguide systems where more than one mode can be propagating. We will review the basic equations of adiabatic mode propagation, and in particular the condition for adiabaticity in optical systems.

We can write Maxwell's equations in a waveguide using operator notation similar to quantum mechanics [63]:

$$\hat{A}\,|\Psi\rangle = -i\frac{\partial}{\partial z}\hat{B}\,|\Psi\rangle \tag{4.1}$$

The operator $\hat{A}$ contains the double curl operator and the medium permittivity, and $\hat{B}$ is the cross product. The vector Ψ represents the field propagating in the waveguide system. Thus, $\langle\Psi_1|\,\hat{B}\,|\Psi_2\rangle$ represents the *unconjugated* Poynting vector $\int(E_1 \times H_2 + E_2 \times H_1)\cdot dS$. Similar results could be derived for the conjugated Poynting vector, but there would be a loss of generality for lossy waveguides. The waveguide modes satisfy the eigenvalue equation

$$\hat{A}\,|\psi_n\rangle = \beta_n\hat{B}\,|\psi_n\rangle \tag{4.2}$$

and we use the unconjugated normalization $\int(E \times H)\cdot dS$. Optical wavegeuide mode solvers solve this equation for the propagation constant β_n and the mode ψ_n. A field propagating inside the waveguide can be expanded in a combination of eigenmodes:

$$|\Psi\rangle = \sum_n c_n(z)\,|\psi_n\rangle \tag{4.3}$$

By replacing (4.3) and (4.2) in (4.1) we can derive an expression for the mode evolution containing the coupling between the instantaneous eigenmodes and their derivatives.

$$\dot{c}_m(z) = i\beta_m c_m(z) - \sum_n c_n(z) \langle\psi_m| \hat{B} \left|\dot{\psi}_n\right\rangle \tag{4.4}$$

Under the unconjugated normalization,

$$\langle\psi_n| B \left|\dot{\psi}_n\right\rangle = 0 \tag{4.5}$$

This can be proven by using $\frac{\partial}{\partial z} \langle\psi_n| B |\psi_n\rangle = 0$. Equation (4.4) could be used when the modes are found numerically and the derivative would be approximated by finite difference. It is however not entirely convenient to find the derivative of eigenmodes analytically. We can differentiate (4.2) to replace $\langle\psi_m| \hat{B} \left|\dot{\psi}_n\right\rangle$

$$\langle\psi_m| \dot{\hat{A}} |\psi_n\rangle + (\beta_m - \beta_n) \langle\psi_m| \hat{B} \left|\dot{\psi}_n\right\rangle = \frac{\partial\beta_n}{\partial z}\delta_{m,n} \tag{4.6}$$

which gives the result similar to perturbation theory:

$$\dot{c}_m(z) = i\beta_m c_m(z) - \sum_{n\neq m} c_n(z) \frac{\langle\psi_m| \dot{\hat{A}} |\psi_n\rangle}{\beta_n - \beta_m} \tag{4.7}$$

We also get the following identity:

$$\langle\psi_n| \dot{\hat{A}} |\psi_n\rangle = \frac{\partial\beta_n}{\partial z} \tag{4.8}$$

The derivative of the $\hat{A}$ operator contains derivatives of ϵ. We can express it as [63]

$$\langle\psi_m| \dot{\hat{A}} |\psi_n\rangle = \int \frac{\partial\epsilon}{\partial z} E_n E_m dS \tag{4.9}$$

Since ϵ represents a moving boundary, this integral becomes a line integral on the moving boundary moving at speed $\frac{\partial h}{\partial z}$:

$$\int \frac{\partial h}{\partial z} \Delta\epsilon E_n E_m dl \tag{4.10}$$

This confirms the intuition that boundaries moving faster, with higher index contrast or where the field is strong will reduce the adiabaticity of the process.

To get an adiabatic process in a waveguide with varying cross-section, the waveguide modes need to change either infinitesimally, or slowly, in order to keep the off-diagonal terms in (4.7) small. This can be accomplished with a linear taper of sufficient length.

A method to reduce the length of an adiabatic taper is to use a nonlinear shape that will have optimal adiabaticity throughout the taper. Indeed, some sections of the device might require a slower taper to prevent coupling to other modes. The optimization of nonlinear tapers is discussed in Section 4.7.

4.2 Simulation of adiabatic devices

Tapered waveguides, like other photonic components, can be simulated using different techniques. This includes 3D vectorial techniques such as finite-difference-time-domain methods (FDTD), or paraxial approximations such as beam propagation methods. A convenient method for simulating long propagating components is the eigenmode expansion method (EME). The modes of the device are computed in several slices along the device. The propagation and mixing of the modes in between the slices can be calculated using either a staircase approximation [64], or first order adiabatic approximation (such as equations (4.4) or (4.7)).

4.3 Simple tapers

Waveguide tapers can be used for several reasons: matching the waveguide width between two different components, expanding the mode for coupling to a fiber, or changing the effective refractive or dispersion to have better control on the phase. In all these cases, the taper length needs to be long enough to prevent exciting higher order modes or radiation modes.

We are interested in estimating the coupling between the fundamental mode and the first higher order mode. Because of symmetry, the TE_0 mode will couple to the TE_2 mode for example.

The simplest example is a linear taper, of width W and taper angle $\theta = \frac{1}{2}\frac{\partial W}{\partial z}$. We assume the taper supports only two modes. Then the propagation equations (4.7) become

$$\frac{\mathrm{d}c_0}{\mathrm{d}z} = -Cc_1 + i\beta_0 c_0 \tag{4.11}$$

$$\frac{\mathrm{d}c_1}{\mathrm{d}z} = Cc_0 + i\beta_1 c_1 \tag{4.12}$$

The second term describes the phase accumulation during propagation. The first term is more complicated to calculate analytically in general. In [65] some results are derived for a low index contrast waveguide (the notation is slightly adapted):

$$\Delta\beta = \beta_1 - \beta_0 = \frac{2\pi\lambda_0}{W^2 n_{\text{eff}}} \tag{4.13}$$

$$C = \frac{\langle\psi_0|\,\dot{\hat{A}}\,|\psi_1\rangle}{\beta_1 - \beta_0} = \frac{3}{4W}\frac{\partial W}{\partial z} \tag{4.14}$$

To simplify the derivation, we assume that the taper is linear, and that the width does not vary much, hence β_i is a constant. To find the solution, we replace

$$a_0 = c_0 e^{i\beta_0 z} \tag{4.15}$$

$$a_1 = c_1 e^{i\beta_1 z} \tag{4.16}$$

and we get

$$\frac{\mathrm{d}a_0}{\mathrm{d}z} = -Ca_1 e^{i\Delta\beta z} \tag{4.17}$$

$$\frac{\mathrm{d}a_1}{\mathrm{d}z} = Ca_0 e^{-i\Delta\beta z} \tag{4.18}$$

The power will oscillate between the two modes inside the taper. We assume now that all the power is in a_0. For lengths shorter than the beat length, the solution is simply

$$|a_1|^2 \approx (Cz)^2 \tag{4.19}$$

For longer propagation lengths (which is usually the case), the solution oscillates with a maximum power of [65]

$$|a_1|^2 \approx 4\left(\frac{C}{\Delta\beta}\right)^2 = \left(\frac{3\theta W n_{\mathrm{eff}}}{2\pi\lambda_0}\right)^2 \tag{4.20}$$

Which gives the simple condition

$$\theta \ll \frac{\lambda_0}{n_{\mathrm{eff}} W} \tag{4.21}$$

This relation can be used as an adiabaticity criterion for designing simple tapers which will have a constant adiabaticity parameter [66].

For example, we simulate a silicon taper from 1 µm width to 2 µm width, at $\lambda = 1.55$ µm. We apply a conservative adiabaticity criteria using 0.1θ and obtain a taper length of 13.5 µm, which is clearly in the adiabatic regime according to Figure 4.1.

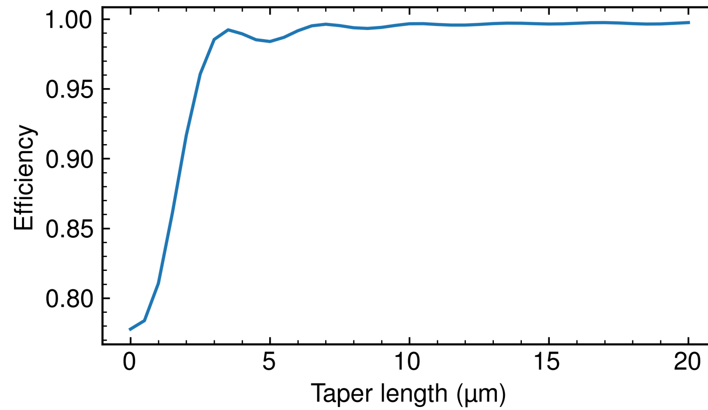


Figure 4.1: Wide taper length sweep

It should be noted that this derivation considers only the coupling between the fundamental mode and the first higher order mode to which there is coupling due to symmetry. This is valid if the width of the waveguide is large.

For narrow tapers, the coupling to the radiation modes need to be considered as the limiting factor, and the condition 4.21 will not be valid. As another example, we simulate a single mode silicon taper from 200 nm to 400 nm width in Figure 4.2. In this case, the adiabaticity criteria would have predicted a taper length of 350 nm, which is clearly incorrect.

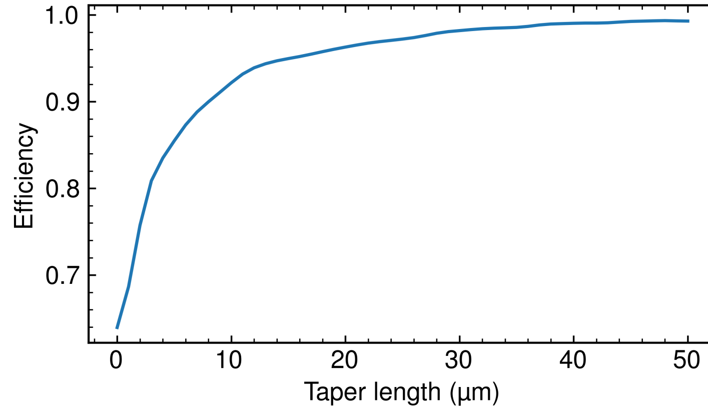


Figure 4.2: Narrow taper length sweep

Avoided crossings and double tapers

In a double taper, we use the index crossing between the modes in two waveguides to steer the energy between the waveguides. This works in a similar manner to adiabatic population transfer in quantum mechanics.

Here we consider a double taper where each waveguide supports only one mode. The waveguides are tapered such that the refractive index difference varies linearly

$$\Delta\beta = \alpha z \tag{4.22}$$

When the two tapers are in close proximity, there is an avoided crossing in the middle of the taper. The coupling between the waveguides is related to the index separation at the avoided crossing $C = \Delta\beta_{\text{crossing}}/2$.

With a similar development to the previous section, we get the system of equations

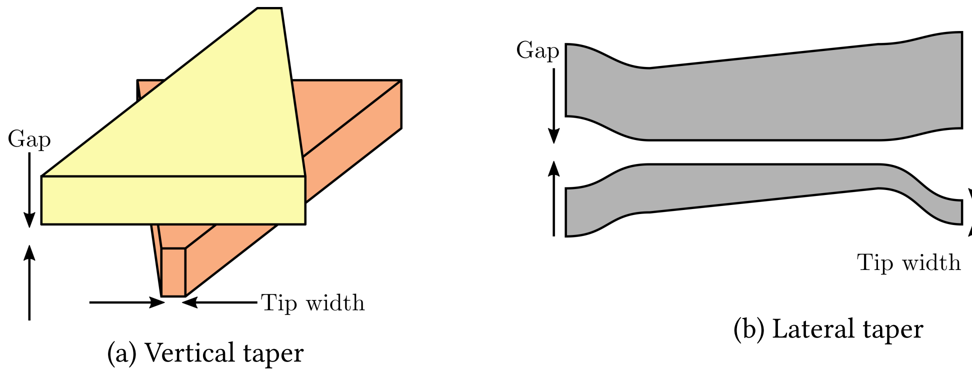
$$\frac{\mathrm{d}a_0}{\mathrm{d}z} = -Ca_1 e^{i\int \alpha z \mathrm{d}z} \tag{4.23}$$

$$\frac{\mathrm{d}a_1}{\mathrm{d}z} = Ca_0 e^{-i\int \alpha z \mathrm{d}z} \tag{4.24}$$

In such a system, the solution limit after the avoided crossing is the Landau-Zener equation [67]:

$$P = e^{-2\pi\frac{C^2}{\alpha}} \tag{4.25}$$

where P is the probability of light remaining in the first waveguide. This gives us some insight on how to make shorter double tapers. A short taper will have large α, which can be compensated by a stronger coupling C. However, the Landau-Zener solution assumes the system starts with modes far apart. In practice this is not always possible since the fabrication limits the smallest achievable taper width. Figure 4.3 shows two possible cases of integrated photonics double tapers. In the vertical case, if the taper tip width is not small enough, the two tapered waveguides need a larger spacing to prevent nonadiabatic coupling at the taper tip. In this case, the lithography resolution for the minimum line will then determine the smallest possible gap between the tapered waveguides. In the lateral taper case, the minimum gap is directly determined by the lithography resolution for the minimum space.



(a) Vertical taper

(b) Lateral taper

Figure 4.3: Minimum feature size of vertical and lateral double tapers

4.4 Bent waveguides

Bent waveguides can be represented as a straight waveguide with a conformal mapping of the refractive index. Several formulations exist, on of them being [68]

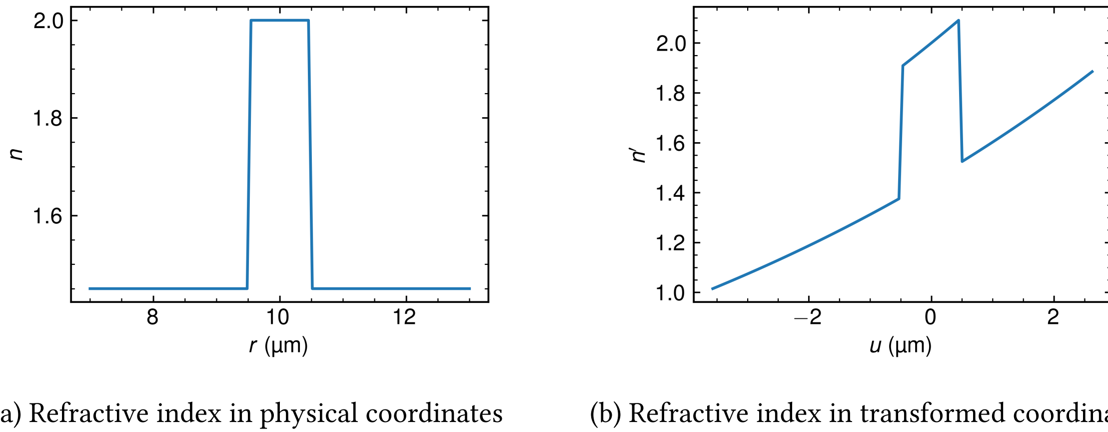
$$u = R \ln \frac{r}{R} \tag{4.26}$$

$$n' = n \exp \frac{u}{R} \tag{4.27}$$

where both the index n and the lateral coordinate r are changed to n' and u (Figure 4.4).

Alternatively, bent waveguides can be simulated in a more rigorous manner in cylindrical coordinates [69]. In this case Maxwell's equations are rewritten to account for the radius of curvature. Another exact method of simulating bent waveguides is by using an anisotropic conformal mapping in cartesian coordinates [70].

The first method gives us an intuition on how the mode profile might change in a bend. We notice that the mode is no longer fully confined: at some distance from the waveguide core, the effective cladding index exceeds the index of the core. This will introduce some loss which



(a) Refractive index in physical coordinates (b) Refractive index in transformed coordinates

Figure 4.4: Refractive index transformation of a waveguide with 10 µm bend radius

increases as the bending radius decreases. This loss can be simulated by adding an absorbing boundary condition on the simulation region. The bent waveguide propagation loss is usually the first factor determining how tight a waveguide can bend by imposing a minimum bend radius.

The conformal mapping method also gives us an intuition on how the modes get distorted in a waveguide bend: they get pushed towards the outside of the core, where the effective refractive index is larger. If the curvature of the waveguide changes abruptly (such as a transition between a straight waveguide and a round bend), there will be loss caused by a mode mismatch. One solution to this problem is to offset the waveguide at the junction of the bend [71].

Alternatively, one can use a continuously varying radius of curvature to create an adiabatic bend that will minimize the mode matching losses. One popular method is to use an Euler spiral combined with a constant radius in the middle of the bend (see e.g. [72]).

More generally, the bend shape can be described with a curvature as a function of propagation distance. This can easily be combined with an optimization algorithm to generate the optimal bend. The optimal bend shape will depend on several factors, such as the effective radius (dictated by space constraints), the total bending angle, and the waveguide width. For narrow waveguides, radiation losses should be minimized and thus the optimal bend will be close to circular. For wider multimode waveguides, it is desirable to minimize conversion to higher modes. In this case, a slower variation of the radius of curvature is desirable.

For example, we compare two bend shapes (circular and adiabatic Euler spiral) for a narrow (400 nm wide) and a wide (1 µm wide) silicon photonics rib waveguide, with 10 µm effective radius. The bend shapes are shown in Figure 4.5, and the bending losses for a 90° bend are listed in Table 4.1. In the narrow waveguide case, the circular bend is superior since it minimizes radiation losses, however the adiabatic bend is advantageous for the wide waveguide bend since it minimizes the excitation of higher order modes.

In the chip layout, the different components can be routed using algorithms such as the Dubins path [73], which finds the shortest path between two oriented points using straight lines and arcs of constant (minimum) radius.

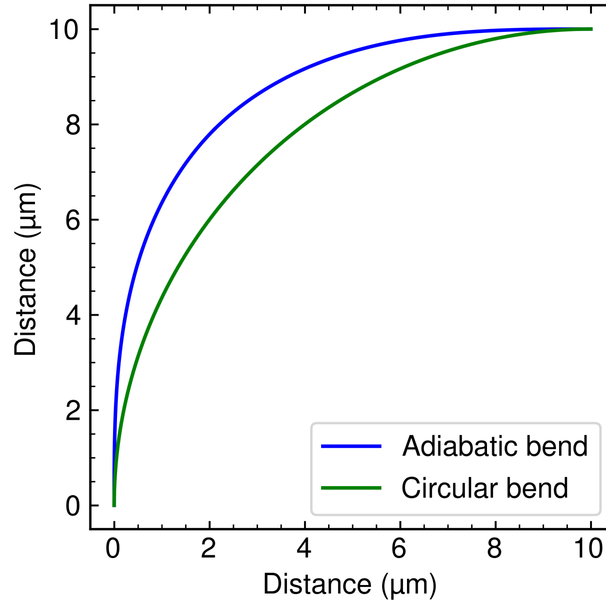


Figure 4.5: Waveguide bend shapes

Bend type	400 nm waveguide loss (dB)	1 µm waveguide loss (dB)
Circular	-0.43	-0.96
Adiabatic	-1.7	-0.14

Table 4.1: Bending loss for narrow and wide waveguides of different shapes

4.5 TE0-TE1 mode converter

A basic building block of integrated photonics is the TE0 to TE1 mode converter. It can be used for example in polarization splitter-rotators [74], efficient modulators [75], or in large bandwidth spectral splitters [76].

The TE0-TE1 mode converter works by coupling two waveguides of different widths, and having a mode crossing between the fundamental (TE0) mode in the smaller waveguide and the TE1 mode in the wider waveguide. An example is given in Figure 4.6. The dimensions are given in Table 4.2, and ensure that the left port has the TE1 mode dominant, while the right port has the TE0 mode dominant (ignoring the TE0 mode of the wide waveguide).

On the left hand side, the input is in TE1 mode, while it is converted to the TE0 mode on the right-hand side. For a silicon photonics rib waveguide and a gap of 250 nm, the TE0-TE1 converter minimum length is around 150 µm. The length can be further reduced using nonlinear tapers (Section 4.7).

Port	Width (μm)
Wide 1	1.0
Wide 2	0.6
Narrow 1	0.25
Narrow 2	0.4

Table 4.2: TE0-TE1 mode converter dimensions

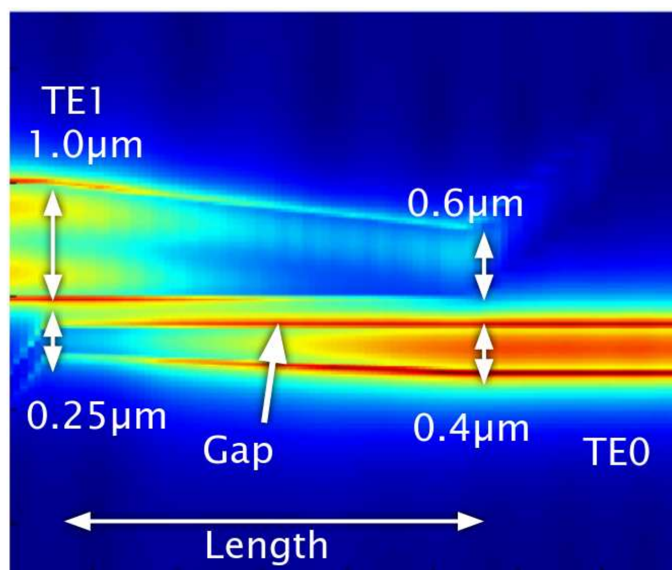


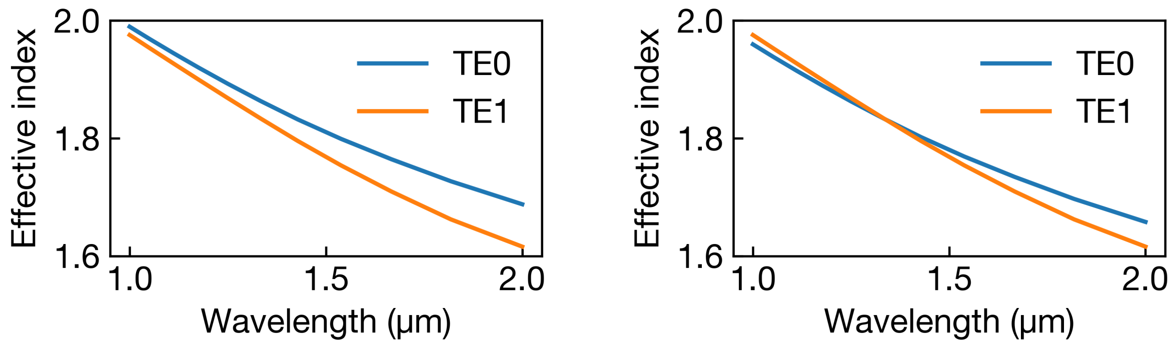
Figure 4.6: TE0-TE1 mode converter

4.6 Spectral filters

One application of adiabatic components is for broadband wavelength selective filters. In this case, we use dispersion between two modes to create a wavelength-dependent mode crossing. This has been shown using waveguides of different materials [77], or using a combination of rectangular and slot waveguides [78]. Another method is to change the coupling strength between the waveguide tapers as a function of wavelength [79]. Since adiabatic wavelength filters do not rely on resonance, their optical bandwidth is inherently very large.

Another method to create modal dispersion is by using higher order modes in waveguides of similar material but different widths. For example, the TE0 mode in a narrow waveguide and TE1 mode in a larger waveguide may have an effective index like shown in Figure 4.7. The wavelength of the mode crossing can be changed by using different waveguide widths. For example, a smaller narrow waveguide will lower the TE0 mode index and push the index crossing to the right. Effectively, this acts as a TE0-TE1 mode converter that operates only at certain wavelengths.

Several of these splitters can easily be cascaded to create a $1 \times n$ spectral splitter. This concept



(a) Modal dispersion at the input of the splitter (b) Modal dispersion at the output of the splitter

Figure 4.7: Modal dispersion between the fundamental and second TE mode in two waveguides of different widths

is shown in Figure 4.8. The last wavelength can be converted back to the fundamental TE0 mode by using a splitter similar to the first one, in reverse.

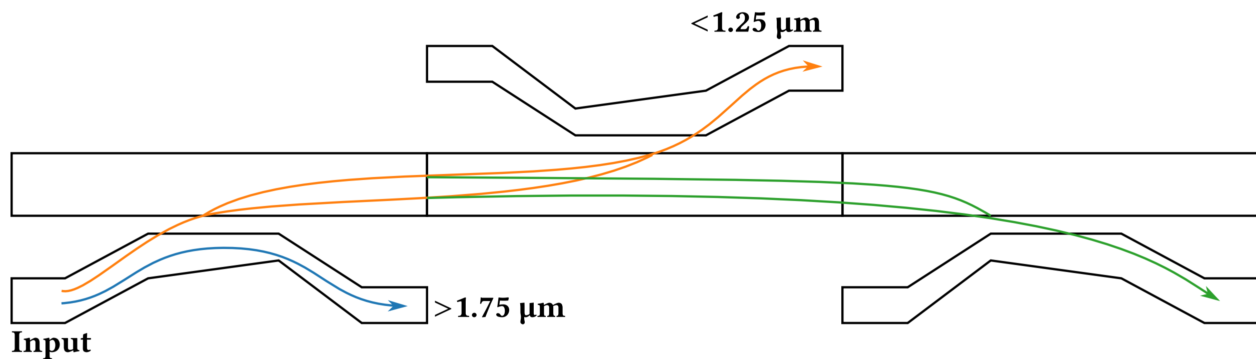


Figure 4.8: Adiabatic spectral splitter concept

We simulate a 1×3 splitter operating at 1 μm, 1.5 μm and 2 μm in Figure 4.9. The crossing wavelengths are 1.25 μm and 1.75 μm. In the mid-wavelength port, the extinction ratio is greater than 10 dB for the extreme wavelengths.

