Large Grain Ge Growth on Amorphous Substrates for CMOS Back-End-Of-Line Integration of Active Optoelectronic Devices

By

Brian Pearson

B.S. Mechanical Engineering
University of California, San Diego, 2010

Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Mechanical Engineering
at the
Massachusetts Institute of Technology

September 2012

Signature of Author:_____
Department of Mechanical Engineering
August 20, 2012

Certified by:_____
Lionel C. Kimerling
Thomas Lord Professor of Materials Science and Engineering
Thesis Supervisor

Certified by:_____
Tonio Buonassisi
SMA Assistant Professor of Mechanical Engineering and Manufacturing
Mechanical Engineering Thesis Reader

Accepted by:_____
David E. Hardt
Ralph E. and Eloise F. Cross Professor of Mechanical Engineering
Graduate Officer

Large Grain Ge Growth on Amorphous Substrates for CMOS Back-End-Of-Line Integration of Active Optoelectronic Devices

By

Brian Pearson

Submitted to the Department of Mechanical Engineering on August 20, 2012 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mechanical Engineering

Abstract

The electronic-photonic integrated circuit (EPIC) has emerged as a leading technology to surpass the interconnect bottlenecks that threaten to limit the progress of Moore's Law in microprocessors. Compared to conventional metal interconnects, photonic interconnects have the potential to increase bandwidth density while simultaneously reducing power consumption. However, photonic devices are orders of magnitude larger than electronic devices and therefore consume valuable substrate real estate. The ideal solution, in order to take advantage of optical interconnects without decreasing transistor counts, is to monolithically implement dense three-dimensional integration of electronics and photonics. This involves moving the photonic devices off the substrate, and into the metal interconnect stack. Moving photonic devices into the interconnect stack imposes two fabrication limitations. First, the available thermal budget allowed for photonic device processing is limited to 450 °C. Second, the metal interconnects are embedded within amorphous dielectrics and therefore there is no crystalline seed to initiate epitaxial growth.

This thesis addresses two major barriers for integration of photonics in the back end: (1) how to fabricate high quality Ge for active regions of optoelectronic devices while adhering to back-end processing constraints, and (2) how to couple optical power to these devices.

First, an approach was developed to fabricate the active region of Ge-based optoelectronic devices. A new technique, known as two-dimensional geometrically confined lateral growth (2D GCLG), has demonstrated single crystalline Ge on an amorphous substrate. This thesis presents the first application of the 2D GCLG technique to fill a lithographically defined SiO_2 trench with large grain Ge, while adhering to back-end processing constraints. A modified design is then proposed to increases the yield of 2D GCLG structures. This trench filling technique is an integral step towards fabricating Ge-based optoelectronic devices that are capable of being integrated into the back-end of a microprocessor.

Once it was established that high quality Ge trenches could be fabricated in the back-end, optical coupling to devices was addressed. For dense three-dimensional integration of photonic devices, vertical coupling between photonic planes is necessary. Therefore, this thesis begins with the design and simulation of vertical couplers. These couplers utilize evanescent coupling between two overlapping inversely tapered waveguides, which ensure efficient coupling due to optical impedance matching. These couplers are designed to exhibit coupling efficiencies in excess of 98.4%, equivalent to a 0.07 dB coupling loss.

The technique of evanescent coupling between overlapping inverse tapers is then applied to electro-absorption modulators (EAMs). A design for low-loss evanescent coupling from a waveguide to a Ge EAM is modeled and optimized. The design implements lateral evanescent coupling from overlapping inverse taper structures. Simulation results show that the coupling efficiency into and out of the modulator can be as high as 99%, equivalent to a 0.04 dB coupling loss.

Thesis Supervisor: Lionel C. Kimerling
Title: Thomas Lord Professor of Materials Science and Engineering

Acknowledgements

Many people have been instrumental in the path towards the completion of this thesis, as it was far from an entirely independent project. My first and foremost thanks goes to my thesis advisor, Professor Kimerling. I joined Kim's EMAT group as a Mechanical Engineering graduate, with only the most basic comprehension of materials science. Therefore, it was safe to say that I had a slow start in getting my project off its feet. However, Kim was always kind and supportive. His approach towards guiding a student-based research team is inspirational. He manages to make me question results and think critically, while maintaining an encouraging attitude. I could not have been more pleased with my experience with this research group and my advisor especially. His guidance, support, and encouragement have been invaluable to my progress in the field of Si photonics.

This thesis would never have been completed without the guidance of Dr. Jurgen Michel. I met with Jurgen, and the rest of the germanium team, every week to discuss research progress. He was immensely helpful with the day-to-day research guidance, as well as long-term planning. He was always accessible, and always willing to help. He seems to have experience and knowledge in every topic even moderately relevant to our research, and I have great respect for him.

I owe many thanks to the entire EMAT team. Dr. Anu Agarwal was the first person in the group that I met, while I was still an undergrad at UCSD. I was so nervous about being at MIT for the first time, but she greeted me with cake at our very first meeting and was nothing but kind and inviting right from the start. Dr. Kevin McComber, whose project I continued, was an excellent mentor for the year that we overlapped in the group. He walked me through his entire process flow, and made a considerable effort to pass on all of the knowledge he had gained about his project. I owe him many thanks for laying the foundational work for my project. As for current group members, I want to specifically thank the Ge team, especially Rodolfo Camacho and Yan Cai. Rodolfo spent many hours with me in the labs teaching me how to use tools, and helping me fix the reactor. Yan was always willing to help with daily research questions, ranging from process flow questions, to characterization techniques. I have also had meaningful discussions and assistance from the rest of the EMAT team.

There have been many people, outside of our research group, which have gone out of their way to help me in the lab. Paul Tierney, of the Microsystems Technology Laboratories, has spent countless hours helping me with the stepper. Lithography was one of the largest limitations with my device processing, and he was always willing to help. Mark Beals and Mayank Bulsara both spent a considerable amount of time in the Substrate Engineering Laboratory with me, helping to fix the broken UHVCVD. It was with their guidance that we were eventually able to get the tool back up and running, after many months of maintenance.

I have to thank the entire Fitzgerald group for putting up with me as the sole Kimerling group member in the Firzgerald group office. I have had countless good times with them ranging from Kunal Mukherjee teaching me semiconductor device physics, to discussing the meaning of life, to drinking beers on Tim Milakovich's balcony.

Finally, I have to thank my friends and family for helping to keep me sane. The support of my friends, sister, and parents were instrumental in helping to keep me motivated. They let me vent, gave me encouragement, and were always there for me.

Table of Contents

List of Figures and Tables

Chapter 1: Introduction

1.1 Motivation

The entire field of electrical interconnection is reaching a bottleneck. In order to keep up with Moore's Law, device dimensions continue to shrink in order to allow for denser integration of transistors in microprocessors. However, as the feature size of silicon complementary metal oxide semiconductor (CMOS) transistors continue to shrink, the industry is obtaining diminishing returns. The main limitations to perpetual reduction of feature size is due to increased interconnect delay, increased power consumption of interconnects, which cause increased heat dissipation requirements.

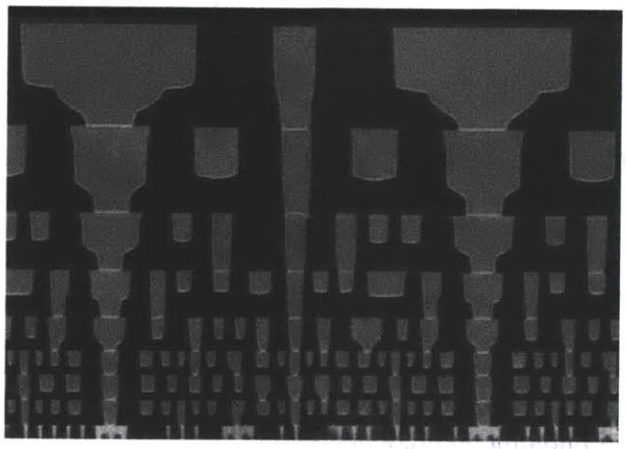

Figure 1.1 An example of a cross-sectional view of an 8 level interconnect stack from Intel. The small metal layers on the bottom are the local interconnects and the large metal layers on top are the global interconnects. From [1].

Metal interconnects have become increasingly complex as device dimensions continue to shrink within integrated circuits. Figure 1.1 shows an example of the metal interconnect stack in one of Intel's integrated circuits in the 48 nm node. As the size of the electronic devices continue

to shrink, the size of the metal interconnects must also shrink. However, this creates an increased RC delay due to increased resistance from smaller metal cross-sections. This effect can be seen in Figure 1.2. Once feature sizes reached 180 nm linewidths, the interconnect delay began dominating the total system delay. The reduced linewidths also create an increased problem with heat dissipation, due to Joule heating, caused by increased electrical resistance. By the 130 nm node, interconnections were already consuming 50% of the total microprocessor power [2]. For context, this means that in the United States, server interconnects consumed more power than all solar power that was generated within the United States in 2007 [3]. The International Tecnhology Roadmap for Semiconductors (ITRS) states that "power management is now the primary issue across most application segments" [4]. In addition, metal interconnects become limited by the Skin Effect at high bandwidths. As bandwidths increase, the electrical signal travels at the surface of the metal, thereby reducing the effective wire cross-sectional area. This increases the resistance of metal interconnect, additionally increasing Joule heating and power consumption. The combination of these factors leads to a limit in the benefits that metal interconnects can attain with increasing data rates and decreasing feature sizes. The recent enhancements, in bandwidth of metal interconnects, have come at the expense of increased latency and increased power consumption [5]. A promising way to overcome these problems is the by the introduction of photonic interconnects.

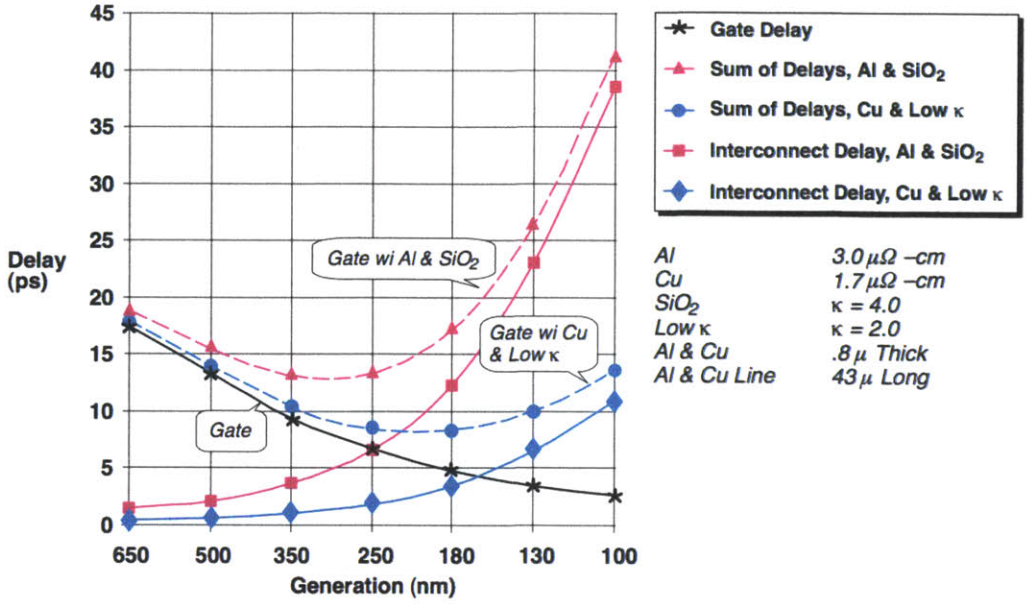

Figure 1.2 Delay as a function of feature size for metal interconnects. The gate delay (switching time) always decreases with decreasing feature size. However, the interconnect delay (propagation time) increases once linewidths are sufficiently reduced. Cu/ low-k interconnects have reached the crossover point, the point at which increasing interconnect delay outweighs decreasing gate delay, at the 180 nm node.

Photonic interconnects can solve the problem of heat dissipation due to the fact that photons do not generate heat while they travel through waveguides. Optical fibers have a very high data capacity, and can transmit data at much higher bandwidths than metal interconnects, which is evident in Figure 1.3. Another benefit of photonic interconnects is the ability for high levels of multiplexing. With optical interconnects, multiple discrete signals can be transmitted at different wavelengths along a single waveguide, without interacting with each other. This is called wavelength division multiplexing (WDM) and allows for much more data to be transmitted though a single interconnect.

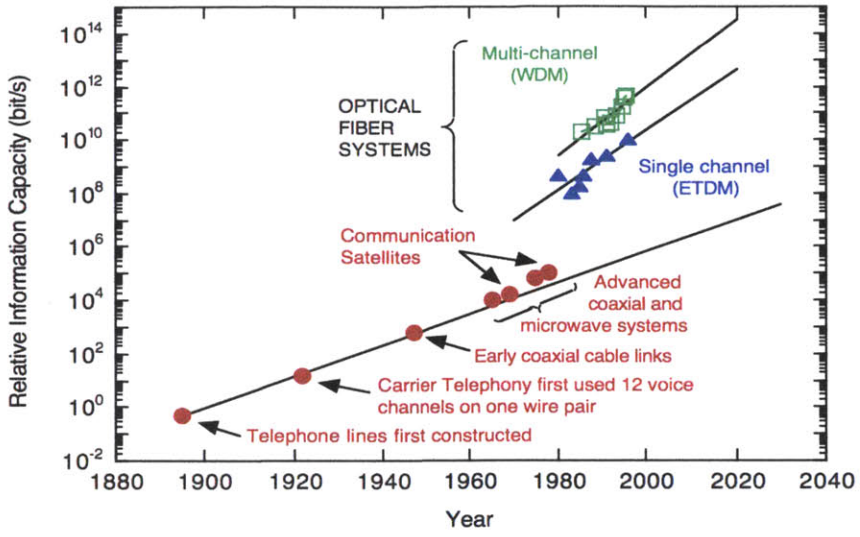

Figure 1.3 Information carrying capacity (bandwidth) of a single communication line. The circles indicate electrical systems while the triangles and squares indicate optical systems. The shift from electrical to optical communications has allowed for drastically increased bandwidths. From [6].

Optical communication has already proved its potential via the introduction of optical fiber systems. This paradigm shift to optical systems has yielded dramatically increased bandwidths. This trend is shown clearly in Figure 1.3. Optical communication is already widespread in telecom types of applications in which high data rates are required over long distances. This is due to the reasons listed above, namely the low loss and high bandwidths of optical fiber systems.

1.2 Benefits of Photonic Interconnects

Electronic interconnects are struggling to continue to increase bandwidths while maintaining low power consumption. It is in the combination of these two limitations that photonic interconnects show their inherent benefits. Optical interconnects have the potential to increase the interconnect density, increasing the bandwidth, while simultaneously reducing the

interconnect energy. This is due to the fundamental physical differences between the two types of interconnections.

The bit rate that can be transmitted through an electrical interconnect is limited by the resistive losses of the metal line. Due to metal resistive losses, in order to transmit high data rates over long distances, repeater amplifiers are required to maintain signal integrity. The total bandwidth that can be transmitted through a given cross-sectional area, known as the bandwidth density, of metal wire is inherently limited. The bandwidth that can be transmitted through a metal wire (B) is given by Equation 1.

$$B \leq B_0 \frac{A}{L^2} \tag{1}$$

In Equation 1, A is the cross-sectional area of the wire, and L is the length of the interconnect. B_0 is a constant with a value of approximately 10^{16} bit/s [7]. For a given interconnect distance, it is impossible to increase the maximum bandwidth of the line without increasing the cross-sectional area of the line, which therefore decreases the bandwidth density. On the other hand, photonic interconnects do not suffer from resistive losses. This means that they can transmit much higher bandwidths over long distances, and within a limited cross-sectional area. In addition, different wavelengths of light do not exhibit cross-talk with each other and hence, multiple signals can be transmitted down the same interconnection line, using wavelength division multiplexing, to further increase the bandwidth density.