We fabricated adiabatic spectral splitter test structures in chalcogenide glass waveguides. An example measurement for a splitter designed at 1.25 μm is shown at Figure 4.10. The fabricated device was analyzed using SEM cross section and was re-simulated to account for fabrication variations, with a predicted 3 dB point around 1.3 μm. The device spectrum was measured using a series of discrete lasers operating between 980 nm and 1.61 μm. The 3dB point shows good matching between the measurement and simulation, and the extinction ratio is better than 20 dB.

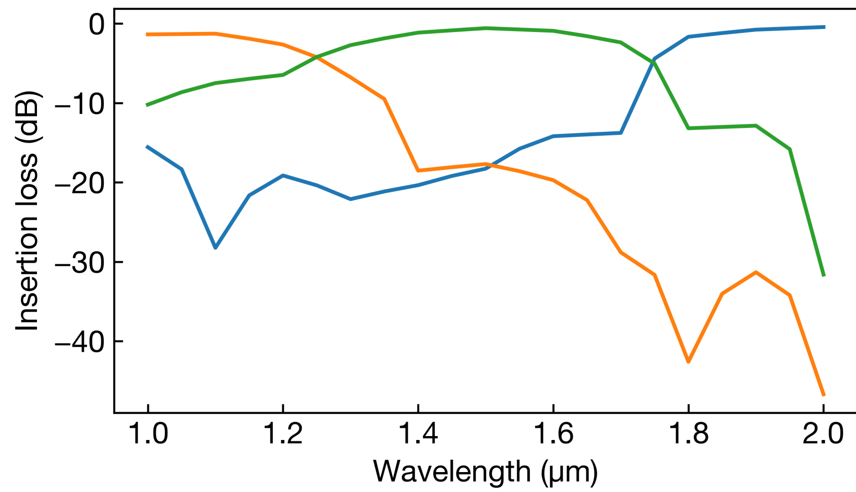


Figure 4.9: Simulation of a 1×3 spectral splitter

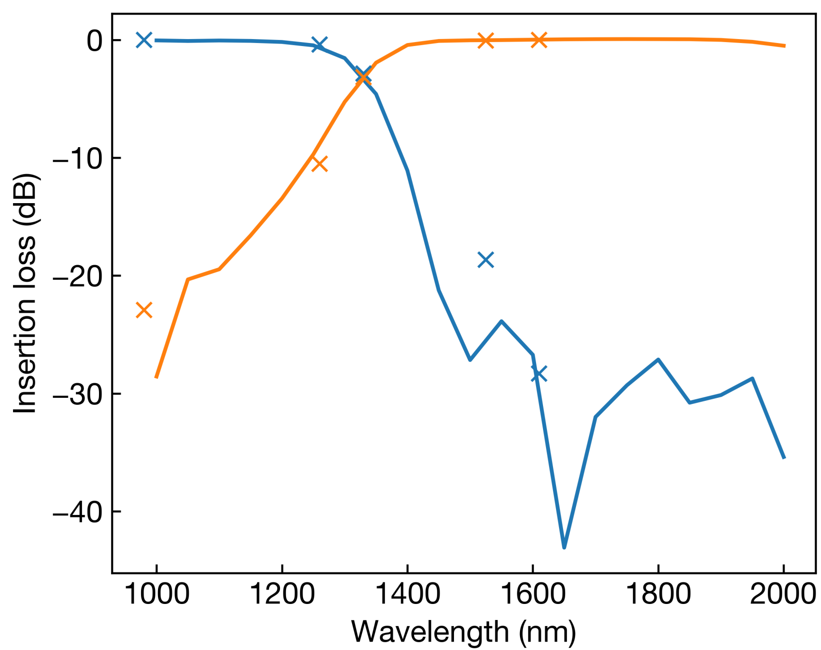


Figure 4.10: Adiabatic spectral splitter performance. Solid line: simulation. Crosses: measurement.

4.7 Optimization of nonlinear tapers

The use of nonlinear tapers can help to reduce the length of adiabatic components. In order to use common optimization algorithms (such as L-BFGS-B available in SciPy [80]), the taper shape needs to be described using a small number of parameters. Since we want the taper to have a well defined shape, we choose to use monotonic splines as a basic function. Such splines are defined as the integral of non-negative M-splines, and are called I-splines [81]. The M-splines and I-splines have an order k and have a number of free parameters n. They will also have $n-k$ nodes, the location of which can also be optimized. There are another $2k$ fixed nodes at the boundaries, for a total of $n+k$ nodes.

The n M_i-splines with nodes t_i are defined by the following recursive relation:

$$M_i(x|1,t) = \frac{1}{t_{i+1} - t_i} \tag{4.28}$$

$$M_i(x|k,t) = \frac{k\left((x-t_i)M_i(x|k-1,t) + (t_{i+k}-x)M_{i+1}(x|k-1,t)\right)}{(k-1)(t_{i+k}-t_i)} \tag{4.29}$$

The I-splines can also be defined in terms of the M-splines with

$$I_i(x|k,t) = \sum_{m=1}^{k} (t_{m+i+k+1} - t_{m+i}) M_{m+i}(x|k+1,t)/(k+1) \tag{4.30}$$

for $t_i < x < t_{i+k}$, 0 for $x \leq t_i$, and 1 for $x \geq t_{i_k}$.

For example, we optimize the TE0-TE1 mode converter using 1 free node and order 2, i.e. $k=2$ and $n=3$. The shape is optimized using the L-BFGS-B algorithm, and the merit function is the efficiency calculated using eigenmode expansion. The final shape is shown in Figure 4.11. The $\times$ symbol represents the location of the node. Intuitively, we would expect the taper to slow down where the mode crossing occurs, which is what this optimization converges to. The optimal TE0-TE1 converter can be made shorter than the linear taper version. In Figure 4.12, we sweep the converter length (scaling the shape proportionally) to determine a minimum length of 100 µm (instead of 150 µm for the linear taper).

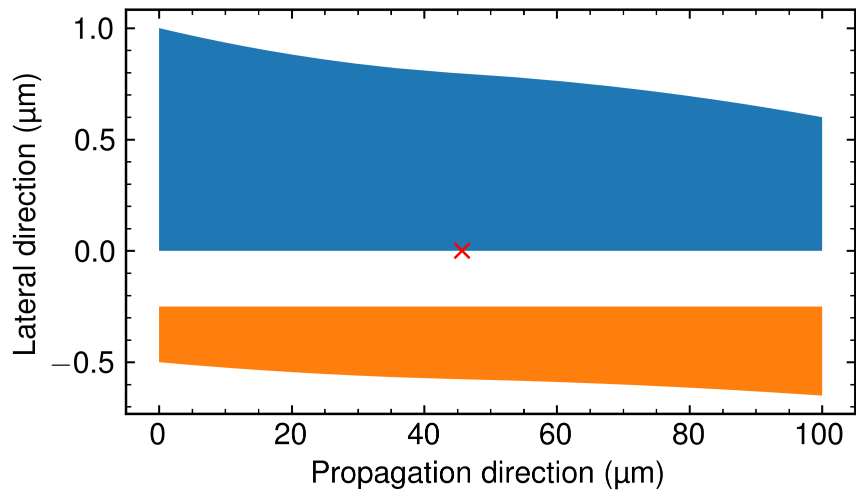


Figure 4.11: Optimized taper shape for the TE0-TE1 converter

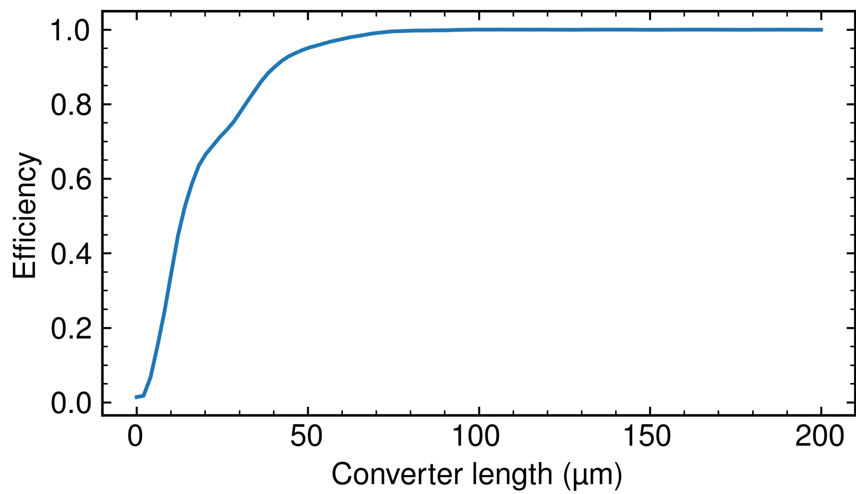


Figure 4.12: Length sweep for the optimized TE0-TE1 converter

Chapter 5

Edge coupling

In Chapter 2, we presented a concept of an on-chip self-referenced frequency comb system which uses different integrated photonic platforms for different applications: indium phosphide for lasing, chalcogenide glass for supercontinuum generation, and lithium niobate for frequency doubling. These chips need to be coupled optically, and a straightforward manner is to couple the light from the edge of the chip.

In the first part of this chapter, we explain a fabrication process for dicing silicon substrates that is suitable for coupling to the edge of the chip, in particular for coupling to another chip. In the second part, we discuss different methods for coupling from the edge of the chip to an optical fiber.

5.1 Chip-to-chip edge coupling

When singulating dies intended for edge coupling, care must be taken to create optical quality facets. This can be done with traditional mechanical dicing followed by careful facet polishing with a precise fixture and a fine lap. However, this method is time consuming and does not produce chips of precise dimensions.

We have explored using laser dicing technologies such as stealth dicing, which works by creating micro cracks in the silicon substrate using two-photon absorption from an infrared laser. The dies are then separated by adding tension on the dicing tape. This method produces very little particles and high quality facets. Moreover, it can dice chips with minimal kerf and almost arbitrary shapes.

However, we ran into two major issues when using stealth dicing:

- The dicing edge is not perfectly straight, especially when the laser crosses structures of varying reflectivity (such as waveguides stitched across dies).
- The dicing cross section is not perfectly vertical. The cutting is initiated from microcracks in the substrate, but the cracks need to propagate a certain length before they follow a crystal plane.

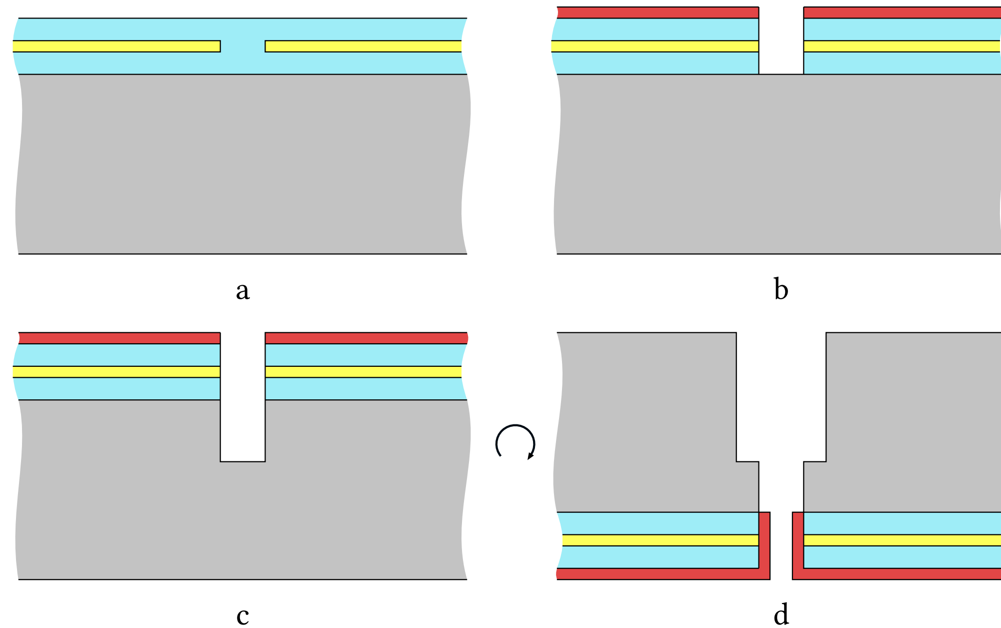


Figure 5.1: Optical chip dicing process. (a) Device wafer ready for dicing. (b) Trench etched through the oxide. (c) Trench etched partially through the wafer. (d) Back-side mechanical dicing to separate the dies.

The first issue can be mitigated by dicing from the back side of the wafer, however this requires further processing such as polishing or grinding of the back side. Nonetheless, because of the second issue, it is typically impossible to get two chips diced in this manner to touch perfectly (less than 1 µm gap) over the whole length of their edges.

To prevent these issues, we developed a dicing process using lithography and plasma etching. While plasma dicing is not a new process, when etching thick (more than 2 µm) cladding the selectivity of the photoresist in the oxide etcher prevents from using the same mask to etch both the cladding and the totality of the substrate thickness. We singulate the dies after a partial etch through the wafer by mechanical dicing from the back of the wafer. This process also creates a nice overhang of the optical facet to prevent the substrate from interfering during alignment.

The dicing process is summarized in Figure 5.1. It is divided in three major steps:

1. Define the trench mask lithographically
2. Etch through the cladding using reactive-ion etching, then through part of the substrate using deep reactive-ion etching (DRIE)
3. Singulate the dies from the back side using a traditional dicing blade, but only partway through the wafer

The trench etch mask has some particular requirements. While the resolution is coarse (on the order of 10 µm to 100 µm), the mask needs to be straight and have low roughness. Both contact

alignment and steppers can achieve this. Some photoresists may reflow during the final baking process, and this can cause distortions at the corners of the dies.

The photoresist mask also needs to be thick enough to withstand the long etching process. In particular, the selectivity between most photoresists and oxide dry etching chemistries is low (2:1 to 3:1). There should be significant photoresist remaining at the end of the etching to prevent erosion and excessive sidewall angle during etching. For silicon photonics devices where the cladding thickness is around 3 µm, a DUV photoresist such as UV26 with a thickness of 4.7 µm is suitable. For devices with lower index contrast waveguides where the cladding thickness can reach 10 µm, we use the AZ P4620 photoresist with a thickness of 8 µm. This latter resist requires either g-line or i-line exposure.

Finally, the dicing mask requires submicron alignment accuracy. In the layout, some margin is left between the tip of the waveguide and the edge of the chip to prevent erosion of the waveguide during the dicing process. The maximum acceptable margin will depend on the divergence angle of the mode at the end of the waveguide. The alignment accuracy of the lithography should be within this margin.

When using a DUV photoresist, the DUV stepper lithography can fulfill all these requirements. When a thicker photoresist such as AZ P4620 is necessary, either a contact aligner or an i-line steppers could be used. In the Nanolab, the GCA8500 i-line wafer stepper has more robust alignment capabilities than the Karl Suss mask aligner.

The resulting edge is shown on the electron microscope image of Figure 5.2. The SiO_2 cladding is vertical and smooth, while the silicon substrate shows the typical scalloping from the deep RIE process. This is not an issue because the silica cladding overhangs above the bulk silicon roughness.

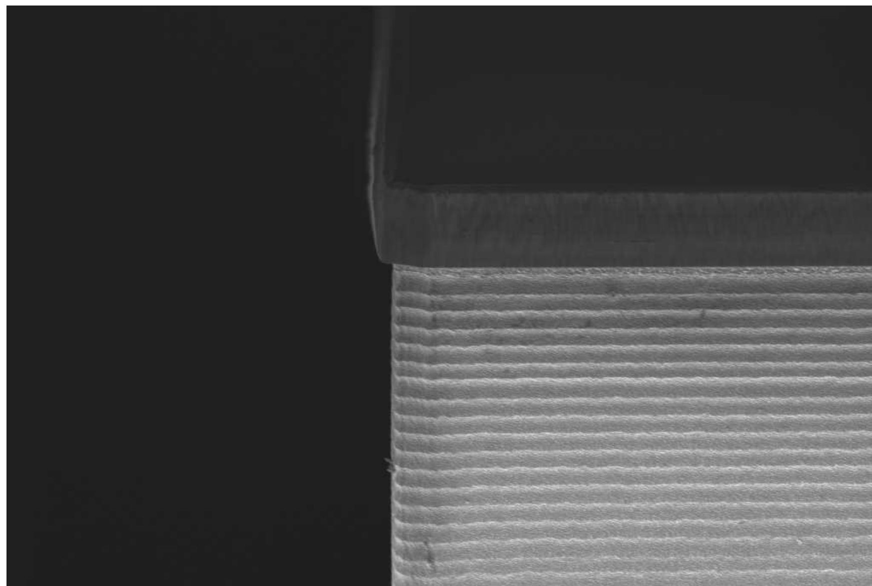


Figure 5.2: Optical facet in SiO_2 after trench etching

As a proof of concept, we packaged a silicon nitride test chip to an indium phosphide laser chip using an automated alignment machine (FiconTEC FL300). Figure 5.3 shows a top-view microscope image after curing the epoxy. The laser chip has three channels: one device channel, one alignment laser channel, and finally a detector channel which is reverse biased to monitor the coupling efficiency during alignment. After the epoxy curing, the chip-to-chip gap is less than 2 μm, and the coupling efficiency is typically 3 dB per facet.

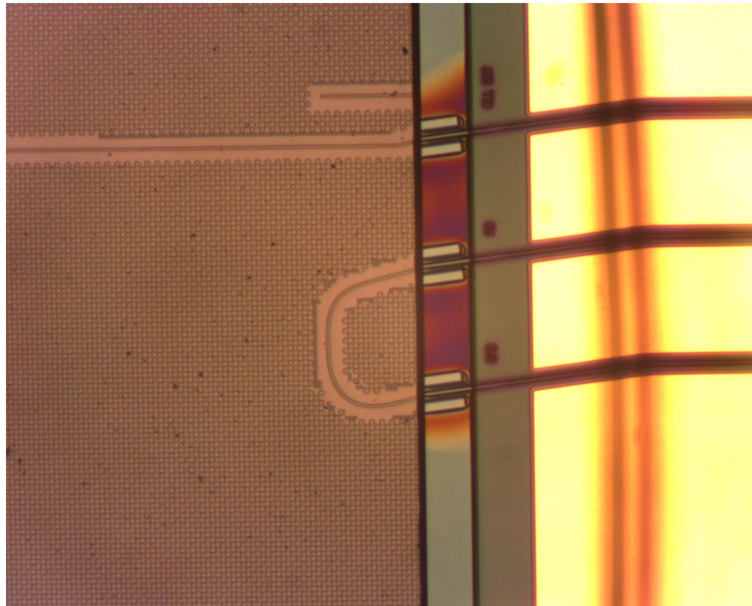


Figure 5.3: Silicon nitride chip edge coupled to an indium phosphide laser chip

For chip to chip alignment with edge coupling, the design of the mechanical carrier (or submount) is critical. Often the chip thicknesses will be different. For example, in Figure 5.4, we align an indium phosphide DBR laser array with a thickness of 100 μm to a silicon photonics chip with a thickness of 650 μm. The outer lasers on the DBR array are used as alignment signal (with one laser reverse biased as a detector), so the InP chip is mounted first to the carrier. A clear adhesive is dispensed and cured at the chip interface to act as index matching material and mechanically stabilize the assembly before the structural adhesive is dispensed and cured. There is some tolerance builtin the submount step height between the DBR array and the silicon chip to account for structural glue thickness variation and the machining tolerance. Furthermore, some adhesive overflow grooves are machined in the submount in case an excess of adhesive is dispensed. Because of the machining tolerance, the glue thickness may vary from 25 μm to 50 μm. This requires the use of low-shrinkage adhesive. For example, 0.08 % shrinkage will result in a 40 nm vertical shift during curing. More importantly, the adhesive should be dispensed and cured evenly to prevent lateral shifts caused by stress in the adhesive. Some solutions include curing more slowly, or curing using multiple UV sources simultaneously.

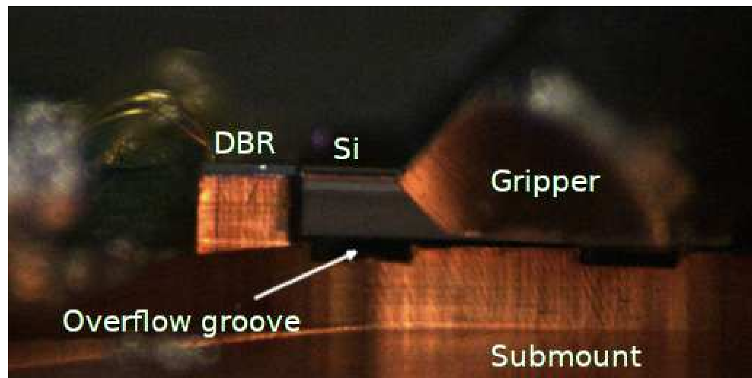


Figure 5.4: Mechanical carrier for chip-to-chip alignment

5.2 Chip-to-fiber edge coupling

Coupling light from the edge of the chip to a fiber is at first glance intuitive: the light exits the on-chip waveguide and enters straight into the fiber. However, a few important things need to be taken into consideration:

- The size and shape of the waveguide mode must match the fiber mode to get efficient coupling
- Both the fiber and the waveguide mode size must be large enough to account for alignment tolerances
- The mode at the chip edge must be converted efficiently to the typically smaller on-chip mode.
- The waveguide at the chip edge must be single mode to prevent coupling to higher-order modes

To give a few examples, typical silicon photonics waveguides are 450 nm wide and 220 nm tall, while lower index materials such as silicon nitride and chalcogenide glasses have a core size around 1 µm square. On the other hand, standard SMF-28 fibers have a mode field diameter around 9 µm, high NA fibers 4 µm and lensed fibers down to 2 µm.

Low NA waveguides

The concept of numerical aperture is often used to characterize optical fibers:

$$NA = \sqrt{n_1^2 - n_2^2} \tag{5.1}$$

$$NA \approx \sqrt{2n\Delta n} \tag{5.2}$$

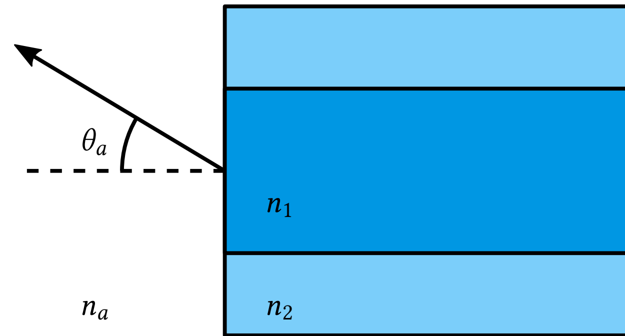


Figure 5.5: Numerical aperture definition

The numerical aperture is related to the index contrast between the core and the cladding, and defines the beam divergence out of the fiber θ_a (Figure 5.5). A similar concept can be applied to rectangular waveguides, although the numerical details will vary.

Low NA waveguides (such as SMF-28 fiber) can be made larger while still remaining single mode. This is desirable because it facilitates alignment:

- Large modes have better lateral misalignment tolerance
- A smaller beam divergence has better tolerance to the gap between the fiber and the chip

Integrated low NA waveguides can be fabricated in different ways. The first is to use low index contrast materials, such as polymers, SiON/SiO_2, phosphosilicated glass, and so on. Another method is to use sub-wavelength structures to effectively create materials with different indices. Such examples include thin SiN rods and sheets, and silicon subwavelength gratings [82, 83, 84].

We have explored these sub-wavelength structures with thin SiN sheets. They can be fabricated in a simple manner by alternating thick SiO_2 cladding and spacer layers and thin SiN core layers. The SiN layers are dry etched in a few seconds and the process is repeated for the total number of layers.

After choosing the number of layers, the structure can be optimized by varying the following SiN layer parameters: layer width, layer thickness and layer spacing. We found that 2 layers provide a large enough solution space to have more than 96% to single-mode fibers.

An example of such a waveguide was fabricated with 20 nm thick and 8 μm wide silicon nitride layers. The fabricated device is shown in Figure 5.6. We measured facet insertion loss as low as 0.3 dB and propagation loss of 0.8 dB/cm.

Mode size converter array

The thin silicon nitride waveguide concept can be extended to fabricate mode size converter arrays, which are suitable to couple to a single mode fiber on one end, and to a silicon inverse taper on the other end. Indeed, we can add another set of thicker layers that are optimized e.g. for a 4 μm mode size. The layers can be interleaved for better coupling (Figure 5.7).