In addition to the bandwidth density benefits of optical interconnects, there may also be energy benefits. The drawback with electrical interconnections is that the entire interconnection line must be charged to the signaling voltage for each bit to be transmitted. The energy per bit for an electrical interconnection (E_e) is given by Equation 2.

$$E_e \geq C_l V_s^2 \tag{2}$$

The capacitance of the line, C_l, is approximately a equivalent for a well-designed line and is equal to ~2 pF/cm [7]. Therefore, reducing the signaling voltage, V_s, is the only way to reduce energy, per bit, of an electrical interconnection line. In contrast, the energy consumption per bit of optical interconnection is determined by completely different physics. In order to transmit an optical signal, there is no need to electrically charge the entire interconnection to the signaling voltage. Instead, the energy per bit is dominated by energy required to charge and discharge the photodetector capacitance, C_d, and the signaling voltage, V_s, which the link is electrically connected to. The total energy per bit of a photonic interconnect, E_p, is given by Equation 3.

$$E_p \geq C_d V_s \frac{\hbar\omega}{e} \tag{3}$$

The photon energy is determined by $\hbar\omega$ where $\hbar$ is the reduced Plank's constant and ω is the angular frequency of the photon. The e term is the elemental charge. In order for the photonic link to require less energy per bit, the charge in the photodetector must be less than the charge required to bring the metal line up to the signaling voltage. Therefore, $C_d\hbar\omega/e < C_l V_s$ must be achieved. This is assuming the external quantum efficiency of the photodetector is unity, and neglecting additional transmission losses in the photonic link. If it is assumed that a wavelength of 1.55 μm is used in the photonic link, $\hbar\omega/e = 0.8\,V$. In comparison, signaling voltages in electrical interconnects can be as low as 0.1 V, or less [3]. Consequently, in order for photonic links to consume less energy than electrical links the relation $C_d \ll C_l$ must be satisfied, with the detector capacitance approximately an order of magnitude lower than the electrical line capacitance. The capacitance of the line increases linearly for the electrical case, while the detector capacitance remains constant for a given design. Therefore, there is a cross-over length, above which photonic links become more energy efficient than electrical interconnects.

Neglecting losses in the waveguides, and energy lost in electrical to optical conversion, this cross-over length can be as short as 50 μm [3]. Other studies yield higher cross-over lengths with values in the several millimeter to ~15 mm range [8], [9].

Since the energy benefits of photonic interconnects are prevalent at longer interconnection lengths, the application of optical links is focused on the replacement of global interconnects, not local interconnections.

1.3 Dense 3D Integration of Photonic Interconnects

There are four main components that make up a typical optical interconnect: a light source, a modulator, a waveguide, and a detector. The light source is the source of the photons. The modulator is a device that encodes the digital signal by effectively switching the light on and off. The waveguide is medium in which the light travels, and directs the light along a given path. The detector converts the optical signal back into an electrical signal. Each of these discrete devices have already been independently demonstrated and each of them are completely compatible with standard CMOS processing constraints [10–13]. Previous efforts have already demonstrated germanium based p-i-n heterojunction diodes that can be used as electro-absorption modulators and detectors [11], [12]. These devices show a lot of promise as the key devices in optical interconnects. However, in comparison to the electronic devices, these devices are very large. They may range in size, but modulators and detectors are in the range of 50 to 80 microns long and 0.5 to 1 micron wide for germanium-silicon active regions. This is orders of magnitude larger than electrical devices, since electrical device feature sizes are in the tens of nanometers size range. Therefore, the introduction of photonic devices will consume a lot of valuable real estate on the silicon wafer.

In order to follow Moore's law of increasing transistor counts, there is no room to put such large photonic devices onto the silicon substrate. A solution to this problem is to introduce three-dimensional integration of the photonic devices. In order to allow transistor counts to continue to increase, the photonic devices need to be taken off of the substrate and integrated in a level above the electronics level. This would introduce all of the benefits of the optical interconnect, without sacrificing any of the valuable real estate on the crystalline silicon wafer. In addition, the initial goal of photonic interconnects is to replace electrical global interconnects. Global interconnects are located at the top of the interconnect stack, furthest from the substrate. Therefore, placing the photonic interconnects high within the interconnect stack reduces the architectural modifications that must be adopted for the integration of photonic interconnects.

The electrical and optical properties of photonic devices are very dependent on their crystalline quality. Ideally, photonic devices have a single-crystalline active region. Grain boundaries and defects can increase dark currents, induce optical scattering, act as carrier recombination centers, and decrease the overall performance of these devices [14–16]. In order to fabricate single crystal devices with low levels of defects and contamination, the devices are typically grown on the substrate via ultra-high vacuum chemical vapor deposition (UHVCVD). Germanium is typically grown at 600 - 700 °C and may be annealed at temperatures up to 900 °C in order to reduce the threading dislocation density [17–19]. An important aspect of this process is that the crystalline silicon substrate is used to seed epitaxial growth of Ge films.

Germanium has emerged as an excellent candidate for the active region in active optoelectronic devices [6], [19–21]. The direct bandgap of Ge, 0.8 eV, corresponds with an optical wavelength of 1.55 μm, which is the wavelength that has the lowest absorption loss in silicon dioxide [22]. At this wavelength, absorption losses in Si-core, SiO_2-clad waveguides can

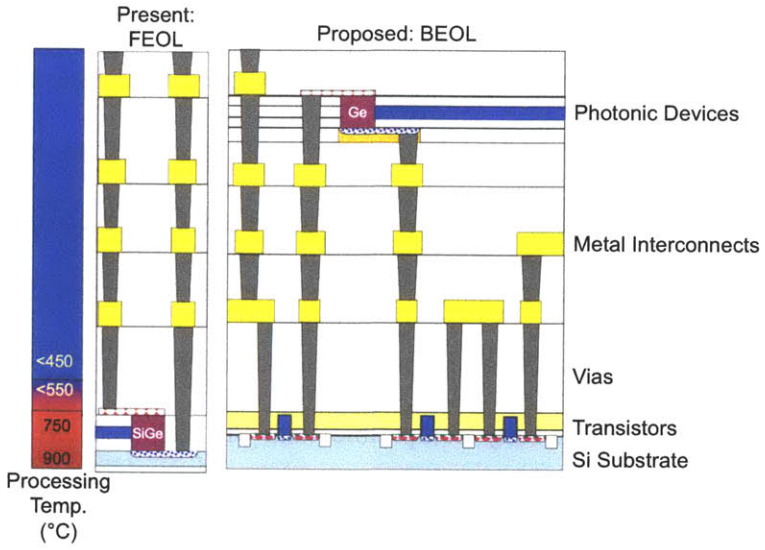

Figure 1.4 Cross-sectional schematic of a microprocessor interconnect stack. Current photonic devices (left) are fabricated in the front-end-of-line (FEOL), in which high-temperature epitaxy is permitted on the crystalline substrate. The right side shows the proposed architecture, in which photonic devices are integrated in the back-end-of-line (BEOL). Photonic devices are in purple and transistors are in blue. Processing temperatures, as a function of distance from the substrate, are shown in the scale on the left. From [21].

A schematic of the proposed microprocessor architecture is shown in Figure 1.4. In the left schematic, the current front-end integration design is shown where the photonic devices are fabricated on the substrate along with the transistors. The schematic on the right side shows the proposed design, in which photonic devices integrated within the interconnect stack. This frees up the substrate to be used for dense integration of transistors, but also restricts the thermal budget allowed for photonic device fabrication to not exceed 450 °C. It also eliminates the potential for epitaxial growth on a crystalline substrate.

Recent research has proven a technique to grow good quality Ge without the crystalline seed, at temperatures below 450 °C [26]. The technique is based on two-dimensional geometrically-confined lateral growth (2D GCLG). Essentially, the Ge is selectively nucleated on a small amorphous Si seed. Many different Ge crystals nucleate and grow in all different

be neglected with respect to scattering losses [23], [24]. In addition to having an appropriate bandgap, Ge has a high carrier mobility, with an electron mobility of 3900 $cm^2/sec \cdot V$ and a hole mobility of 1900 $cm^2/sec \cdot V$, approximately 4 times higher than the carrier mobilities in Si. In addition to the attractive material properties, Ge is completely CMOS compatible. Ge is already included in CMOS technology as a method to create strained Si transistors [25], and therefore Ge is fully compatible with integration into Si electronics and integrated circuits.

Although the Ge itself is fully compatible with Si integrated circuits, there are many significant processing challenges that arise when attempting to moving the germanium optoelectronic devices away from the crystalline substrate. When the devices are grown above the substrate, there is no longer a crystalline seed to grow Ge epitaxially. A completely new growth technique must be utilized in order to obtain large-grain Ge without a crystalline growth seed. In addition, Ge growth and annealing occurs at high temperatures. If the photonic devices are integrated above the electronics level, then the electronics and metal interconnects will already be fabricated. At high temperatures, there may be significant dopant diffusion and silicidation of metal contacts, thereby altering the performance of the electronics that have already been fabricated. Therefore, in order to preserve the electronic devices, processing temperatures must be kept below 450 °C. Therefore, the active region of the devices has to be grown at low temperature, without a crystalline seed.

orientations. Some crystallographic orientations grow faster than others. The amorphous Si seed is recessed within a high aspect-ratio silicon dioxide channel, and therefore only the fastest growing Ge grains emerge from the channel. With the correct design of the channel dimensions, this technique can ensure that the only thing that emerges from the channel is a single grain of crystalline Ge. This growth technique has shown a viable proof of concept, however, it has never been used to fabricate actual devices.

In order to integrate optical interconnects into the back-end of a microprocessor, three significant barriers must be achieved: (1) how to fabricate high quality Ge for active regions of optoelectronic devices, while adhering to back-end processing constraints, (2) how to couple optical power to these devices, and (3) how to metallize and electrically contact these devices. This thesis addresses the first two of these three items, namely, how to fabricate high quality Ge for the active region of optoelectronic devices, and how to couple light to these devices.

1.4 Outline of Thesis

In Chapter 2, a technique to fabricate large grain Ge, constrained by back end processing limitations, is assessed. The fundamentals of 2D geometrically confined lateral growth (2D GCLG) are explored. Then, the 2D GCLG process is implemented in a design that is used to fill lithographically defined oxide trenches with large grain Ge, while confined to back-end processing constraints. Initial fabrication progress proves this method is viable. Finally an improved design is presented that increases reliability and yield.

In Chapter 3, once it was determined that trenches could be filled, given the processing limitations, how to efficiently couple light was addressed. Evanescently coupled devices are designed. This work is in order to optimize individual components that will be necessary in a photonic interconnect. First, a vertical coupler is simulated and designed in order to allow 3D

coupling with nearly perfect coupling efficiency. These devices are necessary if the photonic interconnects are to be densely 3D integrated. This new knowledge of evanescent coupling with tapers is then applied to a Ge electro-absorption modulator. A new coupling design is outlined and modeled that reduces the coupling losses to 0.04 dB loss.

Chapter 4 summarizes the work that has been done in this thesis. It also proposes a plan for future work in the continuing effort to integrate photonic devices into the back-end of a microprocessor.

Chapter 2: Back-End Trench Filling

In order to integrate photonics into the back-end of a microprocessor, a technique must be devised to fabricate high crystal quality germanium on an amorphous substrate, while limited to a maximum processing temperature of 450 °C. For a technique to be applicable to integrated optoelectronic devices, it must be possible to fill lithographically defined trenches with this high quality Ge. This chapter presents a technique that is capable of doing this.

2.1 Two-Dimensional Geometrically-Confined Lateral Growth

In order to integrate Ge optoelectronic devices in the CMOS back-end-of-line, photonic device processing must comply with BEOL restrictions, namely thermal budget and a lack of a crystalline substrate. The processing temperature must be kept below 450 °C and the Ge must be deposited non-epitaxially. These are major processing limitations, which have recently been overcome with a method called two-dimensional geometrically confined lateral growth.

2.1.1 Background Theory

Two-dimensional geometrically confined lateral growth (2D GCLG) is a method to grow single crystal germanium on an amorphous substrate at low temperature [26]. The technique takes advantage of selective deposition, grain growth velocity anisotropy, and twinning.

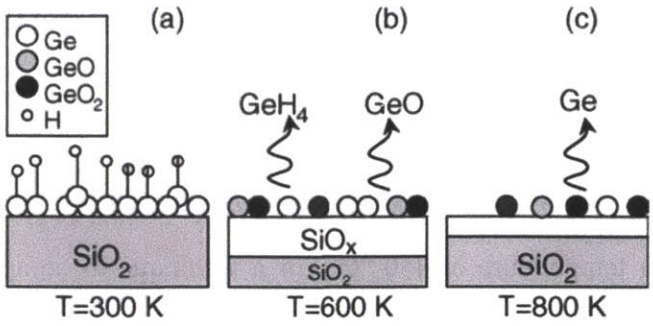

Figure 2.1 Ge interactions with a SiO_2 film. Figure (a) shows GeH_4 decomposition to deposit Ge film at 300 K. Figure (b) shows Ge stripping O from the SiO_2 to create GeO, the subsequent desorption of the volatile GeO, and the formation and desorption of GeH_4 at 600 K. At an even higher temperature of 800 K, (c) shows the further breakdown of GeO_2 and the desorption of Ge directly. From [27].

Ge deposits on a-Si, while it does not deposit on oxide. This is attributed to the volatility of germanium oxide at elevated temperatures. Stanly et. al. determined that when Ge is deposited on SiO_2, the Ge oxidizes, while simultaneously etching the SiO_2 [27]. Beginning at a temperature of 500 K (~227 °C), the Ge adatoms on the SiO_2 surface strip oxygen from the SiO_2 film. This oxidizes the Ge film, creating GeO as well as GeO_2. At this temperature, GeO is volatile and desorbs from the surface. This results in a non-stoichiometric (Si rich) SiO_2 layer at the surface of the film, in addition to an overall decrease in Ge from the film surface. At this temperature, Ge also desorbs in the form of GeH_4. This is caused by the hydrogenation of Ge from surface hydrogen, either from a H_2 carrier gas or GeH_4 decomposition. At further increased temperatures, the GeO_2 decomposes first into GeO, and then is reduced by the non-stoichiometric SiO_x to produce gas phase Ge and stoichiometric SiO_2. This overall process is shown schematically Figure 2.1.

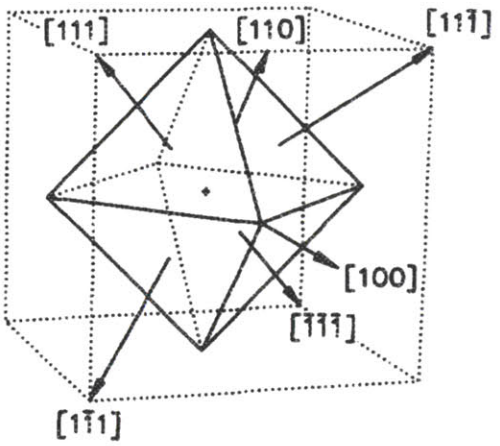

Figure 2.2 Schematic drawing of the octahedron that is formed if a cubic crystal is completely bounded by {111} facets. The edges between the planes indicate the ⟨110⟩ directions, while the tips where four {111} planes meet represent the ⟨100⟩ directions. From [28].

In addition to selective deposition, the 2D GCLG process takes advantage of grain growth anisotropy. Crystalline Ge forms a diamond cubic crystal structure. In a diamond cubic lattice, the {111} planes have the highest atomic packing density, and therefore the lowest surface free energy [29]. This means that the {111} planes are stable, and therefore exhibit slow growth. This yields {111} facets in poly-Ge growth. If only {111} facets are exposed, then two other grain orientations are still possible within the confines of a completely {111} bounded crystal. This hypothetical crystal forms an octahedron, which is shown schematically in Figure 2.2. The edges between the {111} planes create the ⟨110⟩ directions and the corners where the four planes meet create the ⟨100⟩ directions. From a geometric perspective, if all of the {111} planes grow at the same rate, then growth in the ⟨100⟩ directions will be the fastest, growth in the ⟨110⟩ directions will be at an intermediate rate, and growth in the ⟨111⟩ directions will be the slowest. This grain growth velocity anisotropy is shown from a geometric perspective in Figure 2.3. Here, the dotted lines indicate additional monolayer formation on the {111} planes. This monolayer formation will add one lattice spacing, a, in the ⟨111⟩ directions. However,

from the schematic shown, the amount of growth in the [101] direction will be $\frac{a}{\sin(54.74)} \approx 1.22a$. Therefore, growth in the $\langle 110 \rangle$ directions is about 22% faster than growth in the $\langle 111 \rangle$ directions. Following the same logic, in schematic (b) of Figure 2.3, growth in the [010] direction will be $\frac{a}{\sin(35.26)} \approx 1.73a$ for a monolayer addition on the $\{111\}$ planes. For each monolayer, a, addition to the $\{111\}$ planes, the $\langle 111 \rangle$ directions will grow by a distance a, the $\langle 110 \rangle$ directions will grow a distance $1.22a$, and the $\langle 100 \rangle$ directions will grow by a distance $1.73a$. Therefore, even if the crystal is completely bounded by slow growing $\{111\}$ planes, then there will still be grain growth anisotropy such that the growth velocity in the $\langle 100 \rangle$ direction is greater than the growth velocity in the $\langle 110 \rangle$ direction, which is greater than the growth velocity in the $\langle 111 \rangle$ direction.