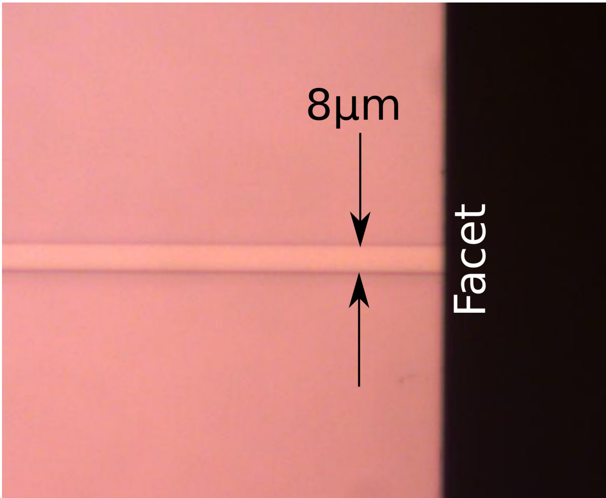


Figure 5.6: Waveguide matched to fiber fabricated with two thin SiN layer

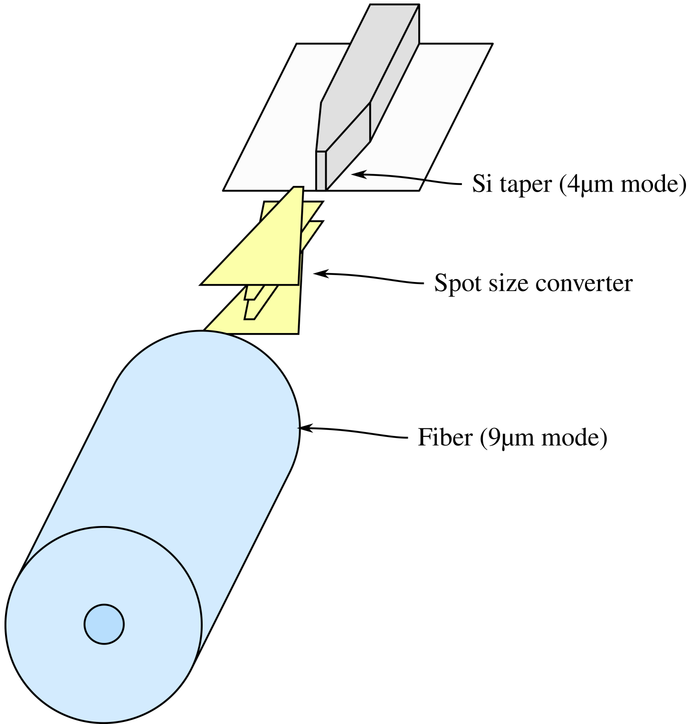


Figure 5.7: Edge spot size converter between silicon taper and single mode fiber

In this example, the inner layers have a thickness of 50 nm, a width of 1 µm and a separation of 270 nm. They are tapered down to 250 nm minimum width where they meet the outer layers that are matched to fiber mode size.

The fabrication process for this device is straightforward, alternating between cladding deposition and thin silicon nitride layer patterning (Figure 5.8). The chips are diced using the process described in section 5.1.

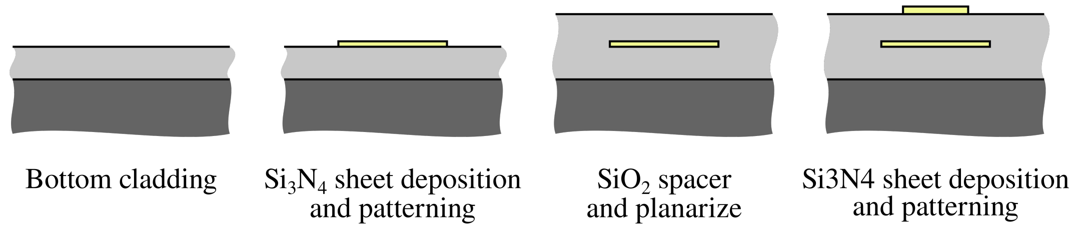


Figure 5.8: Edge spot size converter fabrication process

A array of mode size converter is defined lithographically, such as in Figure 5.9. On the left side are the waveguides matched to fiber, and on the right side the waveguides are matched to the silicon taper. The length of the mode converter array (8 mm) is limited by the bending radius in the pitch reduction section, which is 1 cm.

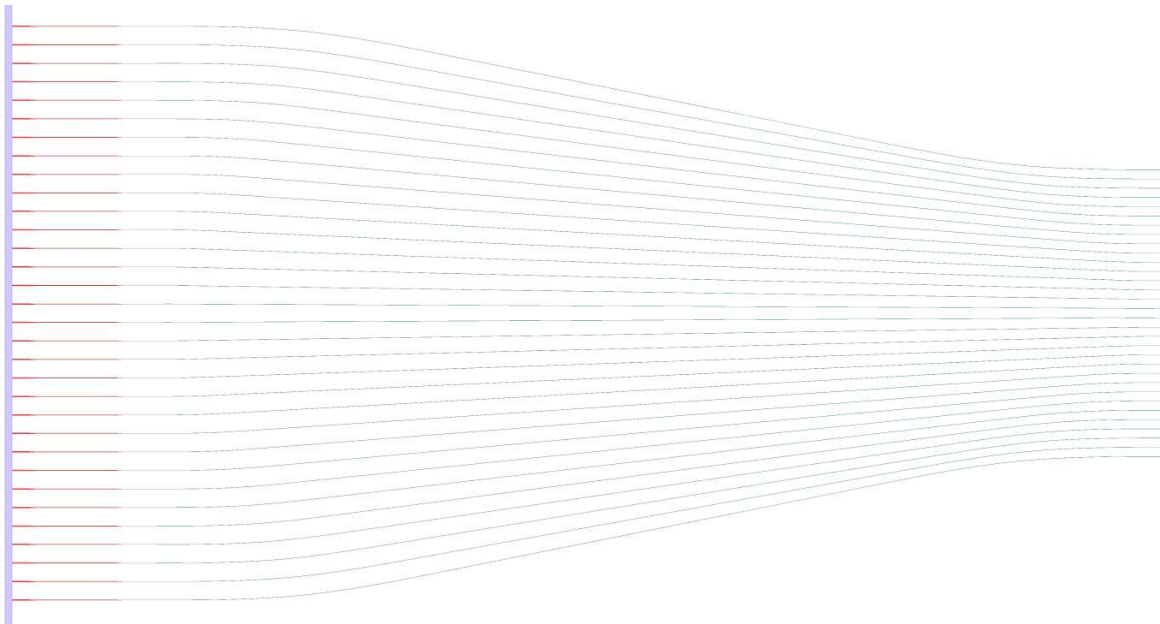


Figure 5.9: Mode size converter array layout

5.3 Substrate leakage

When creating large, dilute modes on-chip, care must be taken to avoid propagation losses due to substrate leakage. When we include a high index substrate (such as silicon) in the waveguide domain, strictly speaking the waveguide mode is not guided anymore. The mode that appears to be localized around the waveguide core mathematically has exponential growth at infinity [85]. Unlike radiation modes, which have a purely imaginary propagation constant, leaky modes have a complex propagation constant. Intuitively, this corresponds to a guided mode with a certain amount of propagation loss. When launching light into the waveguide core, the intensity will initially decrease exponentially with a length given by the imaginary part of the propagation constant.

Indeed, an equivalent formulation to this problem is to consider a finite domain composed of the waveguide core and part of the substrate, with some form of reflection-free absorbing boundary condition (such as perfectly matched layers). The propagation constant in this case is numerically equivalent to the infinite problem. This analysis is also valid for other leaky waveguide problems, such as bent waveguides.

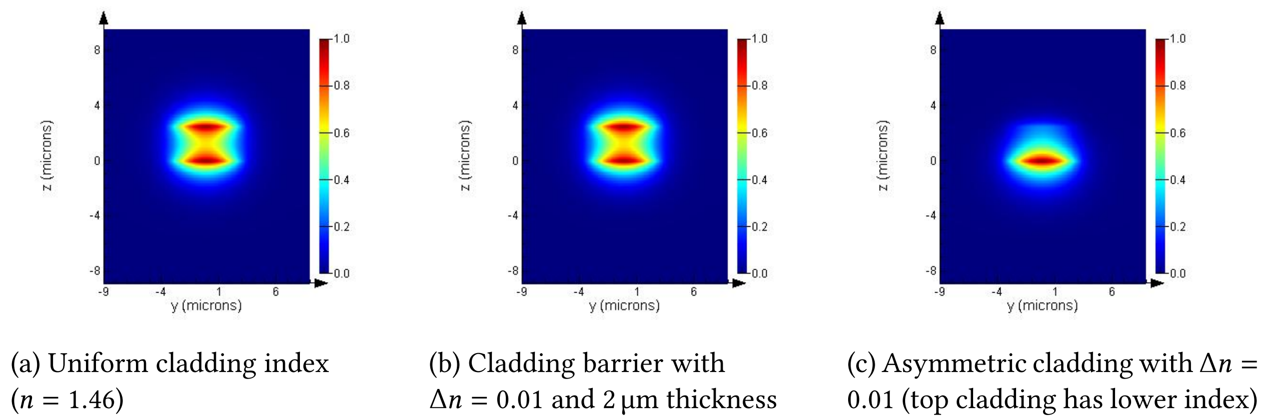
For conventional silicon photonics, a 2 µm bottom cladding layer might be sufficient to prevent substrate leakage. For silion nitride photonics, we use a bottom cladding of 3 µm. For an edge coupler that would couple to a 10 µm single mode fiber mode, at least 6 µm of bottom cladding (and equivalent top cladding) is necessary.

In some cases it might not be desirable to fabricate such a thick cladding. We can borrow some ideas from quantum mechanics and add a cladding barrier layer of lower refractive index to effectively increase the substrate distance by making the evanescent fields decrease quickly in the cladding barrier.

Another issue when dealing with dilute waveguide modes is the asymmetry of top and bottom cladding. From the asymmetric slab waveguide theory, we know that under some condition there might not be a guided mode. In addition, even when the waveguide mode is guided despite the asymmetric cladding, the mode might be pushed towards the substrate which increases the leakage. This might happen for example when the two claddings are deposited using different processes, such as LPCVD vs PECVD.

The effect of the cladding refractive index distribution is illustrated in Figure 5.10. We simulate the quasi-TE mode created by two thin SiN layers, surrounded by an SiO_2 cladding. The dimensions are chosen to closely match the mode size of a single mode fiber:

- SiN layer thickness: 20 nm;
- SiN layer spacing: 2.5 µm;
- SiN layer width: 6.5 µm;
- Cladding thickness: 6 µm below the bottom SiN layer and above the top SiN layer;
- Silicon substrate below the bottom cladding, and air above the top cladding.



(a) Uniform cladding index ($n = 1.46$)

(b) Cladding barrier with $\Delta n = 0.01$ and 2 µm thickness

(c) Asymmetric cladding with $\Delta n = 0.01$ (top cladding has lower index)

Figure 5.10: Effect of cladding index on mode shape and substrate leakage

In 5.10a, the cladding has a uniform refractive index of 1.46. The substrate leakage loss is 0.4 dB/cm. In 5.10b, the bottom 2 µm of the cladding are replaced with a barrier having an index of 1.45. This corresponds to $\Delta n = 0.01$. In this case, the substrate leakage is reduced to 0.1 dB/cm. Finally, in 5.10c an asymmetric cladding is simulated by using $n = 1.45$ for the top cladding. We observe that the mode is pushed towards the substrate, and the substrate leakage is now increased to 1.2 dB/cm. The effect is similar for the TM polarization.

The asymmetric cladding is an undesirable effect that appears when the top and bottom cladding are deposited using different methods or conditions. However, in the case of the cladding barrier, we want to be able to precisely control the refractive index different. This can be done by using doped glass, e.g. phosphosilicate glass (PSG) deposited using LPCVD. The refractive index can be tuned by modifying the gas ratio of phosphine to silane. PSG has the additional advantage of reduced stress compared to undoped LPCVD SiO2. This is desirable especially when fabricating waveguides with a very thick cladding (such as waveguides matched to single mode fibers) where the stress can induce excessive wafer bow or wafer expansion.

Chapter 6

Silicon photonics MEMS switch packaging

With datacenter switch scaling limited by the power consumption of electronics, optical switches have been proposed to increase the bandwidth while improving the efficiency [86]. In particular, silicon photonics is an attractive platform since it can leverage the reliability and scalability developed by electronic foundries. Several methods can be used to switch light in silicon photonics, including electro-optic modulation [87] and and thermo-optic modulation [88]. Even though the electro-optic modulation switches have the advantage of very fast (10 ns) switching times, both schemes suffer from poor scalability because of the cascaded loss of the individual 2×2 switch elements and the analog nature of the electronics required to control them.

In contrast, silicon photonics MEMS switches using a crossbar architecture have been shown to scale up to wafer-size [89]. The basic crossbar architecture is shown in Figure 6.1.

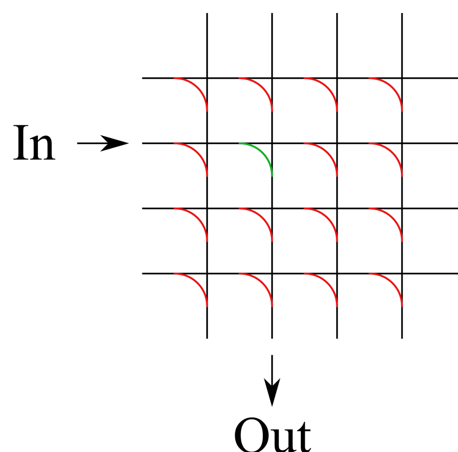


Figure 6.1: Crossbar switch architecture

Each input and output is divided in row and column bus waveguide, with waveguide crossings designed to be low-loss (see e.g. [90]). At each intersection, a vertical MEMS actuator with an

adiabatic coupler is used to turn the light to the required output. Only one actuator per row and column is used at any time. The scaling of this architecture is mainly limited by the waveguide propagation loss and the crossing loss, while the crosstalk remains very low because the adiabatic coupler gap is large in the off state.

In this chapter we first study the electrical and optical packaging of a 4×4 silicon photonics MEMS switch. In the next chapter we will discuss the design of evanescent couplers to improve the coupling efficiency to fibers for wafer-scale switches.

6.1 4×4 silicon photonics MEMS switch

The 4×4 silicon photonics MEMS switch chip was fabricated using a commercial foundry process and released using HF vapor in the UC Berkeley Nanofabrication Laboratory. The 16 switch elements are connected to individual wirebonding pads, and the optical channels are routed to grating couplers. The die is approximately 4 mm by 4 mm in size, and it operates in the O-band.

Figure 6.2 shows an example switch die where the switch elements are visible. The image also shows the pads with wirebonds, and epoxy dispensed on top of the grating couplers for packaging the fiber array.

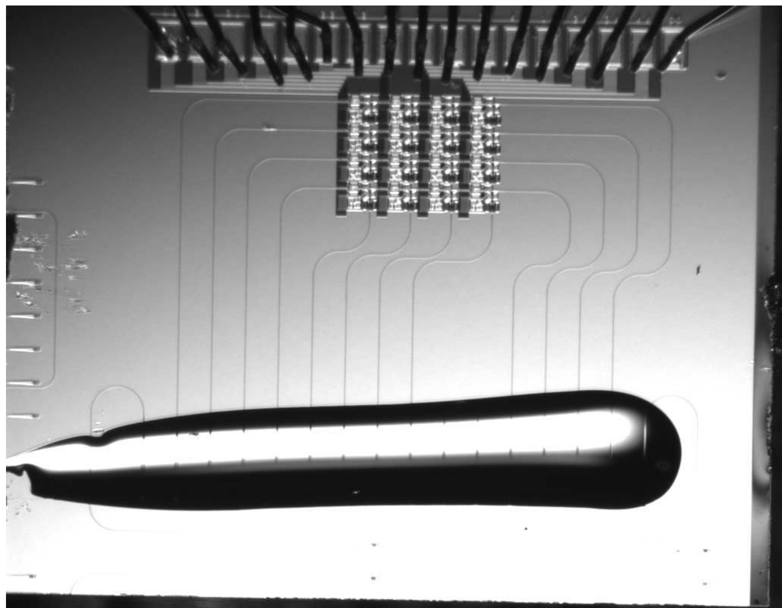


Figure 6.2: Epoxy dispensed on 4×4 switch

Figure 6.3 shows the same die, attached to the PCB. The electrical pads are wirebonded, the fiber array is aligned and the epoxy is cured. This PCB then connects to another PCB with the control electronics via a flat flexible cable.

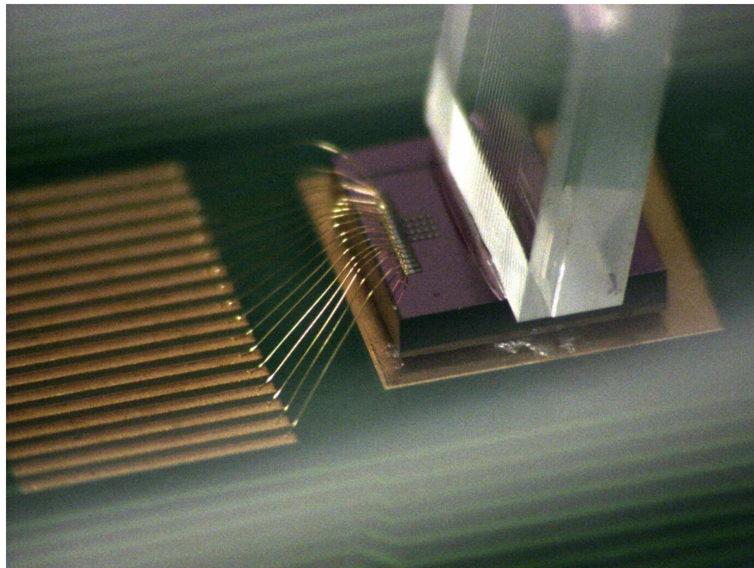


Figure 6.3: Fiber glued to 4×4 switch

Then, the PCB and fiber array are securely mounted inside a box, as in Figure 6.4. The front panel holds the fiber connectors and the USB connector for power and control.

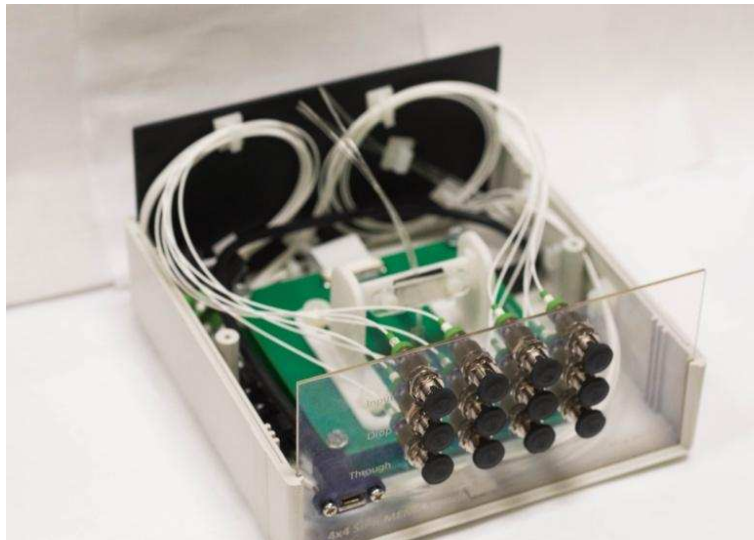


Figure 6.4: 4×4 switch packaged inside box

The packaged switch has a total insertion loss of 10 dB, of which 2 dB is on-chip, and each grating coupler contributes to 4 dB insertion loss. A system experiment is shown in Figure 6.5, with the trigger electrical signal on top, and the switch optical output at the bottom. The optical rise and fall time are less than 1 µs, with a latency around 1 µs caused by the microcontroller operation.

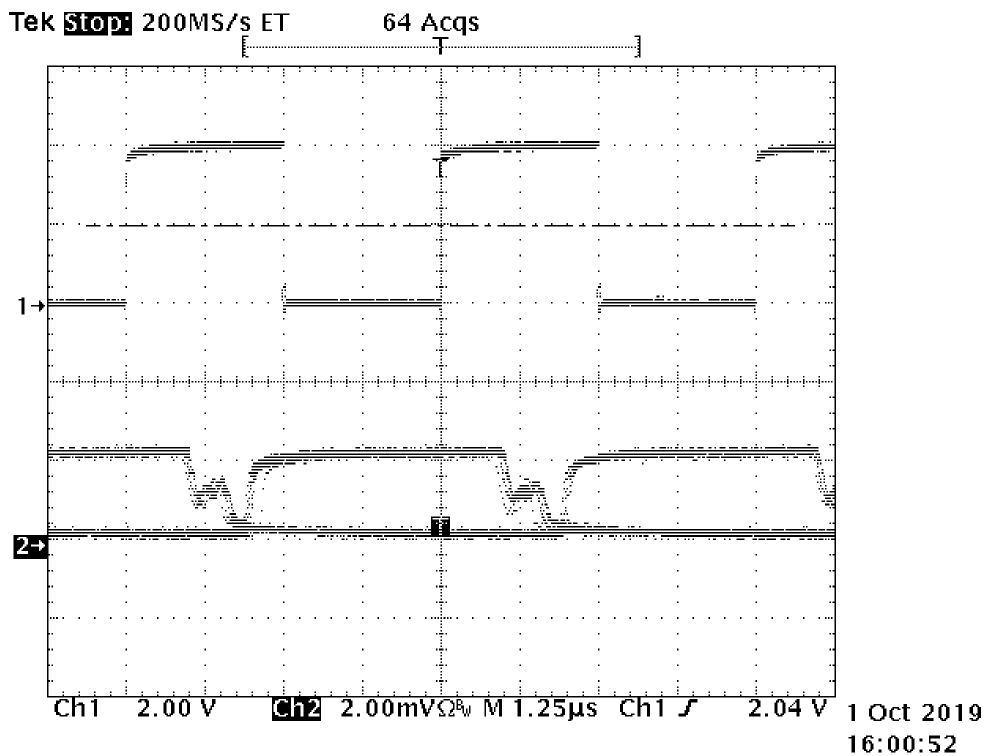


Figure 6.5: 4 × 4 switch optical response time. Top: trigger, bottom: optical signal

6.2 Electrical packaging

Here we review some concepts of electrical packaging, which are relevant for the silicon photonics MEMS switch: wirebonding, and electrical driver circuits.

Wire bonding

Wire bonding is an ubiquitous method for making electrical connections to photonic ICs. For simple prototype devices requiring only a few signals, using a manual wirebonding is simple and fast. For larger ICs (up to hundreds of signals), several vendors offer automated wirebonding services. Two very common materials for wirebonds are aluminum and gold.

Aluminum wirebonds can be performed easily at room temperature. They can easily adhere to common PCB substrates such as ENIG. On the other hand, gold wirebonds require heating of the filament and substrate for optimal bond strength. For commercial products, gold has the advantage of better corrosion resistance. If wirebonding to a PCB, soft gold or ENEPIG pad finish is a good option.

Electrical drivers

Some applications (e.g. large scale MEMS optical switches) require arrays of high voltage drivers to control each element individually. We have developed two circuits for this purpose which will be explained in this section.

The first circuit is a 16 channel driver for a 4×4 switch array. It uses the HV7620 IC from Microchip, which is a 32 channel, high voltage serial to parallel converter. The IC outputs can operate at up to 200V. The circuit also has a high voltage step-up converter powered from a 5V source. We built two variants of this board:

- The first one uses an on-board microcontroller (STM32F303) to control the high voltage driver. The high voltage supply is a coupled-inductor boost converter using the MCP1661 IC (Figure 6.6). It can deliver up to 60 V and the voltage is trimmed using a signal from the microcontroller DAC. The voltage ripple at 50 V output is ±1 V. This board is used mainly for testing the functionality of the high voltage driver and the switch. The switch program is sent via USB, and an external trigger is used to synchronize the microcontroller with external circuits.

- The second version breaks out the HV7620 signals to control directly from an FPGA board. The high voltage supply was changed to a voltage doubling boost converter using the LT8330 IC (Figure 6.7), which switches at a higher frequency to reduce the supply ripple current. The voltage doubling topology is also more flexible regarding the inductor selection. On this board, the output voltage is preset using a fixed resistor. The voltage ripple at 60V is less than ±0.5 V. The purpose of this board is to perform high speed network tests.

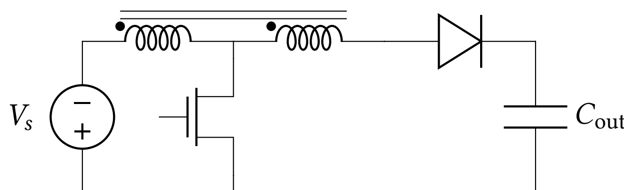


Figure 6.6: Coupled inductor boost converter

The second circuit is a 256 channel driver intended for driving 16×16 switch arrays. It uses two Microchip HV583 high voltage serial to parallel converter. These high voltage drivers feature 128 channels and a maximum output voltage of 80V. The 256 channel driver also has an on-board voltage-doubled boost converter to generate the high voltage supply. The output connector is from the Hirose FX8C series and has 280 pins. A matching connector is used on a breakout board to which is wire-bonded the silicon photonics switch. More details about these PCBs are given in Appendix A.

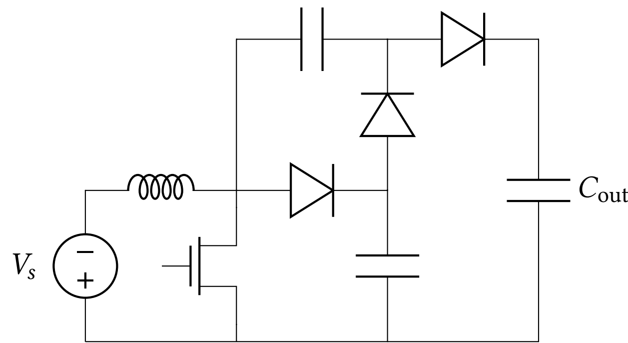
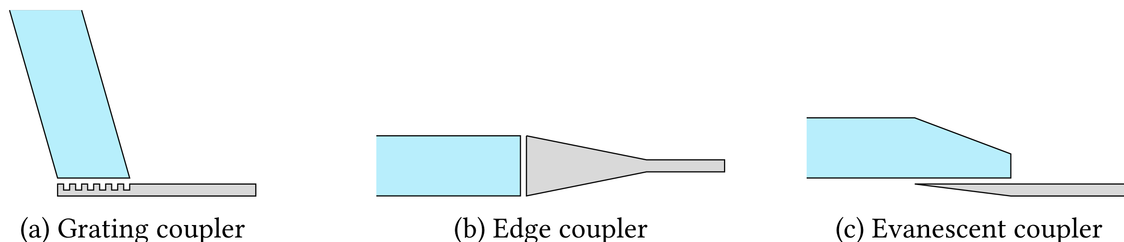


Figure 6.7: Voltage doubled boost converter

6.3 Optical packaging

Depending on the type of optical device, several packaging methods are available. By optical packaging, we mean coupling light in an out of the device, to a standard optical fiber (such as SMF-28), as efficiently as possible, with the fiber attached in a permanent manner.