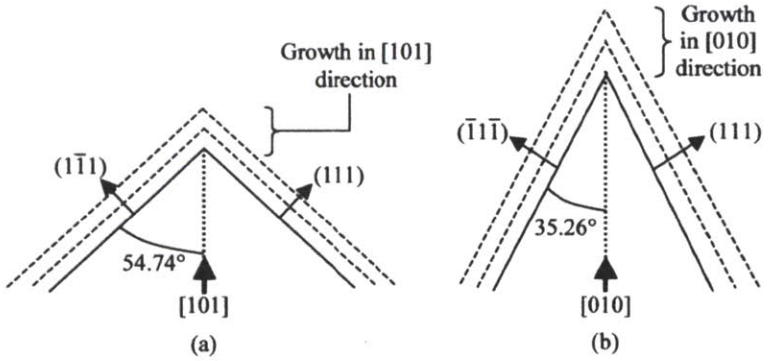

Figure 2.3 A geometric approach to visualizing grain growth velocity anisotropy. The schematics assume that the grains are bounded by {111} planes and growth is characterized by the slow growth rate in the [111] directions. Figure (a) shows the relative growth in a [110] direction bounded by two {111} planes. Figure (b) shows the relative growth in a [100] direction that is bounded by {111} planes. Growth in the [100] and [110] directions are both greater than the growth in the [111] directions, while growth in the [100] is clearly greater than growth in the [110]. From [30].

In a polycrystalline film, there is a distribution of grain orientations that nucleate on the substrate. The orientation of the grain is determined by which crystallographic direction is

oriented such that it is normal to the substrate. A film is said to be textured if one grain orientation is predominant. For example, if a majority of the grains in a polycrystalline film are oriented such that their ⟨110⟩ direction is normal to the substrate, then that film exhibits a ⟨110⟩ texture. This may occur if a certain grain orientation grows much more quickly than the other orientations. If one grain orientation grows much more quickly than the others, then, with a thick enough film, the fast growing grains will overtake the slow growing grains and the film will become textured.

The concepts of grain growth velocity anisotropy and textured films can be combined to predict the texture of a polycrystalline germanium film. This may lead to the hypothesis that a polycrystalline Ge film will exhibit a ⟨100⟩ texture since growth in the ⟨100⟩ direction is predicted to be the fastest. However, this overlooks the effects of twinning. Twin grain boundaries form readily in Ge because there is almost no energy required for the formation of a Σ3 twin grain boundary [31]. This is because Σ3 grain boundaries in Ge are completely coherent. Therefore, there are no dangling bonds within the grain boundary, no additional energy states are formed, and they have a negligible energy of formation. With negligible energy of formation, Ge readily forms twins on {111} planes. These twins occur such that the angle between the original lattice and the post-twin lattice is 60°. When looking at Figure 2.2, it is apparent that no two ⟨100⟩ directions form a 60° angle between them. They are all orthogonal to one another. However, if a grain is growing with a ⟨110⟩ orientation and twins, then it can continue to grow with a ⟨110⟩ orientation since some ⟨110⟩ directions form a 60° angle between them. This means that grains that nucleate with a ⟨110⟩ orientation can sustain long-range growth in the presence of twinning, while ⟨100⟩ orientated grains cannot. Grains with a ⟨100⟩ orientation before twinning no longer have a ⟨100⟩ direction that is normal to the substrate after twinning.

2.1.2 2D GCLG Technique

Two-dimensional geometrically confined lateral growth takes advantage of selective deposition, grain growth velocity anisotropy, and twinning, in order to deposit single crystal Ge on an amorphous substrate. A schematic of the basic process flow is shown in Figure 2.4. First, a silicon substrate is oxidized in order to create an oxide pseudo-substrate. This is the amorphous substrate that the Ge deposition occurs on. Then, a thin a-Si film is deposited on the oxide. The film thickness is typically around 50 nm. The a-Si is then patterned into a thin line (100 – 300 nm in width). After the a-Si line is patterned, a 200 – 500 nm thick oxide overlay film is deposited by plasma-enhanced chemical vapor deposition (PECVD). Then, a reactive ion etch (RIE) is performed to etch a window through the oxide overlay, to the oxide pseudo-substrate, exposing the a-Si line. Then, the a-Si line is etched through with a selective RIE. At this point, there is a thin a-Si line embedded in oxide, and exposed to air at one end. The schematic in part (a) of Figure 2.4 shows what the processing looks like up to this point. Next, a selective wet etch is performed with TMAH in order to undercut the a-Si. At this point, shown by part (b) in Figure 2.4, there is an a-Si line that is recessed within a high aspect-ratio channel that has oxide sidewalls. Finally, the structure is ready for Ge growth. The growth takes place within an ultra-high vacuum chemical vapor deposition reactor (UHVCVD) at 450 °C. The process is completely CMOS compatible and all processing takes place at or below 450 °C so that the process can be implemented in BEOL integration.

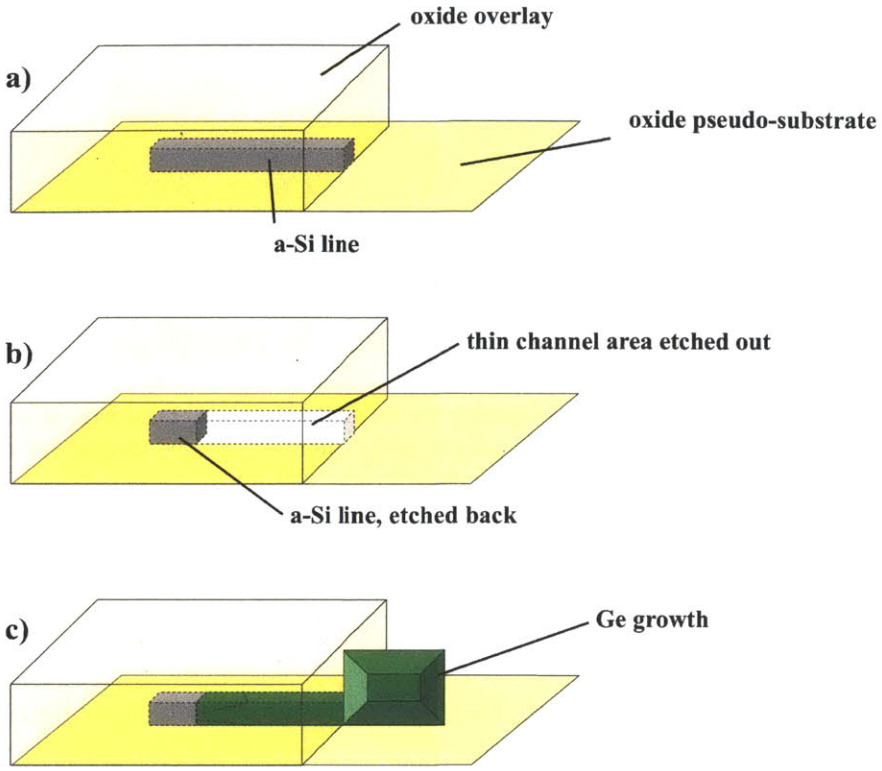

Figure 2.4 Fabrication process for 2D GCLG. First a thin a-Si line is patterned on an oxide pseudo-substrate. Then, a PECVD oxide is deposited on the line. Next, a window is opened in the oxide, exposing the a-Si line, ash shown in (a). Then, as shown in (b), the a-Si is undercut etched to provide a growth channel. Finally, Ge is selectively deposited on the a-Si and not the oxide via UHVCVD. The final structure is shown in (c). From [30].

During the Ge growth in the UHVCVD, the Ge selectively deposits on the Si, and not on the oxide. Therefore, polycrystalline Ge deposition initiates on the a-Si seed at the back of the oxide channel. If the correct oxide channel geometry is chosen, then, by the time the Ge emerges from the channel, only a single grain emerges and the emerging Ge is crystalline. Once the Ge emerges from the channel, the exposed Ge is now single crystalline, and further growth can occur epitaxially. This is shown schematically in part (c) of Figure 2.4. A representative example of a Ge growth via the 2D GCLG method is shown in Figure 2.5. The large facets on the Ge crystallite indicate that the growth is crystalline.

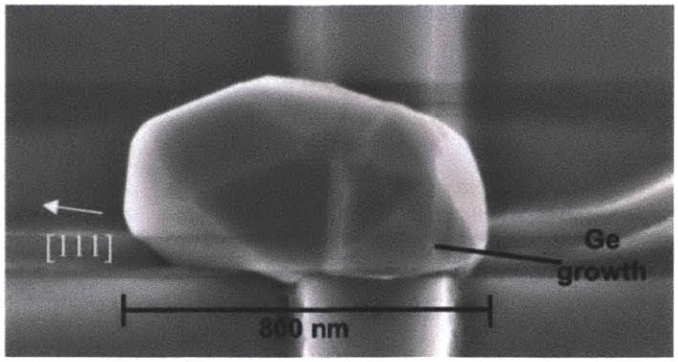

Figure 2.5 A representative example of Ge grown by the 2D GCLG method. The image is obtained by plan-view SEM. The bright vertical line below the Ge growth is the confining channel from which the Ge emerged. The large facets indicate that the Ge growth is crystalline. From [26].

2.1.3 2D GCLG Mechanism

The two-dimensional geometrically confined lateral growth technique takes advantage of grain growth velocity anisotropy, Ge selective deposition, and twinning in order to obtain single crystalline Ge on amorphous substrates at low temperature.

When the Ge growth begins, the Ge nucleates on the Si only, and not on the SiO_2. Since the Si growth seed is amorphous, the Ge grains which nucleate have a randomized orientation. The GeH_4 decomposes via pyrolysis and Ge adsorbs onto the Si surface. The then diffuses along the surface until a stable cluster of adatoms forms. The stable clusters form all over the Si surface and are the initial grains from which the polycrystalline film grows. The initial grains that nucleate have a random grain orientation. During the growth process, the faster growing grains overtake the slow growing grains such that the grain growth velocity anisotropy tends to eliminate the slow growing grains. An example of this is shown in Figure 2.6. In the figure, a 2D simulation is performed, in which the growth mechanism is assumed to be monolayer-by-monolayer addition on the low surface energy edges. The tips of the 2D grains grow faster than the edges, purely because of the geometric advantage as described in Figure 2.3. The initial film

is an array of randomly distributed and randomly oriented grains. The figure shows that as the film thickness increases, the film becomes textured, such that the fast growing grain orientations overtake the slow growing grain orientations. This is exactly what happens in the 2D GCLG technique. The slow growing ⟨111⟩ oriented grains are overtaken by the faster growing ⟨110⟩ and ⟨100⟩ oriented grains.

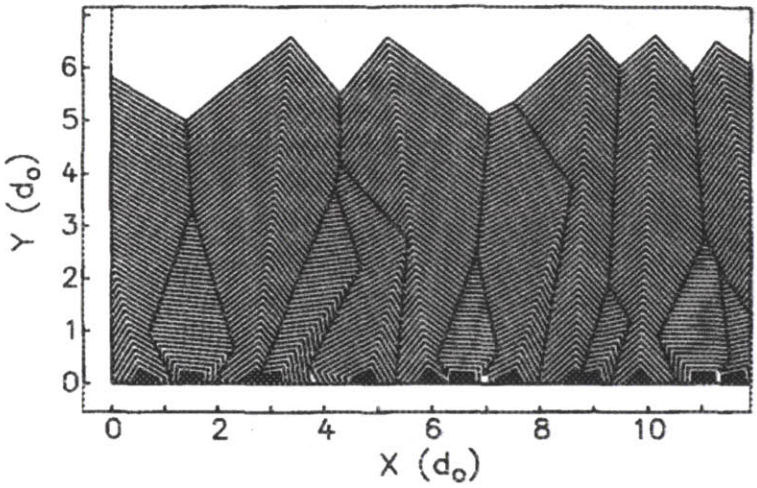

Figure 2.6 Two dimensional simulation of a polycrystalline film that evolves from grain growth velocity anisotropy. The black rectangular and triangular structures at the bottom of the figure (Y=0) are the randomly nucleated grains. The model assumes that the grains grow via monolayer-by-monolayer addition on the low energy edges of the grain. Purely due to the geometry of the grains, the tips of the grains grow faster than the edges. As the film grows thicker, the fast growing grains (tip oriented towards surface) overtake the slow growing grains and the film becomes textured. From [28].

If this grain growth anisotropy model is followed, then it is predicted that the fastest growing grain orientation, the ⟨100⟩ orientation, will dominate the growth. However, the Si growth plane is recessed within a high aspect ratio SiO_2 growth channel and the channel sidewalls stop this from occurring. This is due to the tendency for Ge to form Σ3 twins. Even if a grain is oriented such that the ⟨100⟩ direction is oriented towards the opening of the growth channel, it is unlikely that it will emerge from the channel, assuming that the channel aspect ratio is high. This is because it is likely to form a twin before it can emerge from the channel. When

Ge forms a twin, the angle between the newly formed grain and the original grain is 60°. This means that while the original grain had its fast growing direction oriented towards the channel exit, the grain after the twin will have its fast growing direction oriented towards the channel wall. Therefore twinning effectively terminates the ⟨100⟩ oriented grains by causing them to self terminate in the oxide sidewalls. On the other hand, a grain that initially has a ⟨110⟩ orientation can continue to grow in a ⟨110⟩ direction even after twinning. This is because there exists ⟨110⟩ directions that form a 60° angle between them. While the ⟨100⟩ grain will be oriented towards the channel sidewalls after twinning, the ⟨110⟩ grain may still be oriented towards the channel exit even after twinning. Therefore, if the channel geometry is designed correctly, then only the grains that have a ⟨110⟩ orientation will emerge.

Growing Ge from the bottom of a channel allows for selection of specific grain orientations, but the careful design of the channel is required in order to obtain single crystalline growth. The number of grains that emerge from the channel is related to the channel's aspect ratio. If the channel is shallow and wide (low aspect ratio), then the channel walls will not cause significant confinement in order to terminate the twinned ⟨100⟩ grains. In addition, multiple grains will nucleate with their fast growing grains oriented towards the channel exit, and therefore multiple grains may emerge. However, if the channel is narrow and deep (high aspect ratio), then there will be significant confinement of the misoriented grains. In addition, if the Si surface in which the Ge grows on is smaller, then a smaller number of Ge grains will nucleate and it is less probable for multiple grains to emerge from the channel. If the aspect ratio is too high, then no grains may emerge from a given channel. If the channel is very long and very narrow, then it may be that no grains nucleate such that their fast growing direction is oriented towards the channel opening. Therefore, some channel geometry optimization is necessary.

In order to model the channel, the channel structure is assumed to be a rectangular prism. The a-Si plane at the base of the channel has a height, h, and a width, w. The undercut etch defines the channel depth, d. This geometry is shown schematically in Figure 2.7.

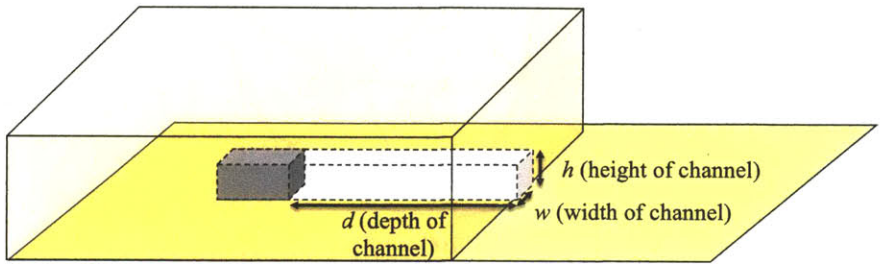

Figure 2.7 Schematic of a channel created by the 2D GCLG process. This is the idealized geometry that is used to model the expected number of grains that emerge from a given channel geometry. The nucleation area is the a-Si at the back of the channel. This schematic is taken from part (b) of Figure 2.4.