In silicon photonics, there are generally three methods suitable for fiber-to-chip packaging: grating, edge and evanescent coupling (Figure 6.8).



(a) Grating coupler (b) Edge coupler (c) Evanescent coupler

Figure 6.8: Optical coupling methods

For edge or evanescent coupling, mode size converters are required. A few examples have been discussed in the previous chapters. Generally, mode size converters can be either built directly in the silicon photonics process, or fabricated as a separate device. Adding the mode size converter to the silicon photonics process has the benefit of not requiring alignment of the mode size converter itself, however it may not be compatible with every process and device. In particular, mode size converters typically require thick (more than 10 µm) cladding to prevent leakage to the silicon substrate.

In the case where the mode size converter is fabricated separately from the main device, several strategies have been developed for passive alignment (e.g. [91]) which enables very low-cost packaging. However, for performance-critical applications, active alignment might be required.

Grating coupling

Grating couplers convert guided waveguide modes into radiation modes which are almost vertical. These radiation modes can then be coupled into an optical fiber. Grating couplers have been studied extensively (see for example [90]) and are well suited for multiple applications, such as telecommunication and phase arrays. They have the advantage being accessible from the top of the wafer, therefore they can be tested during the processing. In contrast, edge couplers can only be used after the device is diced. However, grating couplers have several limitations:

- They operate over a limited bandwidth
- They typically are designed to operate at only one polarization
- The lithography requirements for the high efficiency designs can be quite stringent, especially at shorter wavelengths

Mainly because of the limited bandwidth, they are not often used for applications involving e.g. supercontinuum generation. Moreover, elaborated designs are required when both TE and TM polarizations should be coupled to the photonic chip.

Edge coupling

Edge coupling refers to butt coupling a fiber or a chip directly to a waveguide on the edge of another chip. This requires the modes on both sides to be well matched in size and shape. Compared to grating coupling, it has the following advantages:

- Broadband operation
- Potentially small polarization dependence

Even though simple in concept, the designs of edge couplers can vary greatly. Chapter 5 gave some examples of edge couplers for chip-to-chip and fiber-to-chip applications.

Evanescent coupling

Like edge coupling, evanescent coupling transfers light from a fiber in the same plane as the integrated waveguide. However, in this case we use the lateral transfer of light between two waveguide having an index crossing to transfer the light adiabatically. The main advantage of evanescent coupling compared to edge coupling is the increased alignment tolerance. There is also no need to have a flat chip facet after dicing. This will be discussed in more details in Chapter 7.

6.4 Alignment strategies

When aligning optical components, five to six degrees of freedom need to be controlled with precision. A fiber to chip alignment requires x, y, z, θ_y and θ_z to be controlled, and a fiber array to chip or chip to chip alignment requires all of x, y, z, θ_x, θ_y and θ_z to be aligned.

We use a FiconTEC FL-300 fiber attachment machine to package most of our devices (Figure 6.9). This machine features two six-axis grippers with epoxy dispensing and UV curing. It can be programmed using graphical scripting to perform the desired alignment procedures, and uses feedback from instruments such as power meters or ammeters to optimize the alignment. Each gripper has a force feedback sensor for monitoring when the samples are touching.

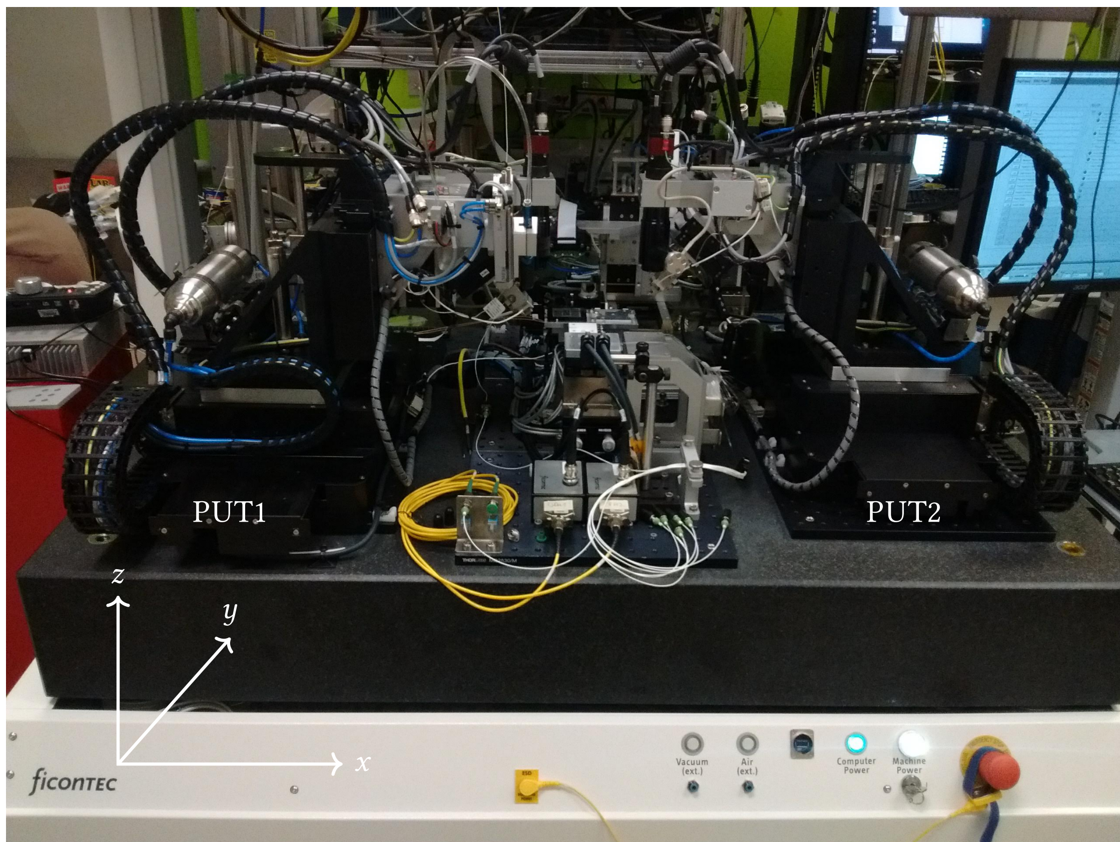


Figure 6.9: FiconTEC FL300

The general alignment procedure is divided in two steps: a coarse alignment using measurements with the cameras, and a fine alignment strategy that uses optical feedback from the device. On the FiconTEC machine, one top view camera is mounted on each arm, and another bottom view camera is fixed to the machine. The origin is defined as the focal point of the bottom view camera in the middle of the image. The top view cameras are referenced to the bottom view camera with a glass slide of a known thickness and refractive index with a metallic fiducial.

Then, the different actuators (e.g. grippers and dispensers) are referenced to the bottom view camera by moving them to the origin point. The chuck is referenced to the top view cameras by locating some key features. In all cases, the $x - y$ position is determined using shape recognition, and the z position is determined by the auto-focus function (which maximizes the image contrast).

The fixed part of the package is mounted on the chuck and located with the top view camera, and the movable part in the gripper is located using the fixed bottom view camera. The coordinate system between these two parts can be converted using the initial reference measured between the two cameras. Additionally, rotation and tilt between the fixed and movable part is measured using three (or more) corners to determine the part's plane. This step should be done iteratively since the part is not necessarily exactly at the center of rotation of the gripper.

During the final fine alignment, two feedback methods can be used:

1. Use one alignment pair from the first to last fiber
2. Have one alignment loopback at the first fiber pair, and another loopback at the last fiber pair

Method 1 requires only one feedback signal, however the loopback waveguide will most likely need to cross other components on the photonic chip. Method 2 requires more fiber channels, but makes it easier to decouple the different degrees of freedom. For example, the alignment pair closer to the center of rotation of the stage can be used for translation alignment, while the other loopback can be used for angular alignment. In both cases, the signal can come from an external laser, or can come from an on-chip laser (in the case of active devices).

Several optimization algorithms are available for alignment. A combination of the following is typically used:

- A spiral search is an efficient method to quickly search a large area
- A simplex search can optimize several coupled degrees of freedom at a time
- A 1D hill climbing strategy can be used for fine tuning the alignment

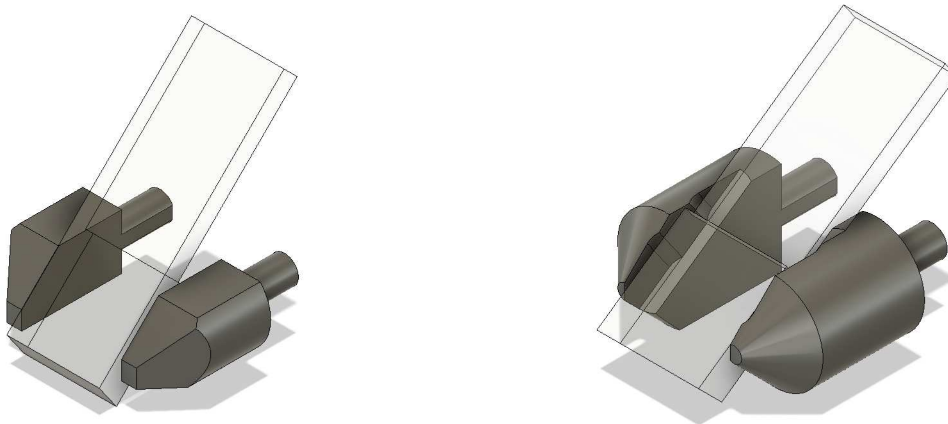
Additional details on the operation and calibration of the FiconTEC machine are given in Appendix B.

Gripper design

The FiconTEC machine was provided with an original set of grippers with a simple design shown in Figure 6.10a. These grippers are used for picking up chips from the loading tray, and manipulating chips and fiber arrays during alignment. While adequate for small chips, these grippers have some limitations:

- The maximum opening is fixed to 10 mm by the gripper actuator

- The flat faces are difficult to align perfectly parallel to each other, in which case the part might rotate easily. This is especially a problem for fiber arrays where the fibers exert a non negligible force on the array (Figure 6.11a)
- There is no good way to load a part at a preset angle (e.g. 10° for fiber arrays coupled to grating couplers).



(a) FiconTEC original gripper design

(b) Improved gripper design

Figure 6.10: Gripper designs

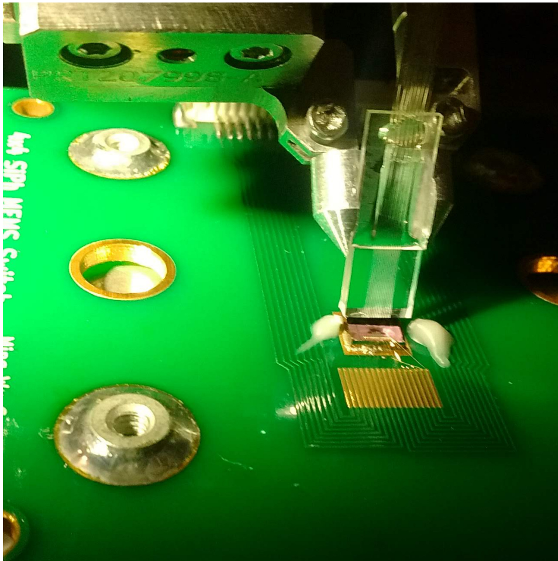
(a) Picking up part with unaligned grippers

(b) Picking up part with improved gripper design

Figure 6.11: Gripper alignment considerations

To correct these problems, we designed some custom grippers shown in Figure 6.10b. The feature a groove at a fixed angle. We built one gripper for 10° fiber arrays, and another one at 0° for evanescent coupling. The latter can be flipped around for completely vertical fiber arrays as well. The width of the groove can be adjusted so the gripper can take parts wider than 10 mm (with a compromise on the lower limit for the part width). Finally, inside the groove is a feature which mimics human hands, with two fingers on one side and an opposing thumb on the other side. This guarantees that the part is always in contact with at least 3 points, preventing any rotation. This effect is also shown in Figure 6.11b.

The 10° gripper prototype was fabricated from 6061-T6 aluminum using turning and milling. To reduce costs and fabrication complexity, the 0° gripper prototype was fabricated from 316L stainless steel using a commercial 3D printing service. The 10° gripper in use is shown in Figure 6.12a, and the 0° gripper is shown in Figure 6.12b.



(a) 10° gripper holding a fiber array during a silicon photonics switch package rework.

(b) 3D printed horizontal gripper

Figure 6.12: Fabricated gripper designs

The 3D printed gripper has a rougher surface finish and required slight polishing using fine grit sanding paper (e.g. P1500). After this preparation step, it was able to successfully grip and align a 68-channel glass interposer (Figure 6.12b).

6.5 Adhesives

After alignment, one of the critical parts of packaging optical devices is gluing everything in place. The choice of glue on its own is a vast topic, and we will give here a few guidelines. The curing of the glue also needs careful attention.

We will divide the adhesives for optical components in two categories: optically transparent, and non transparent. The opaque glues are used for mechanical stability and some come in low shrinkage version, usually by incorporating some filler material such as tiny glass beads. Adhesives such as OP-67-LS from Dymax can have a shrinkage as small as 0.08 %. On the other hand, the clear adhesive will be used in-between optical interfaces, and will often serve an index-matching purpose as well. In this case, the refractive index of the adhesive might be of importance. Adhesives can be found mostly in the refractive index range of 1.3 to 1.7, although usually this is specified at a visible wavelength.

One of the main characteristic of any adhesive is the curing method. For optical packaging, UV curing is usually preferred since it is fast (less than one minute). It is also convenient because the curing can be delayed until the part is aligned correctly. Other curing methods include chemical (e.g. mixing a two-part epoxy), contact with moisture (e.g. cyanoacrylate), or heat. Other properties of adhesives to keep in mind are the presence of adhesion promoters, which help the strength of the bond on some difficult materials, or oxygen inhibition properties, in which case the adhesive should be cured in an inert environment.

Another characteristic, related to the quantity of adhesive to be dispensed, is the viscosity of the adhesive. The viscosity is often reported in centipoise (cP), which corresponds to $1 \times 10^{-3}\,\mathrm{Pa \cdot s}$. It might range from less than 50 cP to higher than 1000 cP. In our experience, a viscosity around 200 cP, combined with a dispensing tip with 100 µm diameter, provides good control on the amount and location of dispensing for chip-to-chip and fiber-to-chip packaging.

For packaging of the 4×4 silicon MEMS switch with a grating coupler (for example in Figure 6.12a), we use the adhesives from Table 6.1

Adhesive	Viscosity (cP)	Refractive index	Curing time	Intensity
NOA146H	175	1.46	2 min	50 %
OP-67-LS	135 000	N/A	30 s	5 %

Table 6.1: Adhesive selection

The light source is a Panasonic Aicure UJ35 with a ANUJ6186 head and ANUJ6426 lens, providing approximately $7.5\,\mathrm{W/cm^2}$ of intensity at full power. The NOA146H is dispensed automatically, and the OP-67-LS is dispensed manually. The NOA146H can also be cured by heat, if additional curing is required (e.g. for better stability during transportation).

The transparent NOA146H is first cured for a long time, since the light intensity that reaches the corners under the fiber array can be low due to the higher refractive index of the fiber array compared to air. Then, the low-shrinkage structural adhesive OP-67-LS is dispensed on the side of the fiber array and cured slowly to prevent excessive stress which might cause movement during curing. Additionally it should be noted that the glue needs to be dispensed with a similar amount on each side, and cured with symmetrical light to prevent uneven curing and stress.

When packaging optical MEMS devices, care should be taken to not dispense an excessive amount of adhesive which might go over the MEMS portion of the chip. Additionally, the curing of the adhesive needs to be done in a timely manner after the dispensing as the adhesive might be slowly creeping towards the MEMS. The use of an automated alignment tool, which can save the optimal dry alignment position, is very useful in this case.

For temporary measurements, index matching gels are also available. These come in a variety of refractive indices and viscosity. They will not cure, dry or harden after time and can be cleaned easily with acetone. In our measurements, we use the IML150 matching gel from Norland, and the Series A refractive index liquids from Cargille.

Chapter 7

Evanescent coupling

Although many technologies exist to couple a single mode fiber directly to silicon photonics, in some cases it is preferable to convert the fiber mode first to an intermediate mode. For example, the performance requirements (bandwidth, polarization, etc.) might dictate it is not possible to build the mode converter using silicon photonics technology. In other cases, the process might be incompatible with large mode size converters. One such example is the fabrication of silicon photonics MEMS: the thick cladding required for the spot size converter would interfere with the oxide release step.

One solution is to use an intermediate chip with a dedicated process and edge couple this chip to the fiber on one side, and to the silicon photonics on the other side. This method has two caveats: the silicon chip also needs to use a dicing method compatible with optical facets (which might conflict with some MEMS processing steps), and the alignment between the silicon mode and the intermediate chip mode (which is much smaller than the fiber mode) might be difficult or very sensitive.

A preferable solution is to couple the intermediate chip to the silicon photonics using evanescent coupling. The basic concept is to bond the intermediate chip upside down on the silicon photonics, and to rely on inverted tapers to convert the mode adiabatically. In this case, the two chips are coupled using the evanescent fields of the inverted tapers. The alignment tolerance between the waveguides is relaxed, as long as the adiabaticity criterion is met. One way of ensuring this is to make the device long enough (e.g. several millimeters), while keeping the gap between the two chips small enough ($< 2\,\mu\text{m}$)

Evanescent coupling is particularly interesting when packaging large devices, such as wafer-scale silicon photonics MEMS switches shown in Figure 7.1. The combination of relaxed alignment tolerance and simple fabrication process is important in keeping the yield and performance of the fiber to chip coupling very high.

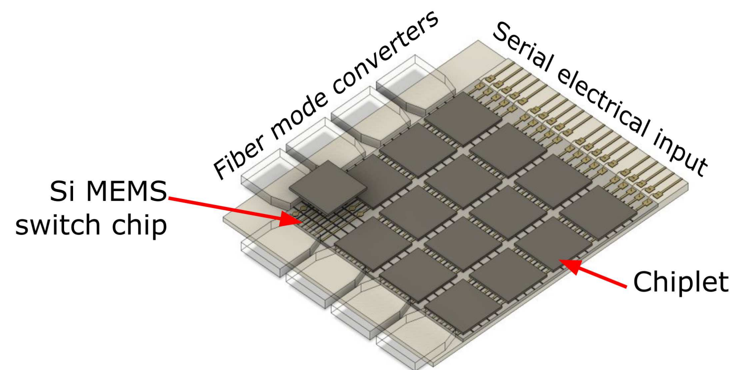


Figure 7.1: Wafer-scale silicon photonics switch concept

7.1 Literature review on evanescent couplers

IBM has been reporting for a few years an evanescent coupler design using a polymer interface chip [92, 91, 93]. The polymer substrate has low index contrast waveguides which are mode-matched to single-mode fiber on one end. The waveguide array pitch is reduced on the polymer chip, and there are additional structures used for passive mechanical alignment to the silicon chip. The polymer waveguides are then evanescently coupled to inverse taper silicon waveguides with 120 nm tip size. The design has been refined for better uniformity during production. The mounting chuck during alignment is designed to apply a non-uniform pressure, guaranteeing a minimal epoxy gap in the coupling region, and a slightly larger gap at the interface between the polymer waveguides and the beginning of the silicon chip. The typical loss reported for this structure is 1.6 dB.

Mellanox has reported on the optimization of a similar design [94]. The process variations are studied extensively, and the theoretical loss of their evanescent coupler is 0.6 dB. Like with the structure proposed by IBM, a thin (approximately 30 nm) silicon nitride passivation layer is present from the CMOS fabrication process, but the silicon tip size is 50 nm. In their study, the authors from Mellanox use the silicon nitride layer to their advantage to optimize the coupling losses between the polymer and the silicon layers. They find that the mode matching efficiency at the edge of the silicon chip is optimal with a nonzero (but not too thick) silicon nitride layer thickness above the silicon waveguide. They also find that the silicon nitride propagation loss (caused by the roughness of the thin film) plays an important role in optimizing the silicon taper length, and can significantly degrade the coupler insertion loss if the silicon nitride film is not smooth enough.

More recently, STMicroelectronics has reported an evanescent coupler device using an ion-diffused waveguide glass interposer [95]. The back end of line oxide is etched to a precise depth to expose the silicon nitride waveguide layer, with a tip size of 180 nm. The silicon nitride waveguides are then coupled to the silicon layer. The typical performance of this design is 1.25 dB to 1.55 dB. When the polarization splitter-rotator is included, the total insertion loss from fiber to silicon is 2.3 dB.

7.2 Platforms for intermediate chip

In this section, we describe potential candidate platforms for the intermediate coupling chip (also called glass interposer later in the text).

Silica on silicon waveguides

Silicon on silicon (sometimes referred to as planar lightwave circuits or PLC) is an optical platform composed of SiO_2 waveguides and cladding built on a silicon wafer. The waveguide core uses doped silica to create an index contrast with the cladding. Typical index contrasts for PLCs are in the range of 1 % to 2 %. This index contrast is slightly higher than optical fibers, but close enough to still be able to able to create fiber edge couplers with good efficiency (less than 1 dB insertion loss) [96, 97, 98]. The PLC can be coupled evanescently to silicon from the top.

Ion-diffused waveguides

Another method to fabricate low index contrast waveguides is to diffuse ions (e.g. Na^+ or Ag^+) inside a glass substrate to form a graded index waveguide core [99]. A mask is defined lithographically on the substrate and ions are diffused through the surface to form a surface waveguide. The waveguide can then be burried inside the substrate by applying an electric field which migrates the ions inside the glass. The properties of the waveguide (dimension, index contrast) can be changed with the mask width, as well as with the diffusion and driving process conditions.

On advantage of the ion-diffused waveguide platform is that the glass substrate is entirely transparent. This simplifies the process of curing epoxy using UV light when packaging to another chip. Because of this, we choose to use a glass interposer with ion-diffused waveguides for our evanescent coupler design.

7.3 Evanescent coupler design in silicon

Because of the high refractive index of silicon, in order to have a mode crossing with a low effective index waveguide, at least one of the following conditions should be met:

1. The waveguide tip should be narrow
2. The film thickness should be small

This is illustrated in Figure 7.2. We plot the contour map where the efficiency due to the tip interface is 95%. Because of birefringence, the TE and TM polarizations have different behaviors. For narrow waveguides (e.g. 50 nm), the TE effective index is lower than TM and it will couple more efficiently. In the case of wide waveguides (e.g. 200 nm) where the thickness is varied, the TE mode required a much thinner film to couple efficiently.

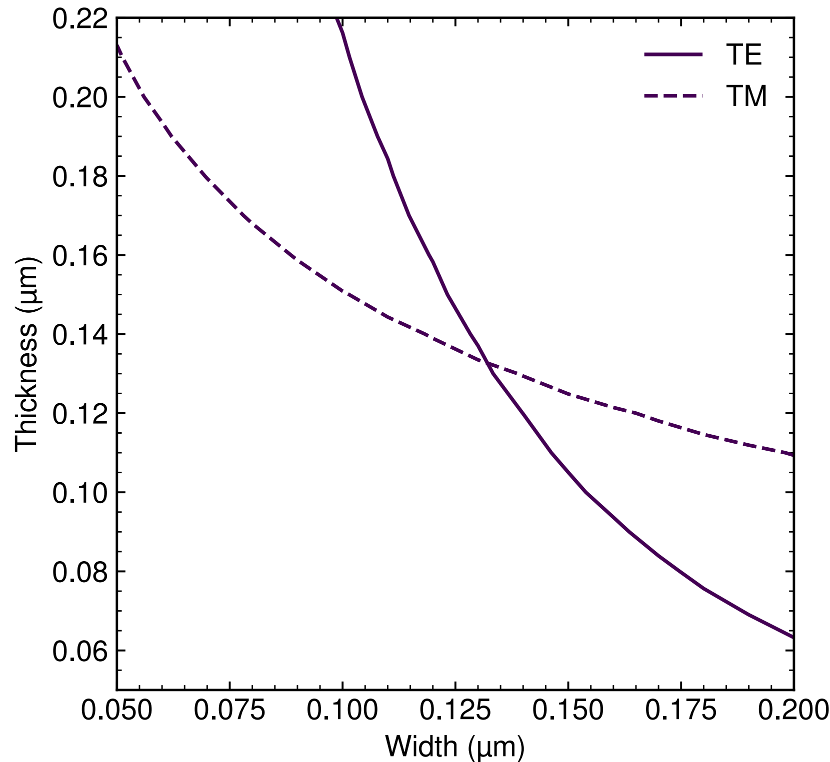


Figure 7.2: Maximum silicon thickness as a function of taper width, for 95 % efficiency

Unfortunately, the waveguide width is limited by lithography resolution. Moreover, the waveguide thickness is often determined by other parameters (substrate availability, dispersion considerations, etc.). Partial etch steps of the silicon are often included in silicon photonics processes. This can be used to create thin inverted tapers which will couple well to the evanescent couplers mentioned previously. Unfortunately, thin waveguides have large aspect ratios and show high birefringence. Moreover, the large step height between the partial etch and the full thickness waveguide can introduce unwanted reflections. To work around these two constraints, we design the silicon evanescent coupler using the following techniques:

1. The tapers shapes are optimized, to minimize the length even if the TE and TM polarizations are coupled at different locations along the taper

2. The taper is designed in 3 steps, with an intermediate partial etch to gradually increase the thickness.

3. The TM polarization is converted into the TE1 mode, which has naturally a higher effective index inside thin waveguides

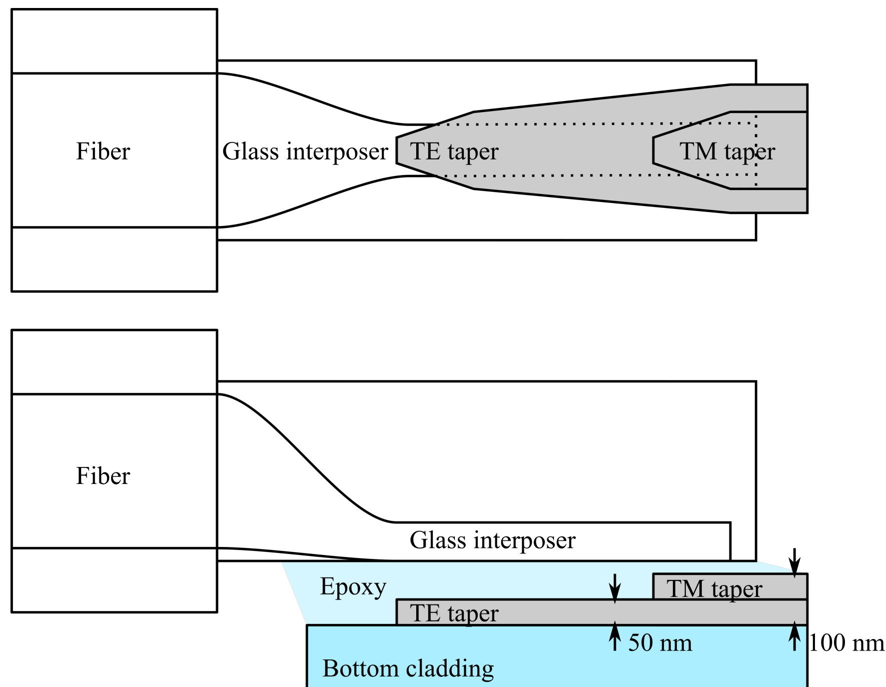


Figure 7.3: 3 step silicon taper design

The 3 step silicon taper is shown in Figure 7.3. The main issue with this design is that the tolerance of the TE and TM tapers are coupled. Indeed, the width of the TE taper at the tip of the TM taper should be narrow enough to not start coupling the TM mode. However, this width is not optimal for minimizing the perturbation of the TE mode at the height transition.

7.4 Evanescent coupler design in silicon nitride

The design of an evanescent coupler in silicon nitride (Figure 7.4) is more straightforward since the refractive index of silicon nitride is smaller than silicon. This lets us use a thicker film without requiring very fine lithography, which greatly reduces the birefringence of the waveguide. In this case we can choose a waveguide dimension where the TE and TM mode will be coupled approximately at the same location inside the coupler. For example, with a 300 nm thick silicon nitride layer, the mode crossing for both the TE and TM modes occurs around 300 nm width. Because of this, the taper can be made very short as there is only one section of the coupler which needs to be tapered slowly.

The silicon nitride coupler has two caveats when used in a silicon photonics platform:

- It requires an additional deposition step and etching in the process. Fortunately many foundries offer a silicon nitride layer in the backend process. However, this might conflict with other processes, e.g. some MEMS process where the oxide release etchant might not be compatible with silicon nitride.

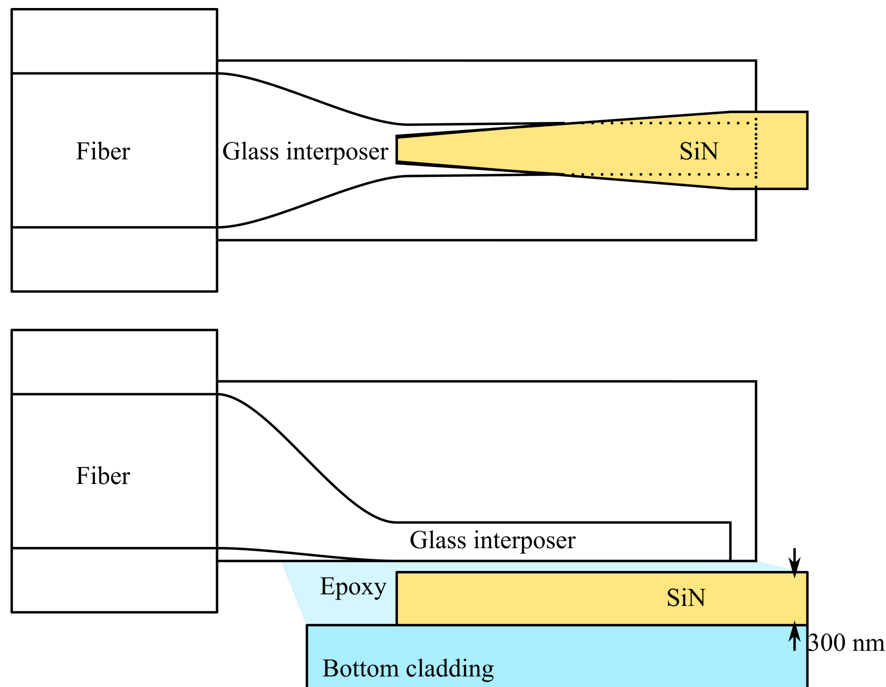


Figure 7.4: Silicon nitride taper design

- The coupling of light from the silicon nitride to silicon is not necessarily trivial. In a foundry process where fine lithography can be achieved (<100 nm), this might not be a problem. However, with typical university cleanroom DUV resolution (250 nm), a thinner silicon layer is needed (100 nm thick) to have a mode crossing with the silicon nitride layer, which adds an additional mask.

This design differs from [95] in that, because of the MEMS structures, there is naturally no oxide cladding on the sides and the top of the silicon nitride waveguide. A back end of line oxide etch is not required in our process. However, since there will be epoxy instead of silicon dioxide on the sides of the waveguide, the effective index of the waveguide is slightly higher. Because of this, our design requires a thinner silicon nitride (300 nm instead of 600 nm), which causes difficulties in coupling the silicon nitride to the silicon layer.

7.5 Splitting taper in silicon

The 3-step silicon taper suffers from poor tolerance to thickness variation. Indeed, there are two thickness steps that are carefully chosen to minimize the total loss from reflection. Moreover, the design depends on the fact that that the TE mode is coupled in the first section and the TM mode in the second section. However, these two sections cannot be optimized independently.

To circumvent this issue, we propose an alternative design in silicon in which the TE and TM coupling sections are separate (and thus the modes are split at the output). This is shown

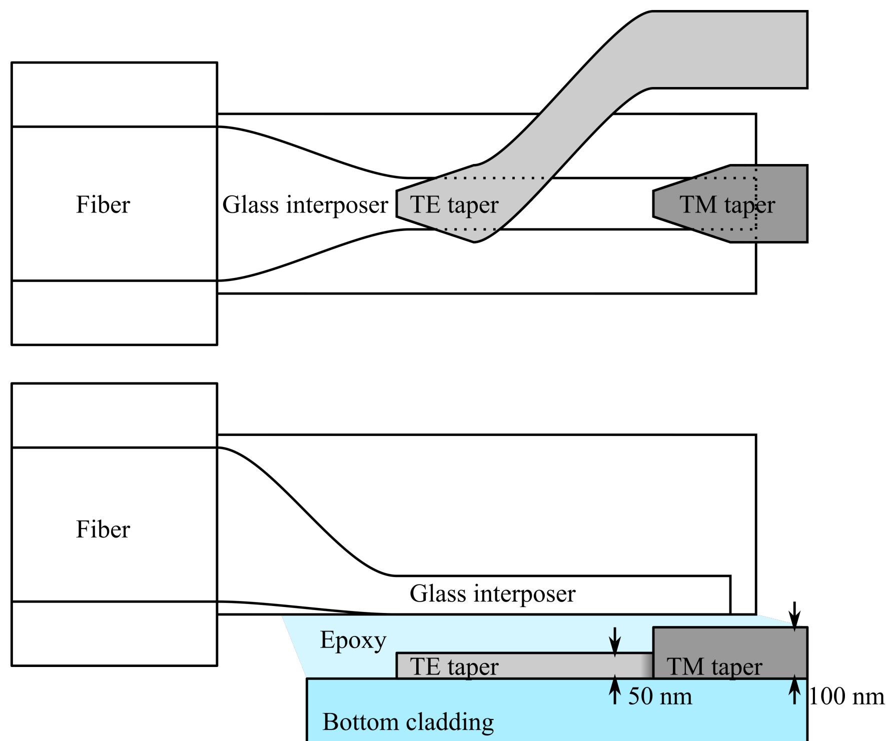


Figure 7.5: Silicon separate TE/TM taper design

in Figure 7.5. This way, we can choose the thickness of each coupler for maximum tolerance to fabrication, without being limited by reflections at height transitions.

Since the TM section of the coupler is thin and the refractive index is asymmetric in the vertical direction, it is expected there will be conversion to the TE1 mode. This feature is desirable in many cases where the silicon photonics circuit prefers to operate in a single polarization (most often TE). Combined with the splitting feature of this coupler, no extra polarization splitter-rotator (PSR) is necessary. Only a mode converter from TE1 to TE0 and a converter to the final silicon thickness are required.

This evanescent coupler has several design parameters. We want to maximize the fabrication tolerance, especially with respect to the etching depth. Based on the minimum feature size and an acceptable loss at the taper tip, we can use Figure 7.6 to determine the maximum thickness allowed for the TE taper. The overlap, for a taper width of 200 nm, reaches 95 % (or 5 % loss) at a thickness of 60 nm. This will be the upper limit of the TE taper thickness.

At this upper limit, we can determine what is the maximum width of the taper based on Figure 7.7a. Indeed, we do not want the TM mode to couple in the TE taper even at the maximum TE taper thickness. To prevent a mode crossing of the TM mode in the TE taper, we choose a taper output width of 650 nm.

Finally, we find the minimum thickness of the TE taper by ensuring there is a mode crossing for the TE polarization before the end of the taper. Based on Figure 7.7b, The minimum thickness

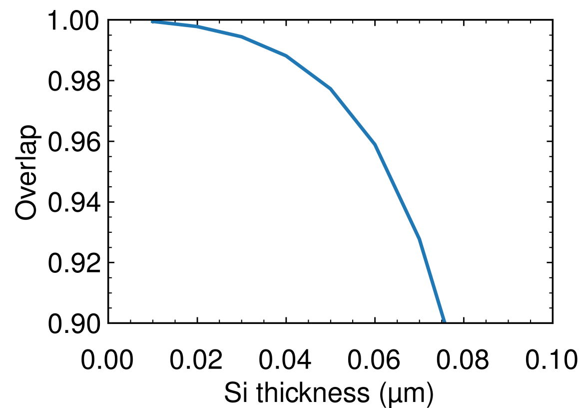
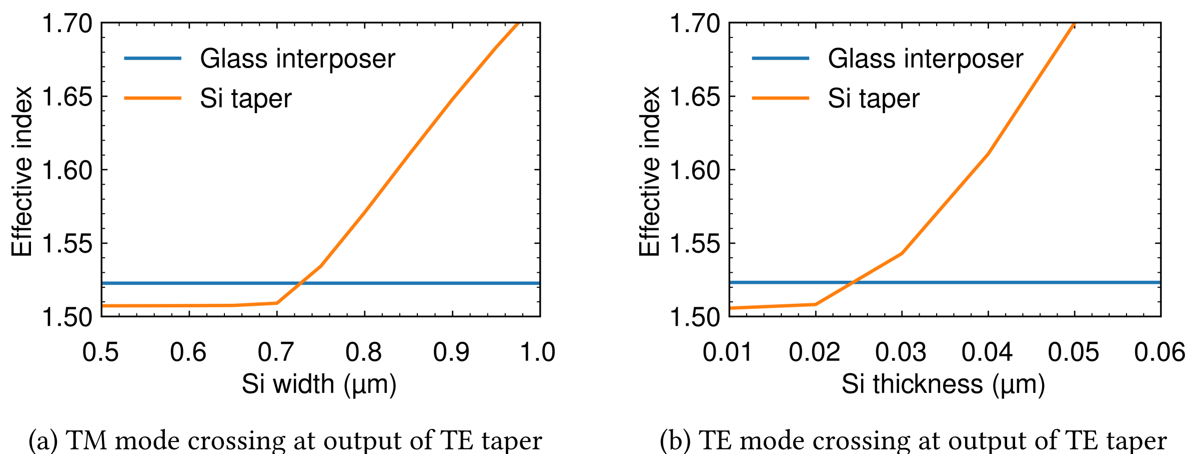


Figure 7.6: TE tip overlap loss for a taper minimum width of 200 nm



(a) TM mode crossing at output of TE taper

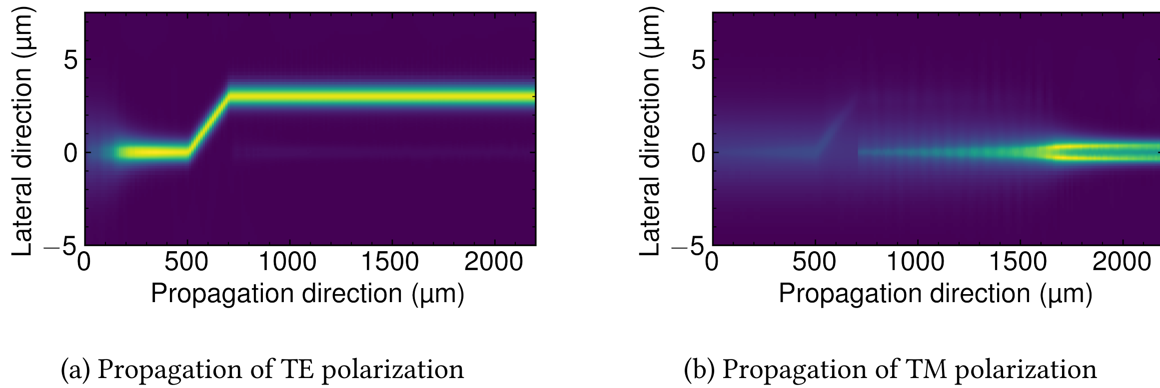
(b) TE mode crossing at output of TE taper

Figure 7.7: Design process of splitting taper evanescent coupler. The TM mode crossing is used to find the maximum TE taper width given the maximum thickness, and the TE mode crossing is used to find the minimum TE taper thickness giving the maximum width.

to have a TE mode index crossing is 30 nm.

For the TM taper, the maximum thickness is found in a similar way (overlap loss limit for given minimum feature). We choose the minimum thickness to match the TE taper fabrication tolerance. Finally, the maximum width of the TM taper is chosen to guarantee a mode crossing even at the minimum taper thickness.

We choose a nominal thickness of 50 nm, which should give a tolerance of at least ±10 nm for 95 % efficiency. With a similar design process, we choose a nominal thickness of 100 nm and a taper output width of 800 nm for the TM taper. Using the eigenmode expansion method we can simulate the propagation along the taper. Figure 7.8 shows the propagation of both the TE and TM input polarizations. The splitting function of the taper is clearly illustrated. Moreover,



(a) Propagation of TE polarization

(b) Propagation of TM polarization

Figure 7.8: E-field propagation along the splitting taper

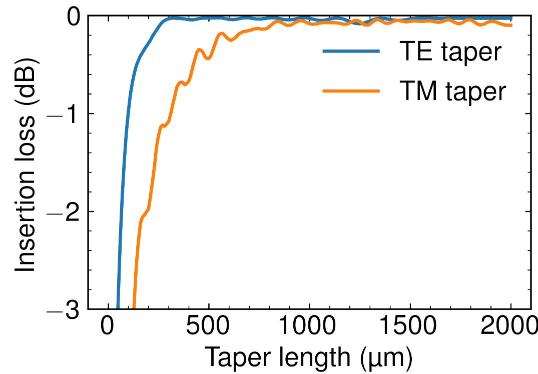


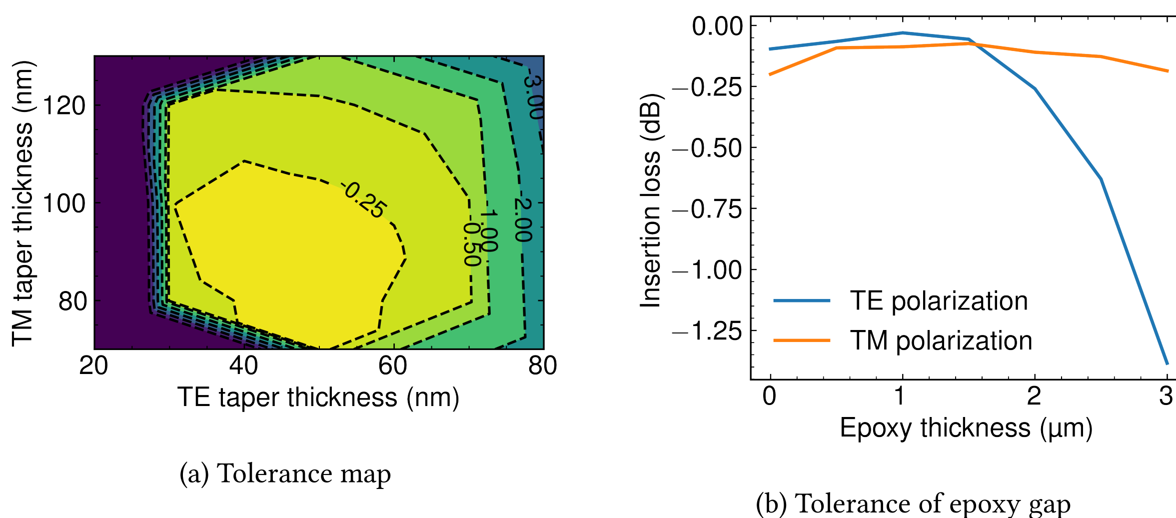
Figure 7.9: Taper length sweep for splitting taper design using eigenmode expansion

because the TM taper has a high aspect ratio at the output, there is rotation from the TM0 mode to the TE1 mode, also shown in Figure 7.8b.

The eigenmode expansion simulation is useful to determine the length of each taper. In Figure 7.9, the insertion loss of each polarization is illustrated as a function of the length of its respective taper. To increase the tolerance to the gap between the silicon waveguide and the glass interposer surface waveguide, we choose final lengths of 450 nm and 1500 nm for the TE and TM tapers respectively. The S-bend for separating the TE taper is 50 µm long.

We plot a tolerance map for the worst-case insertion loss as a function of the TE and TM taper thicknesses in Figure 7.10a. This considers no gap between the silicon and glass interposer. The −0.25 dB thickness tolerance is ±10 nm, and the −0.5 dB thickness tolerance is ±20 nm. These tolerances are readily achievable with silicon dry etching processes. Finally, an epoxy thickness of less than 2 µm will give an excess insertion loss of less than 0.5 dB (Figure 7.10b).

This evanescent coupler design was fabricated in the Marvell Nanofabrication Laboratory, and coupled to a glass interposer from TEEM Photonics. The insertion loss was measured using



(a) Tolerance map

(b) Tolerance of epoxy gap

Figure 7.10: Splitting taper evanescent coupler fabrication tolerance

a 1310 nm laser and test structures were included on the chip to measure the loss of individual components of the evanescent coupler. The breakdown of the individual loss components is listed in Table 7.1. The total insertion loss for the TE mode is 5.3 dB, and for the TM mode is 3.0 dB. The major sources of loss are the propagation loss (caused by the very thin waveguides), and the multiple height transitions.

Component	TE loss (dB)	TM loss (dB)
Glass interposer	0.9	0.9
Propagation	1.5	1.0
Transitions/mode conversion	0.6	0.6
Tip mode overlap	0.2	0.2
Gap	2.1	0.3
Total	5.3	3.0

Table 7.1: Breakdown of loss of individual components of separate taper evanescent coupler

7.6 Oxidized taper evanescent coupler

Another method to reduce the effective index of a silicon waveguide is by oxidation. Silicon thermal oxidation is a well characterized process that consumes 46 % of silicon for each unit of oxide grown. Generally this oxidation is isotropic, but it can be masked using e.g. silicon nitride. Silicon oxidation is routinely used in CMOS fabrication, e.g. LOCOS for transistor insulation or

gate oxide formation. In silicon photonics, oxidation has been used in several process, such as high quality etchless waveguides [100, 101], vertical tapering [102], taper width reduction [103], defining waveguide on bulk substrates [104], rib to strip waveguide transitions [105], and sidewall roughness reduction [106].

In this section we develop a process to shrink mainly laterally the silicon waveguide dimensions using thermal oxidation. This would have the advantage of not increasing the propagation loss as with the separate taper evanescent coupler. The process flow is illustrated in Figure 7.11. An example is given for both an oxidized full etch waveguide, and a partial etch waveguide.

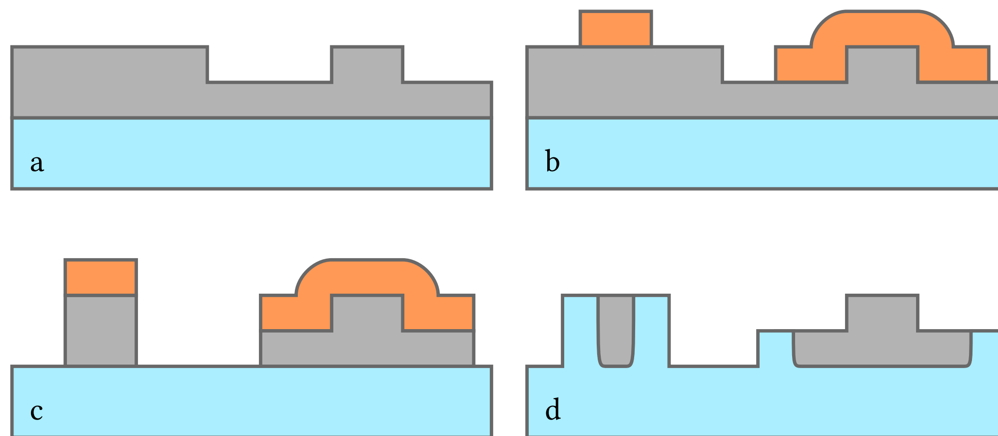


Figure 7.11: Silicon taper oxidation process (a) Partial etch. (b) Silicon nitride mask. (c) Full etch. (d) Oxidation and nitride removal.

The initial wafer is SOI with a 220 nm thick device layer and a 3 µm or greater bottom oxide (BOX) layer. We start by defining the device rib waveguides with a partial etch step. The partial etch is 110 nm deep. Then, silicon nitride is deposited to mask the oxidation process. The silicon nitride mask is patterned with an overetch step to reduce stringers in the regions where there is topography from the partial silicon etch. Because of the selectivity between silicon nitride and photoresist during dry etching is low, it might not be possible to completely remove the nitride stringers. However, the small remaining stringers do not significantly block the oxidation and are removed at the same time as the silicon nitride mask.

In the partial etch waveguide region, the silicon nitride will block the oxidation. In the evanescent coupler region, the mask is also used to pattern the full etch waveguide. The photoresist is removed before the silicon full etch to use only the nitride as a masking layer, which improves the edge roughness. After the full 220 nm etch, the wafer is ready for oxidation. We use dry oxidation at 1050 °C for 4 hours. After the oxidation, the silicon nitride mask is removed. The surface of the silicon nitride is partially oxidized and requires a quick (20 s) dilute (50:1) HF dip, before proceeding with the main wet etching [107, 108]. The etchant is 85 % phosphoric acid, heated at 160 °C with a drip system to compensate for evaporation and keep the bath concentration constant. About 20 minutes of etching are required to fully remove the silicon nitride layer.

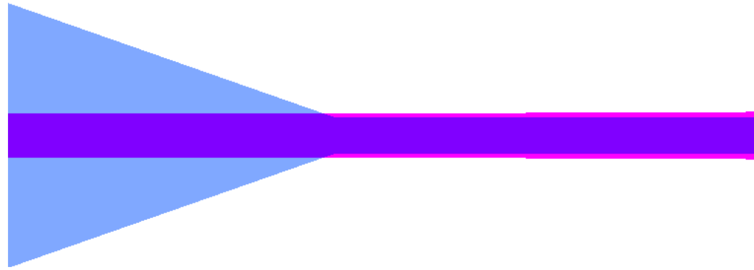
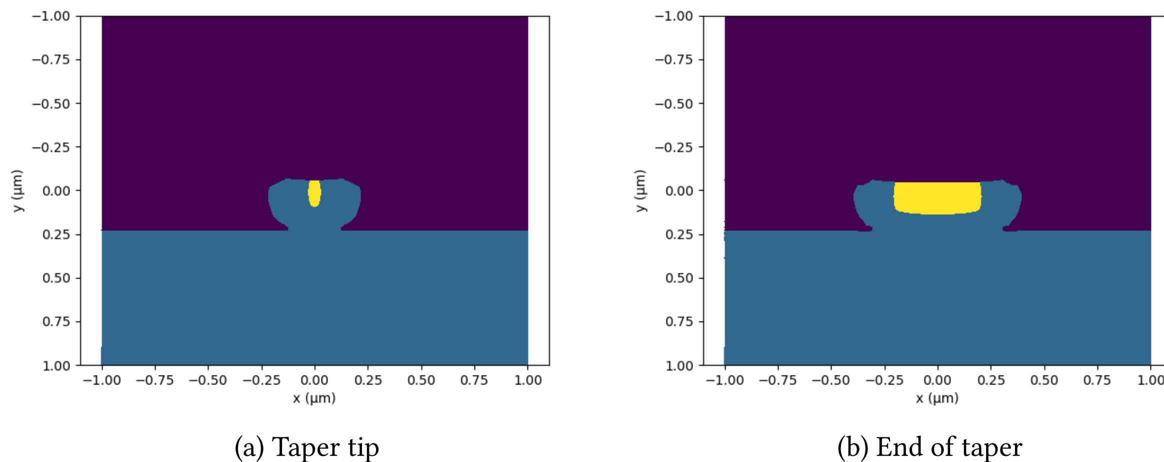


Figure 7.12: Transition between full etch (left) and partial etch (right) region in the oxidized taper process. The red layer is the full etch mask while the blue layer is the partial etch mask.