The goal of the model is to predict the number of grains, N_G, which will emerge from a given channel geometry. The channel geometry is determined entirely by its height, h, width, w, and depth, d. In reference [26], first the ratio of Ω_c/Ω_n is calculated. Here, Ω_c denotes the solid angle of the channel opening as seen from the center of the a-Si growth plane at the bottom of the channel. Ω_n denotes the solid angle bounded the four standard stereographic triangles surrounding a single ⟨110⟩ pole. Figure 2.8 shows an example of the standard stereographic projection of a cubic crystal. The shaded area denotes the four standard stereographic triangles surrounding a single ⟨110⟩ pole. Therefore, Ω_n represents the range of misorientation of a single ⟨110⟩ pole before a different specific ⟨110⟩ pole would be closer to the a-Si substrate normal direction.

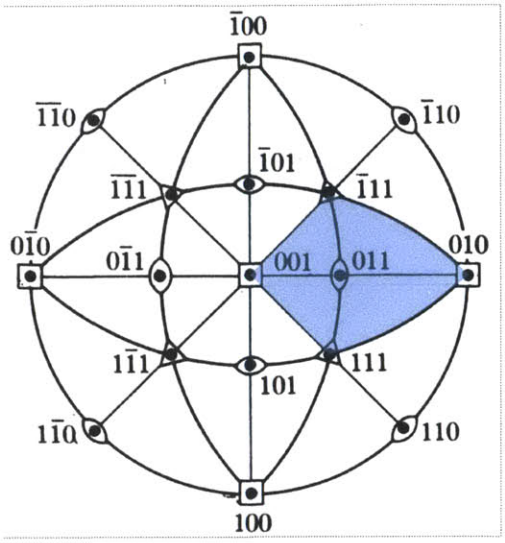

Figure 2.8 The standard stereographic projections of a cubic crystal. The squares at the intercepts represent the ⟨100⟩ poles, the ovals represent the ⟨110⟩ poles, and the triangles represent the ⟨111⟩ poles. An example of the four standard stereographic triangles surrounding a ⟨110⟩ pole has been shaded. The solid angle represented here is an indication of the range of misorientations of a single ⟨110⟩ pole before a different ⟨110⟩ pole is closer to the a-Si substrate normal direction.

The values of Ω_c and Ω_n are calculated by Equations (4) and (5).

$$\Omega_c = 4\sin^{-1}\left[\frac{hw}{\sqrt{(4d^2+w^2)(4d^2+h^2)}}\right] \tag{4}$$

$$\Omega_n = 2\pi \cdot \frac{4}{24} = \frac{\pi}{3} \tag{5}$$

The ratio of Ω_c/Ω_n indicates the probability of a randomly nucleated grain to have its ⟨110⟩ direction oriented such that it will intersect with the opening of the channel. This gives the likelihood of a randomly oriented grain exiting the channel within a sufficient growth time. It assumes that if a grain is oriented such that its ⟨110⟩ direction intersects with the channel wall, then the growth will terminate and will not emerge from the channel.

In order to complete the model, the probability of a grain emerging from a channel (Ω_c/Ω_n) is multiplied by the number of grains that nucleate on the a-Si plane. The number of

nucleated grains is calculated by dividing the area of the a-Si growth plane by the area of the base of a Ge grain on a-Si. The area of the a-Si growth plane is simply the product of the height and width of the channel and is calculated as hw. A_G denotes the average area of the base of a Ge grain on a-Si. Therefore, the total number of grains expected to emerge from a given channel (N_G) is given by Equation (6):

$$N_G = \frac{\Omega_c}{\Omega_n} \cdot \frac{hw}{A_G} \tag{6}$$

With the given model for N_G, the number of grains expected to emerge from a given channel geometry can now be predicted. Although this gives the number of grains expected to emerge from a channel, there are some practical limitations imposed on actual channel design. For example, if a channel is very short and wide ($w \gg h$), then a channel geometry could still be designed such that $N_G = 1$. However it is not necessarily realistic for this geometry to completely confine the growth. If the channel width is large, it is unlikely for any grain to be able to overtake all other competing grains before emerging from the channel. Therefore, in order to insure the model accuracy, it is ideal to have an equal channel height and width. This rough design constraint aides in allowing the fast growing ⟨110⟩ grains to overtake the slower growing grains. If the constraint that $h = w$ is imposed, then the equation for N_G can be reduced. By combining Equations (4), (5), and (6), and setting $h = w$, the expression in Equation (7) is obtained.

$$N_G = C \cdot h^2 \sin^{-1}\left(\frac{h^2}{4d^2 + h^2}\right) \tag{7}$$

$$C = \frac{12}{\pi A_G} \tag{8}$$

Equation (7) is the model for the number of grains expected to emerge from a channel with a square cross-section. The coefficient, C, in front of the expression is defined in Equation

(8) and is a constant; given the assumption that A_G is a constant. Upon inspection of Equation (7), it is evident that there are two ways to reduce N_G. To reduce the number of grains emerging from the square channel, the channel can either be designed to have a smaller cross section (reduce h), or be deeper (increase d). When h is reduced, the h^2 term in front of the expression dominates and N_G is rapidly reduced. However, when the d term is increased, the inverse sine term asymptotically reduces to zero. Therefore, N_G is more sensitive to the cross section dimension and reducing this more rapidly provides the opportunity for single crystal growth. The approach of reducing h instead of increasing d also has an additional benefit. The channel depth, d, effectively defines the amount of Ge growth that must occur before the Ge emerges from the channel as a single crystal. Therefore, a large d would yield long growth times. Since Ge deposition occurs at low temperature (450 °C), the growth rate is slow and minimizing d can save hours of growth time. Putting this all together, in order to obtain single crystalline growth, it is ideal to have (1) a low N_G, (2) a square cross-section ($h = w$), and (3) a small d.

2.2 Novel Design to Utilize 2D GCLG to Fill Trenches

The technique of two-dimensional, geometrically confined lateral growth is a method that was employed to fabricate high quality Ge for the active region in optoelectronic devices. In order to fabricate optimized devices, it is important for the device designer to have control over the device geometry and dimensions. The most basic geometry for the active region in Ge optoelectronic devices is a rectangular prism. There are two main ways in which this geometry is typically fabricated. In the first technique, a Ge thin film is deposited, then patterned using lithography, and then the Ge is etched into the desired structure [12], [32–34]. The second method that is commonly used, is selective growth of Ge in oxide trenches [11], [35–37]. In this

method, the Si wafer substrate is oxidized so that there is an oxide film covering the crystalline Si. Then, the oxide is patterned and etched until trenches are opened, exposing the Si substrate below, in lithographically defined regions. Then, Ge growth occurs within the trench, taking advantage of selective deposition, with Ge only depositing on the Si wafer, and not on the oxide. Therefore, the patterning step occurs in the oxide, instead of the Ge directly. Both techniques allow for geometry control since the bounds of the Ge are controlled lithographically. In addition, both techniques can be used to produce a large range in active region sizes. The thickness of the active area is limited by the Ge growth thickness, but the length and the width of the Ge regions are only limited by lithography. In order for the 2D GCLG technique to be useful for application in a Ge optoelectronic device, the method must be able to create Ge that fills a lithographically defined region. If the device is to be waveguide integrated, which is the ultimate goal, then the ideal shape is a rectangular prism that is on the order of 0.5 to 1 micron in width, similar height, and tens of microns in length. There may be some variations of these dimensions and the exact geometry, but this is a generalized goal.

The technique of two-dimensional, geometrically confined lateral growth is a method that can be used to obtain single crystalline Ge on an amorphous substrate. However, it is a very different crystalline seed than a typical crystalline Si substrate. Although the emerging Ge grain is crystalline, it has a very irregular geometry and does not serve as a typical substrate to seed planar epitaxial growth. Typically, substrates for epitaxial growth are two dimensional and planar. This allows for controlled, uniform epitaxy with regards to growth rate, thickness control, and large area uniformity. However, the crystallite that emerges from the 2D GCLG technique is a small 3D structure. An example of the Ge crystallite that emerges from a channel is shown in Figure 2.5. It still works as a seed for crystalline growth, but the exact shape of the crystallite is

unpredictable. In addition, it is too small to seed epitaxial growth to fill a trench, that is tens of microns long, within a reasonable growth time. In order to use the 2D GCLG growth technique for the active area of Ge optoelectronic devices, a novel new technique was designed in order to fill lithographically defined trenches.