In the transition between the oxidized full etch and the masked partial etch waveguide, the widths are adjusted to account for the silicon consumed during oxidation. The layout of the transition is given in Figure 7.12. The taper angle in the partial etch mask is meant to minimize the impact of misalignment.

We simulate this process using Sentaurus Process. The silicon nitride thickness is 150 nm to prevent the nitride film from having excessive strain which would result in rounded silicon corners [109].

An example of the oxidation simulation is given in Figure 7.13. We notice that, in addition to the reduced lateral dimension, the vertical dimension of the waveguide is also reduced. Even though the top of the waveguide is protected by the silicon nitride layer, the bottom of the waveguide is oxidized by diffusion through the bottom oxide layer. This also has the effect of elevating slightly the waveguide, which should help bringing the evanescent coupler in close proximity to the glass interposer during the packaging, while increasing the volume the epoxy can fill-in to prevent overflow on the rest of the chip.



(a) Taper tip

(b) End of taper

Figure 7.13: Oxidation simulation using Sentaurus

From the simulation, we extract the dimensions of the silicon waveguide after oxidation. We

plot the trajectory of the waveguide dimensions on Figure 7.14. The parameter that is varied on this trajectory is the initial waveguide width, from 250 nm to 600 nm.

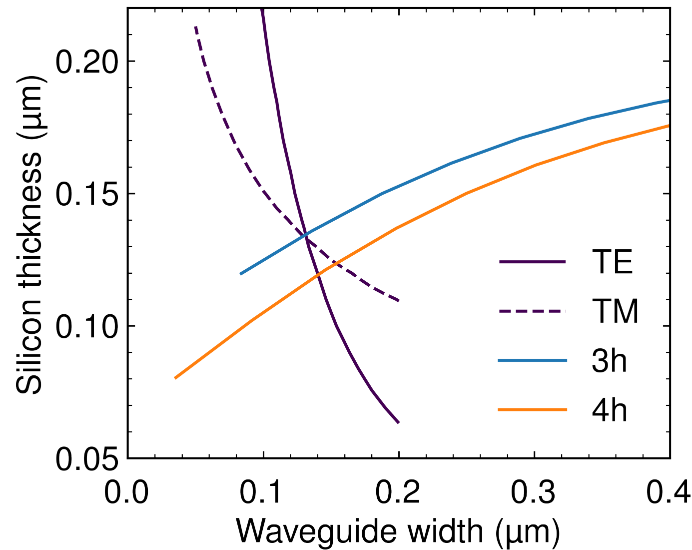


Figure 7.14: Oxidized silicon taper design space. The shape (width and thickness) after oxidation is compared to the contour lines of 95 % efficiency for both the TE and TM polarizations.

On the same figure, we repeat the contour lines for 95 % tip interface efficiency. This illustrates if the oxidation is sufficient to have low loss at the evanescent coupler tip. In this case, 3 h is enough to cross to the left region of the plot where the efficiency is greater than 95 %. However, a longer oxidation time of 4 h can be used to lower even more the tip interface loss and improve the tolerance to the initial waveguide dimension. In both cases, the oxidation has also the added benefit of reducing the birefringence of the waveguide since there is vertical oxidation in addition to the lateral oxidation.

We use a maximum drawn taper width of 600 nm (approximately 400 nm oxidized width) to prevent a TM0/TE1 mode crossing inside the evanescent coupler. The minimum drawn width is 250 nm according to the lithography resolution limit of the DUV stepper. The taper length is determined by an eigenmode expansion sweep (Figure 7.15). Since the epoxy gap might be nonzero in practice, we use a length of 2 mm to increase the robustness to particles on the chip before packaging.

We verify the maximum tolerable epoxy gap is reasonbale for the chosen taper length by varying the distance between the glass interposer core and the silicon waveguide in the eigenmode expansion simulation. This is shown in Figure 7.16, where a gap of less than 2 µm will give less than 1 dB excess loss for both the TE and TM modes. This design is more robust to excess gap than the three etch step design since the oxidation process elevates the waveguide tip closer to the glass interposer surface.

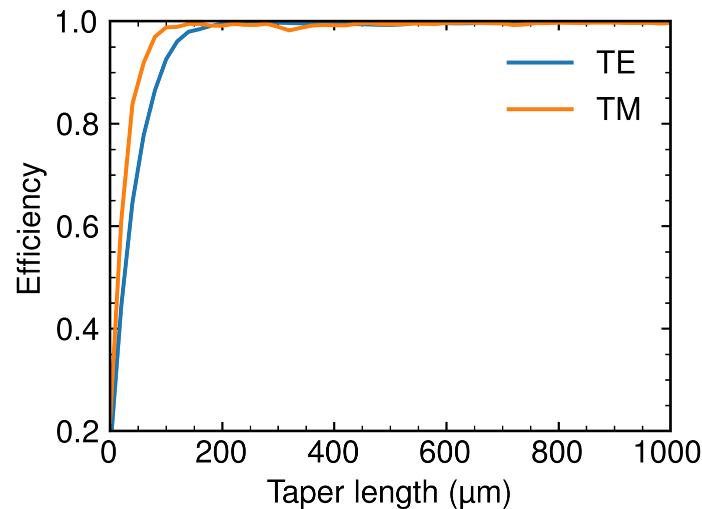


Figure 7.15: Taper length sweep of oxidized taper evanescent coupler

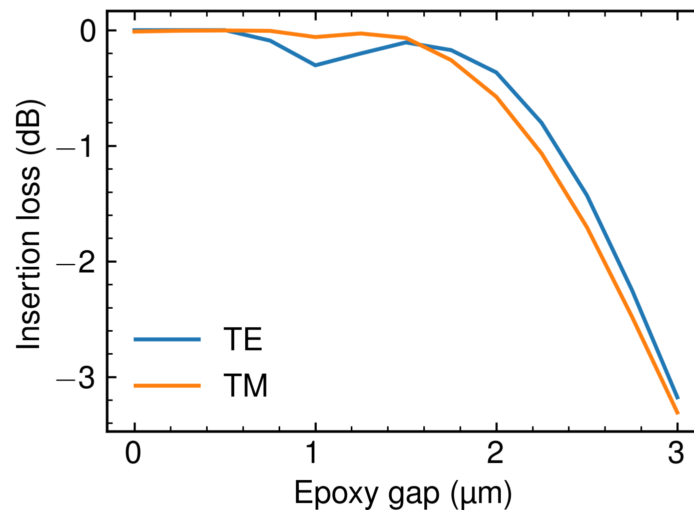


Figure 7.16: Tolerance of oxidized tip evanescent to epoxy gap thickness

The choice of refractive index of the optical adhesive is critical. Indeed, it should not be too high so that the mode in the glass interposer is not guided anymore, and leaks in the adhesive layer. On the other hand, if the adhesive index is too low, the evanescent fields will not be strong enough to couple to the adiabatic taper within a reasonable distance. In our designs with ion-diffused glass interposers, we choose a nominal adhesive index of 1.525, which is slightly higher than the glass interposer bulk index. In this case, we can usually tolerance ±0.01 refractive index

variation before the performance degrades significantly. For example, with an index over 1.535 the leakage loss is significant in the adhesive layer (>10 dB/cm, Figure 7.17), while an index below 1.515 will not couple efficiently if the epoxy thickness is more than a few hundred nanometers (e.g. because of particles).

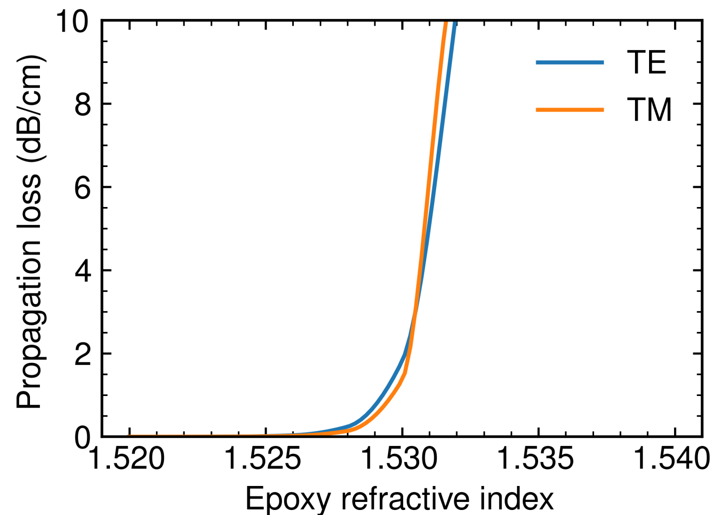


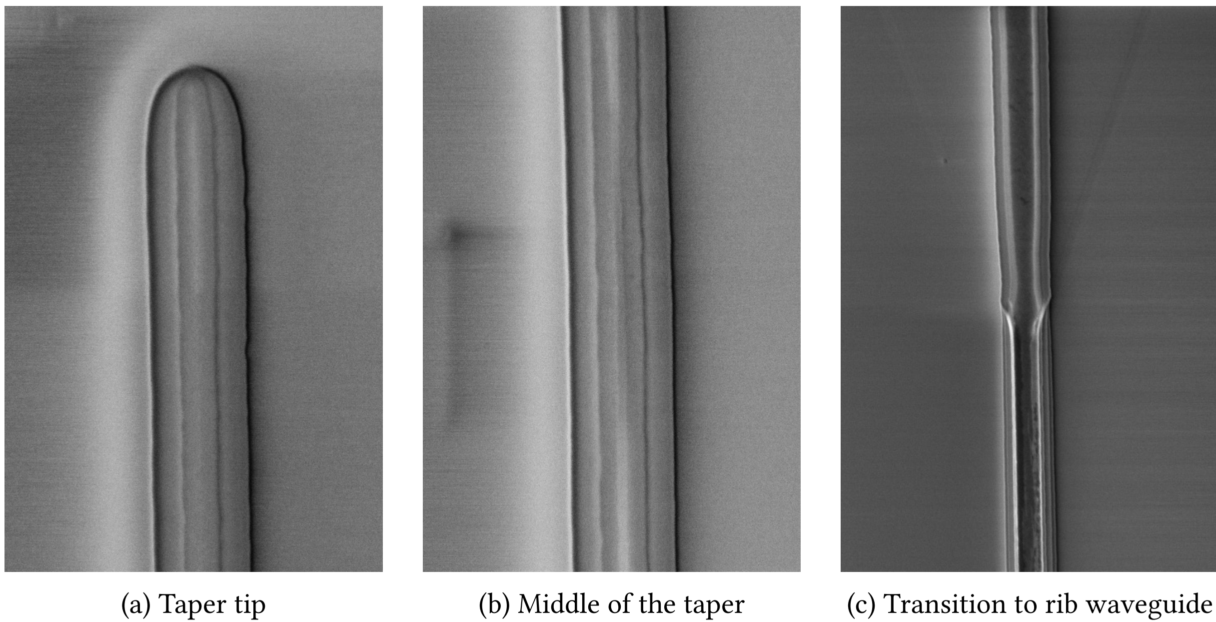
Figure 7.17: Glass interposer leakage loss as a function of epoxy index

In practice when measuring test samples, it is advantageous to use index matching liquids instead of adhesives to reuse the glass interposer or the sample. These index matching liquids can be cleaned easily with acetone, and can come in kits with fine increments. Like with adhesives, these index matching liquids might only have their refractive index specified in the visible (589.3 nm usually). When measuring in the O-band (1310 nm), the refractive index will drop by about 0.02.

Some oxidized taper test structures were fabricated in the UC Berkeley Nanolab according to the process above. The devices are shown in SEM top view in Figure 7.18 and in cross section in Figure 7.19.

On the taper top view (Figure 7.18b), we can see the narrow silicon waveguide surrounded by the grown oxide layer. Even though the silicon nitride mask layer is removed, we can clearly see its lithographically defined contour that remains. This contour is transferred to the oxide during the dilute HF step since HF also etches silicon dioxide. On the tip image (Figure 7.18a), the silicon waveguide has almost zero width. Finally, the transition region (Figure 7.18c) shows the alignment between the oxidized waveguide and the rib waveguide section.

The effect of the oxidation is well illustrated in the cross section image. In particular, we can see the bottom of the waveguide was oxidized and elevated the waveguide with respect to the initial bottom cladding layer, while the top of the waveguide was indeed masked by the silicon nitride mask. The sides of the waveguide were oxidized as expected by the simulation. We also



(a) Taper tip (b) Middle of the taper (c) Transition to rib waveguide

Figure 7.18: Fabricated oxidized waveguide

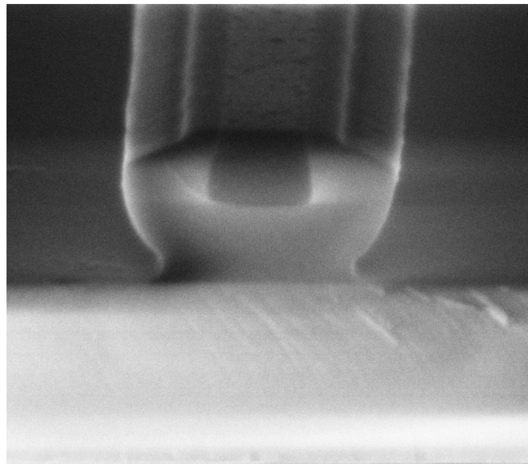


Figure 7.19: Cross section of oxidized silicon taper

notice the smoothing effect of the oxidation by comparing the roughness of the silicon region to the roughness of the contour of the silicon nitride mask.

The measured insertion loss for the oxidized tip evanescent coupler is listed in Table 7.2. The total insertion loss is 2.4 dB for the TE mode, and 1.8 dB for the TM mode. Compared to the separate taper evanescent coupler, the simpler design limits the loss in the transition sections, and the propagation loss is also greatly reduced. The propagation loss of the TE mode is larger than the TM mode because it interacts more with the sidewalls. It can also be confirmed in simulation using e.g. Figure 3.4.

Component	TE loss (dB)	TM loss (dB)
Glass interposer	0.9	0.9
Propagation	0.8	0.6
Transition/rotator	0.1	0.1
Splitter	0.2	0.2
Gap	0.4	0.0
Total	2.4	1.8

Table 7.2: Breakdown of loss of individual components of oxidized tip evanescent coupler

The insertion loss as a function of wavelength (measured using a tunable laser) is shown in Figure 7.20. Over 100 nm bandwidth, the TM mode is better than 2 dB and the TE mode is better than 3 dB. The TM mode performance compares favorably to [95], since it includes the polarization splitter-rotator. The TE mode performance is mainly limited by the excess loss from the epoxy gap, and could be improved with better sample cleanliness.

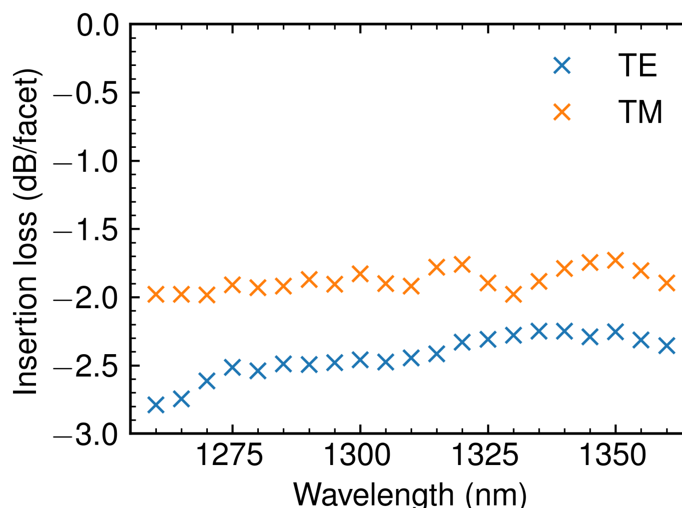


Figure 7.20: Fiber to silicon insertion loss as a function of wavelength for evanescent coupler with oxidized tip

In order to get the total insertion loss below 1 dB for very low loss wafer-scale switch applications, we detail potential improvements in Figure 7.21. With design improvements, the glass interposer loss can be reduced to 0.5 dB in total. Moreover, with process improvements regarding the edge roughness of the silicon waveguides, the propagation loss and transition losses can be reduced to 0.5 dB, for a total of 1 dB per facet.

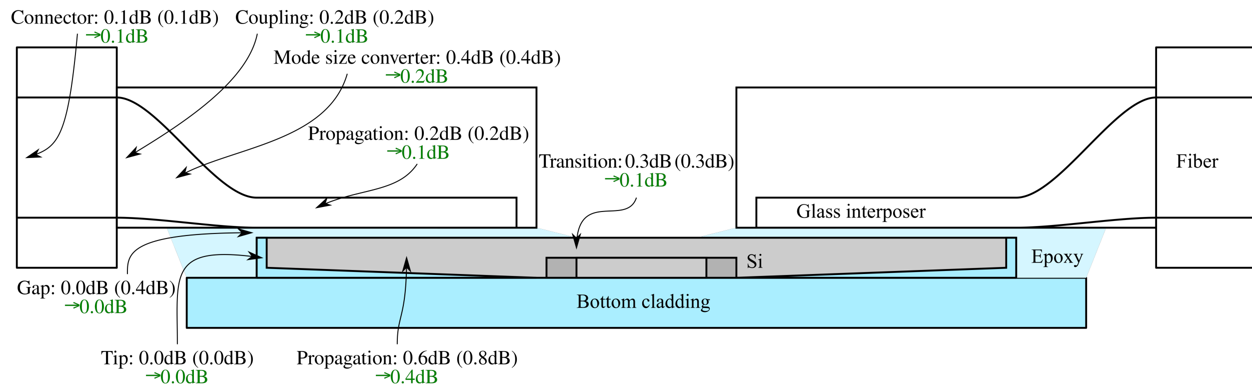


Figure 7.21: Potential improvements to evanescent coupler loss. TM mode in black, (TE mode in parenthesis), improvements in green.

7.7 Evanescent coupler alignment

In the case where one of the chips is transparent (such as glass interposers for evanescent coupling), another fine alignment method consists of using thin-film interference in the air-gap between the two parts to measure the relative tilt.

Using the air-wedge interference formula, the gap thickness at the nth fringe is

$$t = \frac{n\lambda}{2} \tag{7.1}$$

We deduce the relative tilt θ from the fringe period P using

$$\theta = \frac{\lambda}{2P} \tag{7.2}$$

This phenomenon is illustrated on Figure 7.22. The field of view is approximately 1 mm. When less than one fringe is visible in the camera, the misalignment is less than 0.02°, or less than 300 nm across the field of view.

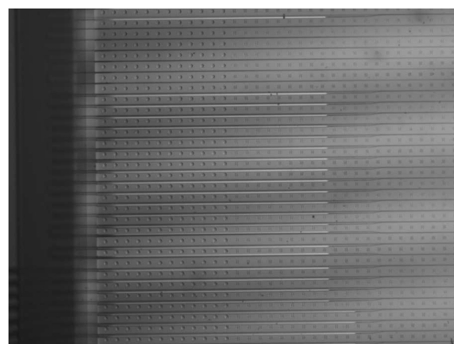


Figure 7.22: Interference fringes during evanescent coupler alignment

Chapter 8

Conclusion

In this thesis, we presented the design and fabrication of several nonlinear and low-loss integrated photonic components. These components were used for two applications: on-chip supercontinuum generation, and large-scale silicon photonics MEMS switches. Moreover, we also discussed about packaging techniques for both these applications. In this process, we designed three major components that were fabricated and measured.

First, by using the nonlinear Schrödinger equation to optimize the dispersion, we fabricated a nonlinear waveguide in chalcogenide glass. The rib waveguide with 1.2 µm width has anomalous dispersion which is suitable for supercontinuum generation. With input pulses of 240 fs and 77 W peak power, a 2 cm long waveguide can generate an octave-spanning supercontinuum (from 1 µm to 2 µm wavelength) which is simulated to be coherent. The use of a dispersive wave helped in increasing the power in the 1 µm part of the supercontinuum.

Then, we presented a broadband spectral splitter that can separate the f and $2f$ components of the supercontinuum. The spectral splitter uses the concept of mode evolution, and the fact that higher order modes have different dispersion than the fundamental mode. A simple TE0 to TE1 converter with the proper dimensions can be used to convert the blue part of the spectrum without affecting the red part. Another converter can be cascaded to convert back to the fundamental mode. The device was fabricated on a chalcogenide glass chip with a cutoff wavelength of 1300 nm. The measurement shows good agreement with the simulation. In particular, the extinction ratio is greater that 20 dB

Finally, we studied several designs of evanescent couplers between a glass interposer and a silicon waveguide. The most promising design consists of a partially oxidized silicon taper. The oxidation is used to reduce the width at the tip of the taper down to less than 50 nm, enabling efficient coupling to the glass interposer without using advanced lithography. We proposed a process for selectively oxidizing the silicon taper, and fabricated an array of evanescent couplers on a silicon on insulator wafer. The evanescent coupler has 1.8 dB insertion loss for the TM mode, and 2.4 dB for the TE mode. We also propose areas of improvements to reduce the insertion loss to 1 dB.

Bibliography

[1] R. Soref. “The Past, Present, and Future of Silicon Photonics”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 12.6 (2006), pp. 1678–1687.

[2] B. Jalali and S. Fathpour. “Silicon Photonics”. In: *Journal of Lightwave Technology* 24.12 (2006), pp. 4600–4615.

[3] R. Soref and B. Bennett. “Electrooptical effects in silicon”. In: *IEEE Journal of Quantum Electronics* 23.1 (1987), pp. 123–129.

[4] G. T. Reed et al. “Silicon optical modulators”. In: *Nature Photonics* 4.8 (2010), pp. 518–526. ISSN: 1749-4893.

[5] Laurent Vivien et al. “42 GHz p.i.n Germanium photodetector integrated in a silicon-on-insulator waveguide”. In: *Opt. Express* 17.8 (Apr. 2009), pp. 6252–6257.

[6] Alexander W. Fang et al. “Electrically pumped hybrid AlGaInAs-silicon evanescent laser”. In: *Opt. Express* 14.20 (Oct. 2006), pp. 9203–9210.

[7] Alan D. Bristow, Nir Rotenberg, and Henry M. van Driel. “Two-photon absorption and Kerr coefficients of silicon for 850–2200nm”. In: *Applied Physics Letters* 90.19 (2007), p. 191104.

[8] Vilson R. Almeida, Roberto R. Panepucci, and Michal Lipson. “Nanotaper for compact mode conversion”. In: *Opt. Lett.* 28.15 (Aug. 2003), pp. 1302–1304.

[9] Andrew Michaels and Eli Yablonovitch. “Inverse design of near unity efficiency perfectly vertical grating couplers”. In: *Opt. Express* 26.4 (Feb. 2018), pp. 4766–4779.

[10] Th. Udem, R. Holzwarth, and T. W. Hänsch. “Optical frequency metrology”. In: *Nature* 416.6877 (2002), pp. 233–237.

[11] David J. Jones et al. “Carrier-Envelope Phase Control of Femtosecond Mode-Locked Lasers and Direct Optical Frequency Synthesis”. In: *Science* 288.5466 (2000), pp. 635–639.

[12] Victor Brasch et al. “Self-referenced photonic chip soliton Kerr frequency comb”. In: *Light: Science & Applications* 6.1 (2017), e16202–e16202. ISSN: 2047-7538.

[13] David R. Carlson et al. “Self-referenced frequency combs using high-efficiency silicon-nitride waveguides”. In: *Opt. Lett.* 42.12 (June 2017), pp. 2314–2317.

[14] M. Malinowski et al. “Amplified octave-spanning supercontinuum from chalcogenide waveguides for second-harmonic generation”. In: *2017 IEEE Photonics Conference (IPC)*. 2017, pp. 261–262.

[15] J. D. Jost et al. "Counting the cycles of light using a self-referenced optical microresonator". In: *Optica* 2.8 (Aug. 2015), pp. 706–711.

[16] A. Bhardwaj et al. "A Monolithically Integrated Racetrack Colliding-Pulse Mode-Locked Laser With Pulse-Picking Modulator". In: *IEEE Journal of Quantum Electronics* 56.4 (2020), pp. 1–8.

[17] Ashutosh Rao et al. "Actively-monitored periodic-poling in thin-film lithium niobate photonic waveguides with ultrahigh nonlinear conversion efficiency of 4600 %W−1c
In: *Opt. Express* 27.18 (Sept. 2019), pp. 25920–25930.

[18] Govind Agrawal. *Nonlinear Fiber Optics (5th Edition)*. Elsevier, 2013.

[19] L. Petit et al. "Nonlinear optical properties of glasses in the system Ge/Ga-Sb-S/Se". In: *Opt. Lett.* 31.10 (May 2006), pp. 1495–1497.

[20] Ju Won Choi et al. "Nonlinear characterization of GeSbS chalcogenide glass waveguides". In: *Scientific Reports* 6.1 (Dec. 2016), p. 39234.

[21] Samuel Serna et al. "Nonlinear optical properties of integrated GeSbS chalcogenide waveguides". In: *Photon. Res.* 6.5 (May 2018), B37–B42.

[22] D. T. H. Tan et al. "Group velocity dispersion and self phase modulation in silicon nitride waveguides". In: *Applied Physics Letters* 96.6 (2010), p. 061101.

[23] M. Dinu, F. Quochi, and H. Garcia. "Third-order nonlinearities in silicon at telecom wavelengths". In: *Applied Physics Letters* 82.18 (2003), pp. 2954–2956.

[24] Neetesh Singh et al. "Midinfrared supercontinuum generation from 2 to 6 μm in a silicon nanowire". In: *Optica* 2.9 (Sept. 2015), pp. 797–802.

[25] Johan Hult. "A Fourth-Order Runge–Kutta in the Interaction Picture Method for Simulating Supercontinuum Generation in Optical Fibers". In: *J. Lightwave Technol.* 25.12 (Dec. 2007), pp. 3770–3775.

[26] John M. Dudley and Stéphane Coen. "Coherence properties of supercontinuum spectra generated in photonic crystal and tapered optical fibers". In: *Opt. Lett.* 27.13 (July 2002), pp. 1180–1182.

[27] R. Paschotta. "Noise of mode-locked lasers (Part I): numerical model". In: *Applied Physics B* 79.2 (July 2004), pp. 153–162.

[28] John M. Dudley, Goëry Genty, and Stéphane Coen. "Supercontinuum generation in photonic crystal fiber". In: *Rev. Mod. Phys.* 78 (4 Oct. 2006), pp. 1135–1184.

[29] Jean-Étienne Tremblay et al. "Picojoule-level octave-spanning supercontinuum generation in chalcogenide waveguides". In: *Opt. Express* 26.16 (Aug. 2018), pp. 21358–21363.

[30] J. Tremblay et al. "Large bandwidth silicon nitride spot-size converter for efficient supercontinuum coupling to chalcogenide waveguide". In: *2017 Conference on Lasers and Electro-Optics (CLEO)*. 2017, pp. 1–2.

[31] Lian-Wee Luo et al. “High quality factor etchless silicon photonic ring resonators”. In: *Opt. Express* 19.7 (Mar. 2011), pp. 6284–6289.

[32] Dane R. Austin et al. “Dispersive wave blue-shift in supercontinuum generation”. In: *Opt. Express* 14.25 (Dec. 2006), pp. 11997–12007.

[33] Adrea R. Johnson et al. “Octave-spanning coherent supercontinuum generation in a silicon nitride waveguide”. In: *Opt. Lett.* 40.21 (Nov. 2015), pp. 5117–5120.

[34] A. Michaels, M. C. Wu, and E. Yablonovitch. “Hierarchical Design and Optimization of Silicon Photonics”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 26.2 (2020), pp. 1–12.

[35] Chris A Mack. *Inside PROLITH: a comprehensive guide to optical lithography simulation : for the PROLITH family of lithography simulation tools, v5.0.* en. Austin, Tex.: FINLE Technologies, 1997. ISBN: 978-0-9650922-0-3.

[36] Shalin B. Mehta and Rudolf Oldenbourg. “Image simulation for biological microscopy: microlith”. In: *Biomed. Opt. Express* 5.6 (June 2014), pp. 1822–1838.

[37] Malcolm C. A. White et al. “PyKonal: A Python Package for Solving the Eikonal Equation in Spherical and Cartesian Coordinates Using the Fast Marching Method”. In: *Seismological Research Letters* 91.4 (June 2020), pp. 2378–2389.

[38] Steven J. Byrnes. *Multilayer optical calculations.* 2019. arXiv: 1603.02720.

[39] F. P. Payne and J. P. R. Lacey. “A theoretical analysis of scattering loss from planar optical waveguides”. In: *Optical and Quantum Electronics* 26.10 (Oct. 1994), pp. 977–986.

[40] F. Grillot et al. “Size influence on the propagation loss induced by sidewall roughness in ultrasmall SOI waveguides”. In: *IEEE Photonics Technology Letters* 16.7 (2004), pp. 1661–1663.

[41] S. M. Lindecrantz and O. G. Hellesø. “Estimation of Propagation Losses for Narrow Strip and Rib Waveguides”. In: *IEEE Photonics Technology Letters* 26.18 (2014), pp. 1836–1839.

[42] Daniel K. Sparacin, Steven J. Spector, and Lionel C. Kimerling. “Silicon Waveguide Sidewall Smoothing by Wet Chemical Oxidation”. In: *J. Lightwave Technol.* 23.8 (Aug. 2005), p. 2455.

[43] Matthew Borselli, Thomas J. Johnson, and Oskar Painter. “Beyond the Rayleigh scattering limit in high-Q silicon microdisks: theory and experiment”. In: *Opt. Express* 13.5 (Mar. 2005), pp. 1515–1530.

[44] S. T. S. Cheung et al. “Low-loss and High Contrast Silicon-on-Insulator (SOI) Arrayed Waveguide Grating”. In: *Conference on Lasers and Electro-Optics 2012.* Optical Society of America, 2012, CM4A.5.

[45] Michael R.E. Lamont et al. “Supercontinuum generation in dispersion engineered highly nonlinear ($\gamma = 10$ /W/m) As2S3 chalcogenide planar waveguide”. In: *Opt. Express* 16.19 (Sept. 2008), pp. 14938–14944.

[46] M. R. Karim et al. “Design of dispersion-engineered As2Se3 channel waveguide for mid-infrared region supercontinuum generation”. In: *Journal of Applied Physics* 123.21 (2018), p. 213101.

[47] M. R. Karim, B. M. A. Rahman, and Govind P. Agrawal. “Mid-infrared supercontinuum generation using dispersion-engineered Ge11.5As24Se64.5 chalcogenide channel waveguide”. In: *Opt. Express* 23.5 (Mar. 2015), pp. 6903–6914.

[48] Guang Yang et al. “A photo-stable chalcogenide glass”. In: *Opt. Express* 16.14 (July 2008), pp. 10565–10571.

[49] Jeff Chiles et al. “Low-loss, submicron chalcogenide integrated photonics with chlorine plasma etching”. In: *Applied Physics Letters* 106.11 (2015), p. 111110.

[50] Martin H. P. Pfeiffer et al. “Photonic Damascene process for integrated high-Q microresonator based nonlinear photonics”. In: *Optica* 3.1 (Jan. 2016), pp. 20–25.

[51] Kevin Luke et al. “Overcoming Si3N4 film stress limitations for high quality factor ring resonators”. In: *Opt. Express* 21.19 (Sept. 2013), pp. 22829–22833.

[52] Kevin Luke et al. “Broadband mid-infrared frequency comb generation in a Si3N4 microresonator”. In: *Opt. Lett.* 40.21 (Nov. 2015), pp. 4823–4826.

[53] H. El Dirani et al. “Crack-Free Silicon-Nitride-on-Insulator Nonlinear Circuits for Continuum Generation in the C -Band”. In: *IEEE Photonics Technology Letters* 30.4 (2018), pp. 355–358.

[54] C. H. Henry et al. “Low loss Si3N4–SiO2 optical waveguides on Si”. In: *Appl. Opt.* 26.13 (July 1987), pp. 2621–2624.

[55] Hugo Dupont. “Low loss Silicon Nitride Waveguides for Photonic Integrated Circuits”. PhD thesis. Mar. 2019.

[56] Jianfeng Luo and David A. Dornfeld. “Wafer-Scale Modeling of CMP”. In: *Integrated Modeling of Chemical Mechanical Planarization for Sub-Micron IC Fabrication: From Particle Scale to Feature, Die and Wafer Scales*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 255–284.

[57] Chen Sun et al. “Single-chip microprocessor that communicates directly using light”. In: *Nature* 528.7583 (2015), pp. 534–538.

[58] Jie Sun et al. “Large-scale nanophotonic phased array”. In: *Nature* 493.7431 (2013), pp. 195–199.

[59] M. Haverlag, G. S. Oehrlein, and D. Vender. “Sidewall passivation during the etching of poly-Si in an electron cyclotron resonance plasma of HBr”. In: *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena* 12.1 (1994), pp. 96–101.

[60] Mutumi Tuda, Kenji Shintani, and Junji Tanimura. “Highly Selective Removal of Residual Deposited Films and Oxide Hard Masks on Polysilicon Gate Electrodes in Anhydrous HF Gases”. In: *Japanese Journal of Applied Physics* 43.3 (Mar. 2004), pp. 945–951.

[61] Chang Weon Lee et al. "The cleaning effect analysis on post treatment of polysilicon etching". In: *Journal of the Korean Institute of Chemical Engineers* 38 (2 Apr. 2000).

[62] Wai Mun Lee. "Cleaning compositions for removing etching residue with hydroxylamine, alkanolamine, and chelating agent". US5672577A. Sept. 1997.

[63] Steven G. Johnson et al. "Adiabatic theorem and continuous coupled-mode theory for efficient taper transitions in photonic crystals". In: *Phys. Rev. E* 66 (6 Dec. 2002), p. 066608.

[64] P. Bienstman and R. Baets. "Optical modelling of photonic crystals and VCSELs using eigenmode expansion and perfectly matched layers". In: *Optical and Quantum Electronics* 33.4 (2001), pp. 327–341.

[65] A. Milton and W. Burns. "Mode coupling in optical waveguide horns". en. In: *IEEE Journal of Quantum Electronics* 13.10 (Oct. 1977), pp. 828–835.

[66] Yunfei Fu et al. "Efficient adiabatic silicon-on-insulator waveguide taper". In: *Photon. Res.* 2.3 (June 2014), A41–A44.

[67] Curt Wittig. "The Landau-Zener Formula". In: *The Journal of Physical Chemistry B* 109.17 (2005), pp. 8428–8430.

[68] M. Heiblum and J. Harris. "Analysis of curved optical waveguides by conformal transformation". In: *IEEE Journal of Quantum Electronics* 11.2 (1975), pp. 75–83.

[69] Michael Krause. "Finite-Difference Mode Solver for Curved Waveguides With Angled and Curved Dielectric Interfaces". In: *J. Lightwave Technol.* 29.5 (Mar. 2011), pp. 691–699.

[70] D. M. Shyroki. "Exact Equivalent Straight Waveguide Model for Bent and Twisted Waveguides". In: *IEEE Transactions on Microwave Theory and Techniques* 56.2 (2008), pp. 414–419.

[71] T. Kitoh et al. "Bending loss reduction in silica-based waveguides by using lateral offsets". In: *Journal of Lightwave Technology* 13.4 (1995), pp. 555–562.

[72] Florian Vogelbacher et al. "Analysis of silicon nitride partial Euler waveguide bends". In: *Opt. Express* 27.22 (Oct. 2019), pp. 31394–31406.

[73] David A. Anisi. ""Optimal Motion Control of a Ground Vehicle"". MA thesis. Royal Institute of Technology (KTH), Stockholm, Sweden, 2003.

[74] Wesley D. Sacher et al. "Polarization rotator-splitters in standard active silicon photonics platforms". In: *Opt. Express* 22.4 (Feb. 2014), pp. 3777–3786.

[75] You-Chia Chang et al. "Resonance–Free Light Recycling in Waveguides". In: *Conference on Lasers and Electro-Optics.* Optical Society of America, 2017, SF1J.5.

[76] J. Tremblay et al. "Large Bandwidth Waveguide Spectral Splitters Using Higher-Order Mode Evolution". In: *2018 Conference on Lasers and Electro-Optics (CLEO).* 2018, pp. 1–2.

[77] Y. Shani et al. "Integrated optic adiabatic devices on silicon". In: *IEEE Journal of Quantum Electronics* 27.3 (Mar. 1991), pp. 556–566.

[78] Emir Salih Magden et al. "Transmissive silicon photonic dichroic filters with spectrally selective waveguides". In: *Nature Communications* 9.1 (2018), p. 3009.

[79] Eric J. Stanton et al. "Multi-octave spectral beam combiner on ultra-broadband photonic integrated circuit platform". In: *Opt. Express* 23.9 (May 2015), pp. 11272–11283.

[80] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272.

[81] J. O. Ramsay. "Monotone Regression Splines in Action". In: *Statist. Sci.* 3.4 (Nov. 1988), pp. 425–441.

[82] Martin Papes et al. "Fiber-chip edge coupler with large mode size for silicon photonic wire waveguides". In: *Opt. Express* 24.5 (Mar. 2016), pp. 5026–5038.

[83] M.-J. Picard et al. "Novel spot-size converter for optical fiber to sub-μm silicon waveguide coupling with low loss, low wavelength dependence and high tolerance to alignment". In: *2015 European Conference on Optical Communication (ECOC)*. 2015, pp. 1–3.

[84] Marie-Josée Picard et al. "CMOS-compatible spot-size converter for optical fiber to sub-um silicon waveguide coupling with low-loss low-wavelength dependence and high tolerance to misalignment". In: *Silicon Photonics XI*. Ed. by Graham T. Reed and Andrew P. Knights. Vol. 9752. International Society for Optics and Photonics. SPIE, 2016, pp. 132–138.

[85] Jonathan Hu and Curtis R. Menyuk. "Understanding leaky modes: slab waveguide revisited". In: *Adv. Opt. Photon.* 1.1 (Jan. 2009), pp. 58–106.

[86] Tae Joon Seok et al. "Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers". In: *Optica* 3.1 (Jan. 2016), pp. 64–70.

[87] N. Dupuis et al. "An 8×8 Silicon Photonic Switch Module with Nanosecond-Scale Reconfigurability". In: *2020 Optical Fiber Communications Conference and Exhibition (OFC)*. 2020, pp. 1–3.

[88] Keijiro Suzuki et al. "Low-Insertion-Loss and Power-Efficient 32 × 32 Silicon Photonics Switch With Extremely High-Δ Silica PLC Connector". In: *J. Lightwave Technol.* 37.1 (Jan. 2019), pp. 116–122.

[89] Tae Joon Seok et al. "Wafer-scale silicon photonic switches beyond die size limit". In: *Optica* 6.4 (Apr. 2019), pp. 490–494.

[90] Andrew Michaels. "A Hierarchical Approach to the Design and Optimization of Photonics". PhD thesis. EECS Department, University of California, Berkeley, Dec. 2019. URL: `http://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-163.html`.

[91] Tymon Barwicz et al. "A Novel Approach to Photonic Packaging Leveraging Existing High-Throughput Microelectronic Facilities". In: *IEEE Journal of Selected Topics in Quantum Electronics* 22.6 (Nov. 2016), pp. 455–466.

[92] Tymon Barwicz and Yoichi Taira. "Low-Cost Interfacing of Fibers to Nanophotonic Waveguides: Design for Fabrication and Assembly Tolerances". In: *IEEE Photonics Journal* 6.4 (Aug. 2014), pp. 1–18.

[93] Tymon Barwicz et al. “Advances in Interfacing Optical Fibers to Nanophotonic Waveguides Via Mechanically Compliant Polymer Waveguides”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 26.2 (Mar. 2020), pp. 1–12.

[94] Ioannis Poulopoulos et al. “SiN-assisted Polymer-to-SiPh adiabatic coupler optimization”. In: *Optical Interconnects XIX*. Ed. by Henning Schröder and Ray T. Chen. San Francisco, United States: SPIE, Mar. 2019, p. 22.

[95] Marco Binda et al. “High Capacity Silicon Photonics Packaging”. In: *2019 22nd European Microelectronics and Packaging Conference Exhibition (EMPC)*. Sept. 2019, pp. 1–8.

[96] M. Ishii, Y. Hibino, and F. Hanawa. “Multiple 32-fiber array connection to silica waveguides on Si”. In: *IEEE Photonics Technology Letters* 8.3 (Mar. 1996), pp. 387–389.

[97] Martin Hoffmann and Peter Kopka. “Thermooptical Digital Switch Arrays in Silica-on-Silicon with Defined Zero-Voltage State”. In: *J. Lightwave Technol.* 16.3 (Mar. 1998), p. 395.

[98] Hiroshi Takahashi. “High performance planar lightwave circuit devices for large capacity transmission”. In: *Opt. Express* 19.26 (Dec. 2011), B173–B180.

[99] Antoine Kévorkian. “Passive and Active Glass Integrated Optics Devices”. In: *Fibre Optic Communication Devices*. Ed. by Norbert Grote and Herbert Venghaus. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, pp. 197–261.

[100] Jaime Cardenas et al. “Low loss etchless silicon photonic waveguides”. In: *Opt. Express* 17.6 (Mar. 2009), pp. 4752–4757.

[101] R. Pafchek et al. “Low-loss silicon-on-insulator shallow-ridge TE and TM waveguides formed using thermal oxidation”. In: *Appl. Opt.* 48.5 (Feb. 2009), pp. 958–963.

[102] Guillaume Beaudin, Ali Belarouci, and Vincent Aimez. “Precise localized thinning and vertical taper fabrication for silicon photonics using a modified local oxidation of silicon (LOCOS) fabrication process”. In: *Opt. Express* 23.4 (Feb. 2015), pp. 4377–4384.

[103] Minhao Pu et al. “Ultra-low-loss inverted taper coupler for silicon-on-insulator ridge waveguide”. In: *Optics Communications* 283.19 (2010), pp. 3678–3682.

[104] Nicolás Sherwood-Droz, Alexander Gondarenko, and Michal Lipson. “Oxidized Silicon-On-Insulator (OxSOI) from bulk silicon: a new photonic platform”. In: *Opt. Express* 18.6 (Mar. 2010), pp. 5785–5790.

[105] S. J. Spector and R. B. Swint. “Rib waveguide to strip waveguide mode converter using local oxidation of silicon (LOCOS)”. In: *2009 6th IEEE International Conference on Group IV Photonics*. 2009, pp. 83–85.

[106] Kevin K. Lee et al. “Fabrication of ultralow-loss Si/SiO2 waveguides by roughness reduction”. In: *Opt. Lett.* 26.23 (Dec. 2001), pp. 1888–1890.

[107] S. I. Raider et al. “Surface Oxidation of Silicon Nitride Films”. In: *Journal of The Electrochemical Society* 123.4 (Apr. 1976), pp. 560–565.

[108] Lee M. Loewenstein and Charlotte M. Tipton. “Chemical Etching of Thermally Oxidized Silicon Nitride: Comparison of Wet and Dry Etching Methods”. In: *Journal of The Electrochemical Society* 138.5 (May 1991), pp. 1389–1394.

[109] Boris Desiatov, Ilya Goykhman, and Uriel Levy. “Demonstration of submicron square-like silicon waveguide using optimized LOCOS process”. In: *Opt. Express* 18.18 (Aug. 2010), pp. 18592–18597.

Appendix A

High voltage parallel MEMS drivers

We fabricated three different high voltage driver boards for controlling the silicon photonics MEMS switch. I would like to acknowledge Kerry Yu for helping with the layout and test of these PCBs. The different versions are listed in Table A.1. Two versions are designed for the 4×4 switch (16 channels), and another is designed for the 16×16 switch (256 channels).

PCB Version	Control unit	High voltage driver	High voltage supply
4×4 microcontroller	STM32F303	HV7620 (32 channels, 200V)	MCP1661 Coupled-inductor boost (adjustable 40 V to 60 V)
4×4 FPGA	Cora Z7-07S	HV7620	LT8330 Voltage doubled boost (preset 40 V to 70 V)
16×16 FPGA	Cora Z7-07S	2× HV583 (128 channels, 80V)	LT8330 Voltage doubled boost (selectable 40 V, 50 V, 60 V or 70 V)

Table A.1: High voltage driver PCB versions

A.1 4×4 driver with microcontroller

The first board drives the 4×4 switch using a microcontroller and the HV7620 IC. The system architecture is shown in Figure A.1. The switch is wirebonded to a small PCB which connects to the control PCB via a ribbon cable. The control PCB has onboard a USB connector, a high voltage converter (powered by USB), the microcontroller, and the high voltage driver IC. A hardware trigger port is also added for synchronization with other equipment. The trigger takes 3.3 V CMOS logic voltages and operates on the rising edge.

The PCB is shown in Figure A.2. The switch wirebonding board mounts on top.

The microcontroller is controlled using a set of ASCII commands sent over a USB virtual serial port. Table A.2 lists each available command. The commands are terminated with a newline (\n)

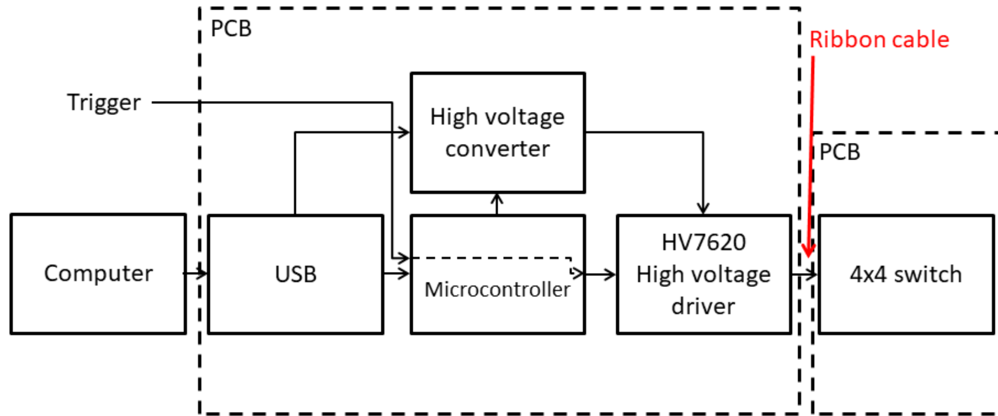


Figure A.1: Block schematic of 4×4 driver with microcontroller

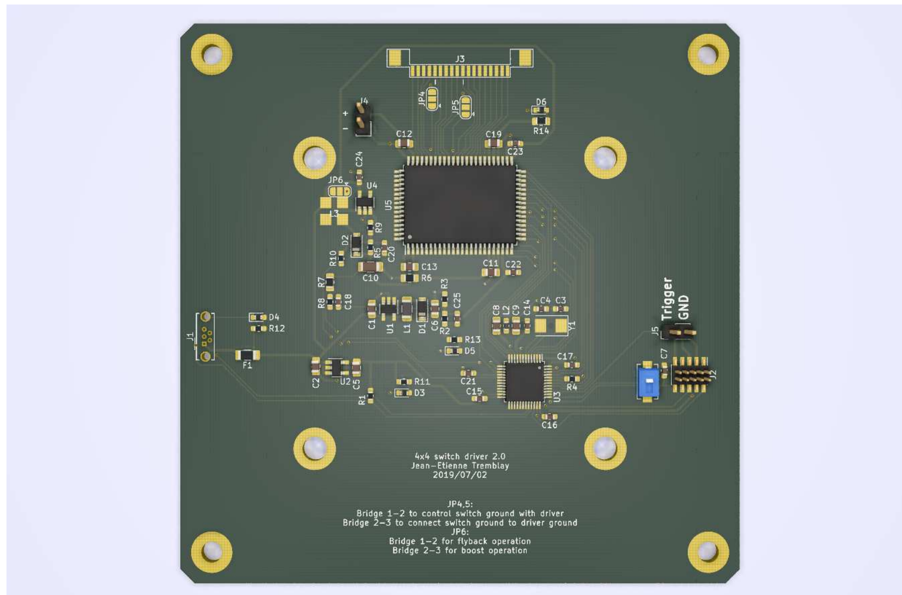


Figure A.2: PCB of 4×4 driver with microcontroller

character. A buffer of up to 32 switch configurations is available, and the microcontroller will cycle through the buffer on trigger (either software or hardware). The user first sets the buffer length n, then sends n S$x_1x_2x_3x_4$ commands to fill the buffer with different switch configurations. For simple testing, a buffer length of 1 can be used, and each S$x_1x_2x_3x_4$ command is followed by a trigger command.

Command	Function	Reply
Vxx	Set voltage in volts	"OK"
M	Measure voltage	Voltage in volts
R	Reset the buffer	OK
Lxx	Set the buffer length. Also resets the buffer.	"OK", or "ERROR" if longer than 32
T	Software trigger, increment the read pointer	Buffer element number which was sent
S$x_1x_2x_3x_4$	Set the enabled column for each row. x is 0 for not enabled, or 1 to 4 for the corresponding column. Increment the write pointer.	Buffer element number which was written, or "ERROR" if wrong format

Table A.2: Microcontroller commands for 4×4 switch

A.2 4×4 driver with FPGA

For more complex experiments, we also designed a breakout board that is controlled using a FPGA. The system is shown in Figure A.3. The switch board is connected using the same ribbon cable. The breakout control PCB contains the high voltage converter and the high voltage driver only. In this version, the output voltage is preset by a resistor. The FPGA board is separate, and we use typically the Cora Z7-07S from Digilent. The FPGA and the driver PCB are connected through a single 2×6 pin PMOD header. The FPGA provides the latch enable, clock, and four data lines to the HV7620 chip. The board includes a level translator that can operate between 1.8 V and 3.3 V, for compatibility with various FPGA technologies.

For testing, the FPGA code uses a similar set of commands as the microcontroller board (except for the voltage control) through a serial to USB adapter, and one of the FPGA pins can be dedicated to the trigger.

The PCB is shown in Figure A.4, and is similar to the microcontroller version.

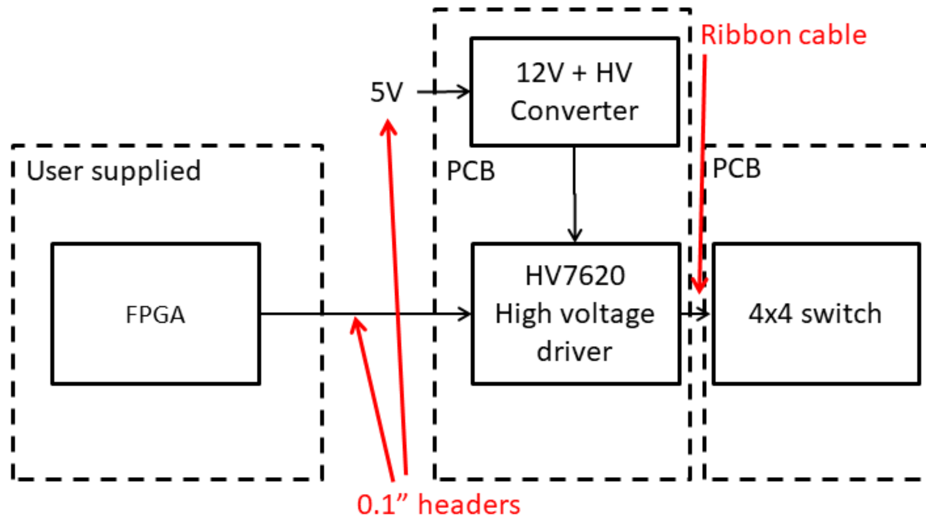


Figure A.3: Block schematic of 4×4 driver with FPGA

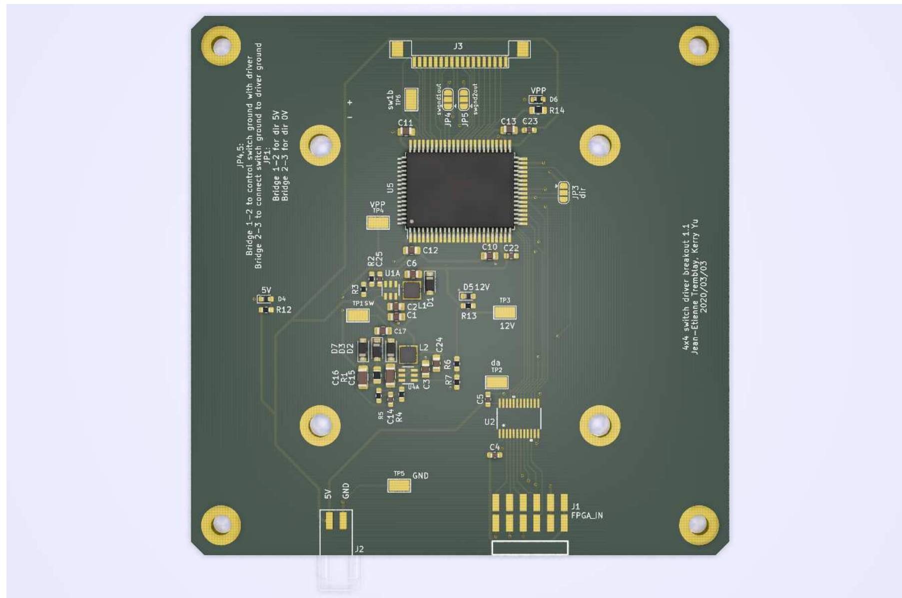


Figure A.4: PCB of 4×4 driver with FPGA

A.3 16 × 16 driver with FPGA

For larger switches, such as the 16 × 16 port switch, we designed a board that can provide up to 256 high voltage channels. This board uses two HV583 chips, which each have 128 channels up to 80 V. The system is shown in Figure A.5. Because of the large number of channels, we replaced the ribbon cable with a 280 pin connector (Hirose FX8C series with 2 × 140 contacts). The 16 × 16 switch is flip-chip bonded to an interposer PCB, which sits on top of the driver PCB. The driver PCB includes a high voltage converter with selectable output voltage, two HV583 chips and a level translator that operates between 1.8 V and 3.3 V. The driver PCB is connected to the FPGA board (we use the Cora Z7 FPGA) via two 2 × 6 pin PMOD headers. The connection uses 8 data channels, and two separate clocks and latch enable signals to operate the two HV583 chips independently. The PCB is shown in Figure A.6.

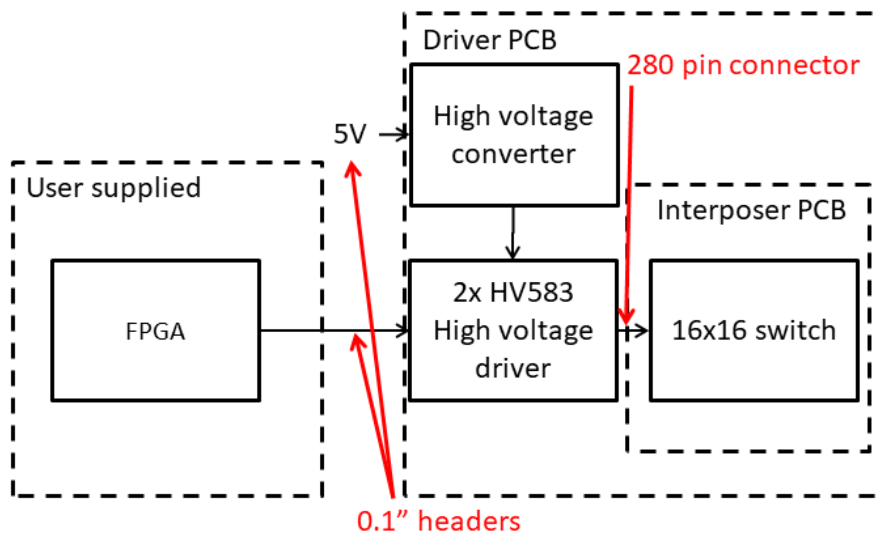


Figure A.5: Block schematic of 16 × 16 driver with FPGA

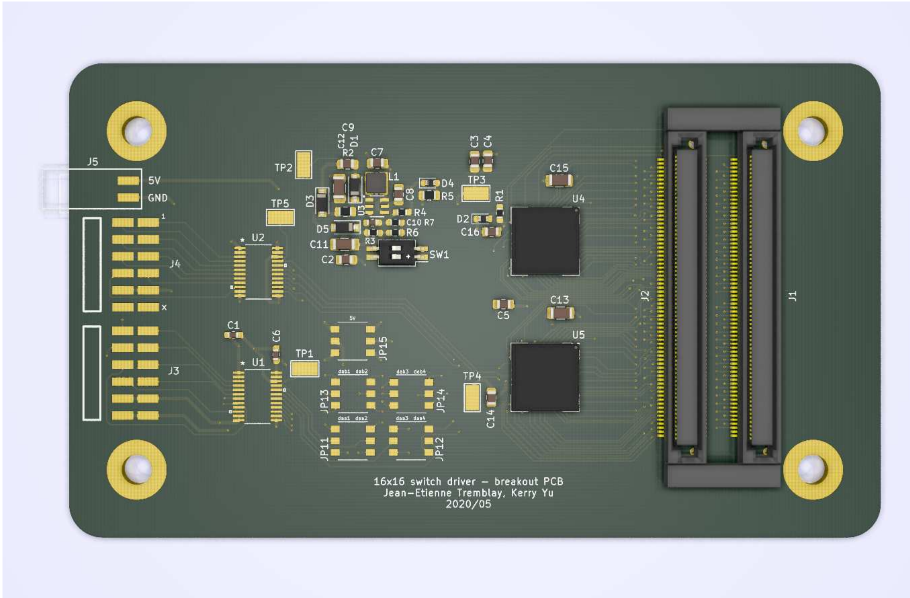


Figure A.6: PCB of 16×16 driver with FPGA

Appendix B

FiconTEC FL300 procedures

In this chapter, we detail the different components, programs, and alignment strategies of the FiconTEC FL300 machine.

B.1 Components

The FiconTEC components are divided into:

- Motion, including the force feedback
- Cameras and vision
- Data acquisition (DAQ), including the vacuum control and the optical/electrical feedback

Motion

The machine has two 6-axis arms for manipulating chips, and another 3-axis arm for the probe card. They can be controlled manually through the jog panel (Figure B.1). The jog panel can also enable or disable axes, and display a relative or absolute position for each axis.

The left arm (or pickup tool) is called PUT1, and the right one is called PUT2. The translation axes are labelled x, y and z (see Figure 6.9), and the rotation around these axes are R_x, R_y and R_z with the usual right-hand convention. The units of translation are µm, and the units of rotation are arcsecond.

Each arm has an addition axis for the gripper. The gripper closed position is 0, and they open towards negative values. The maximum opening is 10 mm.

The probe arm is labelled PR1 and has the x, y and z axis pointing in the same direction as the pickup tools.

PUT1 and PUT2 have a force sensor mounted near the gripper. The force sensor acquisition is routed through the motion system (labelled ACS in the software).

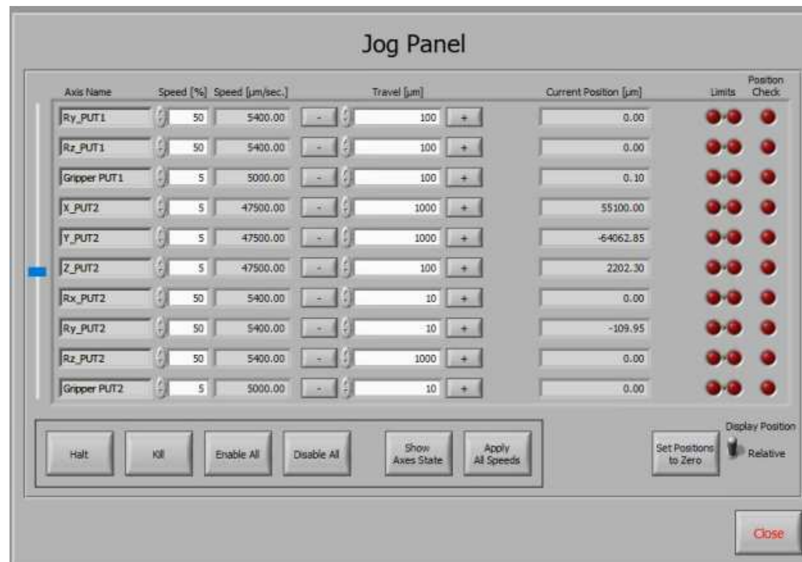


Figure B.1: Jog panel

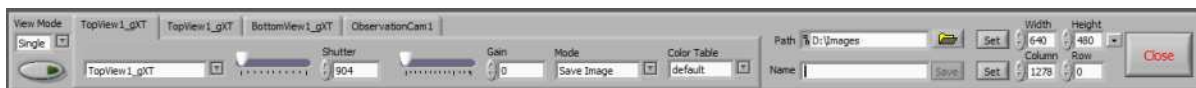


Figure B.2: Camera panel

Cameras

The machine is equipped with 4 cameras (Figure B.2): TopView1 and TopView2 are mounted on the two pickup tools. BottomView1 is located behind the trays and fixed to the machine. Each of these three cameras are monochrome and also have an illumination system. ObservationCam1 is a color camera and can be moved around the machine for monitoring the process.

The vision system has the following useful tools: autofocus, simple vision, and shape detection. The autofocus scans an axis (usually the z axis) and maximizes the camera contrast.

The vision algorithms (Figure B.3) are used to detect simple objects like corners and holes. Some typical algorithms include binary threshold, edge filters like Sobel and Kirsch, dilation and erosion, and circle or rectangle fitting. When the algorithm fails, the vision editor can ask for a manual input from the operator.

The shape editor (Figure B.4) is used to detect more complicated objects, such as fiducial marks. It uses simple edge detection and comparison algorithms, and usually works better when the illumination settings are consistent. This can be done using the "Shutter Measurement" tool to calibrate the exposure.

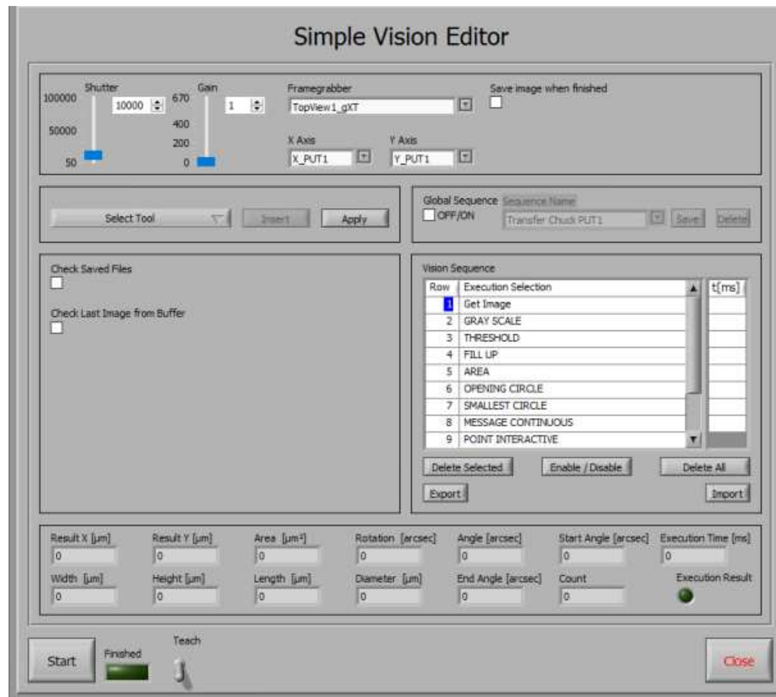


Figure B.3: Vision algorithm editor

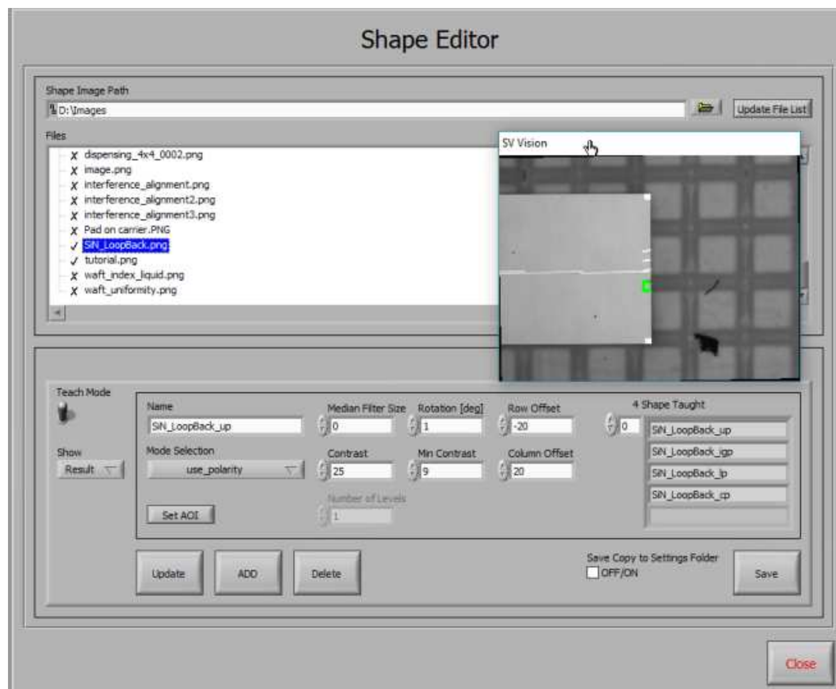


Figure B.4: Shape editor

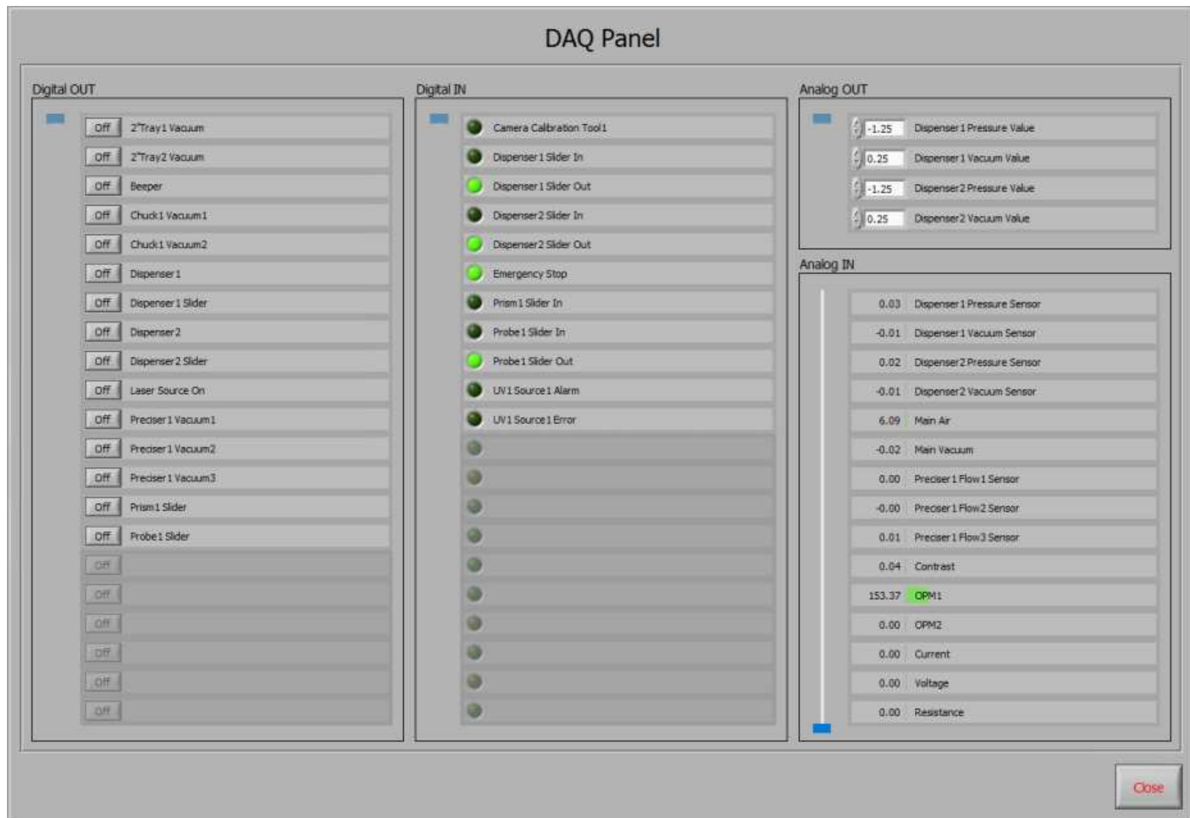


Figure B.5: DAQ panel

DAQ

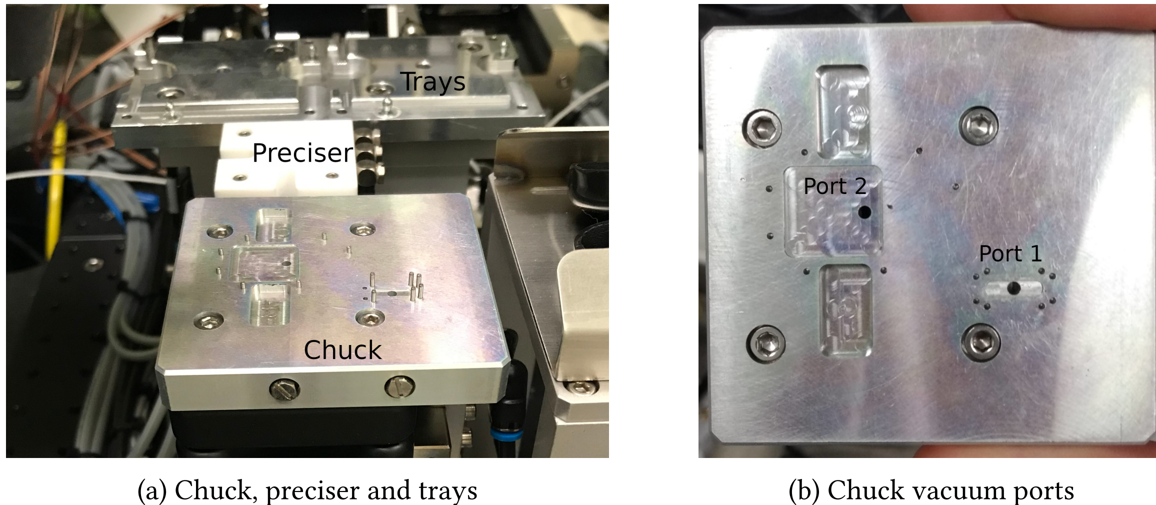
The acquisition system controls all the other analog and digital signals from the machine (Figure B.5).

These include the vacuum system, the compressed air actuators, and the electrical measurement tools.

The vacuum system is used to hold the parts in the various chucks and trays. The chuck in the front of the machine has two vacuum ports. Vacuum1 is the smaller port located on the right of the chuck, and Vacuum2 is the larger port on the left of the chuck (Figure B.6). Behind the main chuck is the preciser, an auxiliary chuck used to remount accurately the parts picked up from the trays. Behind the preciser are two 2 ” trays which can hold Gel-Paks or other similar chip trays. The trays also have a vacuum port to release the chips.

The compressed air controls the adhesive dispenser pressure and the sliders. Each dispenser is mounted on the pickup tool, and can be retracted via a slider. A prism can also be slid above the bottom view camera to convert it to a side view camera.

Various electrical tools are connected to the analog acquisition system. A Thorlabs LDC200C laser current driver is connected to the "Laser Source" signal. The two optical power meters have a wavelength range of 400 nm-1600 nm and 800 nm-2000 nm.



(a) Chuck, preciser and trays

(b) Chuck vacuum ports

Figure B.6: Chucks and mounts

A Keysight B2901A source measure unit is also connected to the machine. The voltage and current measurements can be routed to the process program through virtual analog channels.

Finally, the UV source for adhesive curing can be controller either manually, or through the software. The adjustable parameters are the time and power.

B.2 Programs

Here we list the most useful process programs, and add some notes on how to use them.

Calibration

In order to do precise alignment and packaging, each moving part of the machine needs to be calibrated to a reference. The origin of the system is the focal point of the bottom view camera, in the middle of the field of view. The calibration data is saved in the positions list.

The first step in calibrating the machine is to reference the TopView1 and TopView2 cameras to the bottom view camera. This is done using a glass slide with a metal fiducial (Figure B.7).

After calibrating the cameras, the various moving parts (grippers and dispensers) are calibrated to the bottom view camera using the vision system. They can then be referenced to the top view camera using the top view camera calibration data.

Finally, the trays, preciser and chuck are calibrated to both top view cameras using the vacuum ports as the reference point. They are calibrated separately to each moving arm since they might not be perfectly parallel.

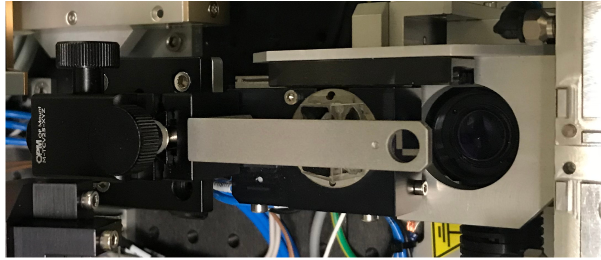


Figure B.7: Bottom view camera calibration tool

Edge coupling

The edge coupling process picks up a silicon chip from the tray, and aligns it to a mechanical carrier with an integrated laser chip mounted on the chuck. It can also be adapted to align a fiber array to a fixed silicon chip. The steps are as follow:

1. Define the location of the chip in the tray
2. Place the mechanical carrier on the chuck
3. Locate and measure the mechanical carrier and laser chip
4. Pick up the silicon chip from the tray
5. Measure the silicon chip using the bottom view camera
6. Align the silicon chip to the mounted laser chip
7. Dispense the adhesive
8. Realign and cure the adhesive

Before the alignment, the program calculates and compensates for the relative tilt between the two parts. After curing the transparent index-matching adhesive, it is recommended to manually add additional low-shrinkage structural adhesive to secure the part before opening the gripper. The different available alignment procedures are discussed in section B.3.

Grating coupling

The grating coupler procedure is simpler since there is no part to pickup. This program is intended to package a fiber array to a chip mounted on a PCB. The grating coupler is first mounted in the 10° gripper manually. Then the procedure is:

1. Load the PCB on the chuck
2. Locate and measure the PCB and chip

3. Locate the fiber array using the bottom view camera
4. Align the fiber array to the chip
5. Dispense the adhesive
6. Realign and cure the adhesive

Before dispensing the adhesive, it is recommended to do dispensing tests on a glass slide. The adhesive dispense pressure and suck-back vacuum can be set in the DAQ panel. Both are relative values, and a recommended starting point for the NOA146H adhesive is -0.1 to 0 for the pressure, and 0.25 for the vacuum.

B.3 Alignment

The choice of the alignment algorithm depends on the type of package, and in general they can be combined for a faster and more robust alignment.

Spiral search

As it name implies, the spiral search start at a given point and spirals around it until a given threshold is found. It works in two dimensions and is useful in finding the initial signal after the coarse visual alignment. It is relatively quick when the optimal position is nearby the initial point.

For grating coupler alignment, the x and y axes will be used for the spiral. For edge coupler, the y and z axes are used instead.

The spiral search can also be combined with a vision routine, e.g. to find parts which are not in the field of view. The image contrast can be fed to a virtual analog channel through the "Live Analysing Process" tool.

After the spiral search, we use the force feedback sensor to verify the point where the two parts start touching, and back-off slightly before the final alignment to prevent any damage.

2D area scan

While slower than the spiral search, an exhaustive 2D area scan can be useful in characterizing the alignment tolerance of a device. It operates on two axes and can sweep in a meander fashion to accelerate the scan.

Simplex search

The simplex algorithm is a powerful optimization procedure which can operate on multiple axes at the same time. Once the initial alignment is found using a spiral search, a simplex search can

be used to align axes which are not decoupled. This is often the case for rotation axes when the part is not exactly in the center of rotation.

1D profile scan

After finding the initial alignment position and optimizing the rotation axes, we perform a fine alignment by iteratively sweeping the most critical axis using a profile scan at low speeds (0.01 %). This is often useful after dispensing the epoxy. Indeed, the parts could have slightly shifted when they come in touch because of the adhesive surface tension.