2.2.1 Design

The crystal quality of the Ge within the trench is important. The more crystalline the Ge is in the trench, the better the optical and electrical properties will be. Since grain boundaries can serve as optical scattering points and carrier recombination centers, ideally the Ge in the trench should be single crystalline. In principle, this could be achieved with a single 2D GCLG crystallite. A single Ge crystallite is all that is needed to seed a crystalline Ge trench. However, the Ge growth rates at low temperature are slow, with growth rates observed at around 75 – 100 nm/hr, depending on growth pressure. Therefore, it is impractical to grow a single crystallite to tens of microns in order to fill the trench because growth times would be several days to weeks long. In addition, this would yield a very large overgrowth above the trench, which would be difficult to planarize. Instead, multiple 2D GCLG channels were used to seed multiple Ge crystallites. With this design, the Ge in the trench will not be single crystalline, however the grains that fill the trench will be large (>1 μm) and the device designer has complete control of the number of grains, as well as the placement of the individual grains.

Before the device itself was designed, first the individual 2D GCLG channels were designed. Since the Ge growth is seeded from these channels, from now on the 2D GCLG channels will be referred to as seeds. The geometry of the individual seed was first designed in order to yield single crystal Ge crystallites. In order to determine the required channel

dimensions, Equation (6) was plotted for range of different channel geometries. This plot is shown in Figure 2.9. Here, it is assumed that the width of the seed is held constant at 100 nm. The purpose was to determine the effect of varying the seed height and depth. The height of the seed is determined by the a-Si film thickness, the width is defined by lithography and dry-etching, and the depth is determined by the TMAH undercut etch step.

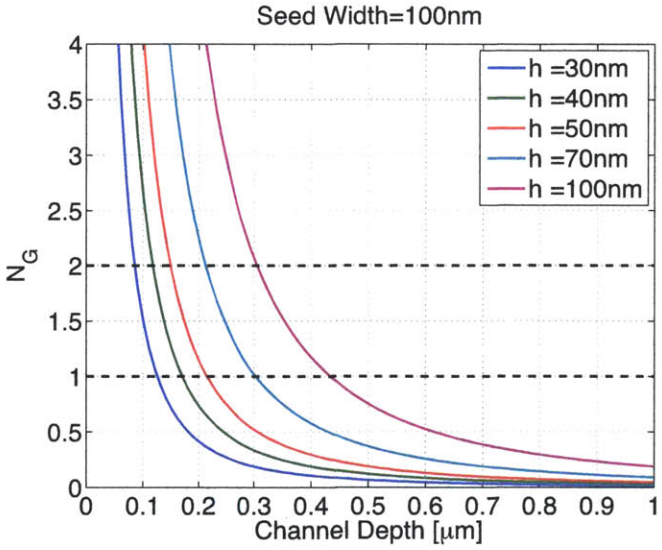

Figure 2.9 The number of grains expected to emerge from a 2D GCLC growth seed (N_G) for a variation of channel heights (h). This figure assumes a channel width of 100 nm. In order to obtain a single crystalline Ge crystallite emerging from the channel, the target N_G is between 1 and 2.

In Figure 2.9, two dotted lines are drawn, one at $N_G = 1$ and one at $N_G = 2$. These should be considered the upper and lower bounds for an acceptable design. If $N_G < 1$, then it is predicted that a Ge grain will not emerge from the channel every time. For example, if $N_G = 0.5$, then it is predicted that a Ge grain will only emerge from 50% of the channels with this given geometry. If $N_G = 2$, then it is predicted that 2 Ge grains will emerge from each channel. The ideal channel design is one in which the Ge emerges from the channel every time, however only one grain emerges. Therefore, the target range of N_G is between 1 and 2. Within this range, it is predicted

that a Ge grain will emerge from the grain each time. In addition, when N_G is less than 2, the model predicts that no more than 2 grains will emerge from each channel. Therefore, the channel should be designed such that the target range for N_G is between 1 and 2.

Each individual line in Figure 2.9 shows the same general trend. N_G decreases as the channel depth increases. This makes intuitive sense because as a channel gets deeper, the aspect ratio gets larger. This means that solid angle of the opening decreases. Therefore, the likelihood of a randomly nucleated grain being oriented such that the ⟨110⟩ direction is pointing towards the channel opening decreases. A second observation is that N_G increases as the height of the channel, h, increases. The explanation of this follows the same logic as before. The larger h yields a larger channel opening, and hence a larger solid angle of the channel opening. The larger solid angle leads to the large chance that a randomly nucleated grain is oriented such that its ⟨110⟩ direction intercepts the channel opening.

In order to narrow in on exact channel dimensions, first the channel height was selected. The main factor limits the channel height is hydrogen incorporation into the a-Si film. The channel height is determined by the a-Si film thickness, which is deposited by plasma-enhanced chemical vapor deposition (PECVD). A-Si that is deposited by PECVD is subjected to significant H incorporation into the Si during deposition [38], [39]. The problem with hydrogen incorporation is that when heated, the H that is trapped within the film may coalesce and form bubbles. The hydrogen bubbles create microvoids which eventually burst, destroying the planar Si film [40]. Therefore, the a-Si film must be kept thin. If it is too thick, then significant amounts of H become trapped within the film, which cause bubbles and hence destroy the film. In order to eliminate this risk, the a-Si layer was kept thin at 50 nm. Once the channel height is fixed, then the channel width and depth are chosen. The ideal width, for a fixed 50 nm channel height, is 50

nm. However, in reality, the width is limited by the resolution of the lithography. Once the width is determined, then the design rules are applied and Figure 2.10 is consulted in order to determine the depth that will yield an N_G value between 1 and 2. For channel widths of 100, 200, and 300 nm, channel depths of 175, 350, and 550 nm were targeted, respectively.

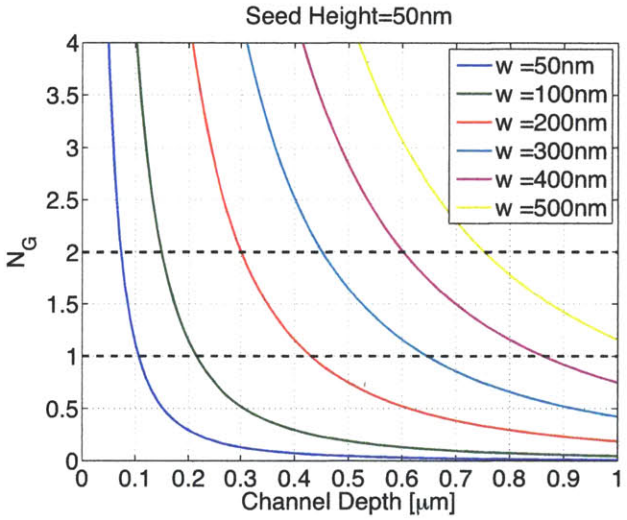

Figure 2.10 2D GCLG design guide. This figure assumes a fixed channel height of 50 nm. The channel width is increased from 50 nm, which is the ideal value, to 500 nm, which is easy to fabricate in a real system. The purpose of this figure is to determine the channel length that yields an N_G value between 1 and 2.

Once the individual channels are designed and optimized, then it is necessary to design structure that will fill a trench. As mentioned, for the 2D GCLG technique to be useful for a real optoelectronic device, it must be capable of filling a rectangular prism shaped trench that has a width of around 500 nm, a height of 200 – 500 nm, and a length of tens of microns. These dimensions are not arbitrary, but they are the approximate dimensions for waveguide integrated Ge optoelectronic devices, such as photodetectors and electro-absorption modulators. The cross-section of approximately 500 nm by 200 nm is because this yields a single-mode Ge waveguide for light with a wavelength of 1.55 μm.

The design approach utilizes multiple 2D GCLG structures to seed multiple Ge grains. The use of more than one seed means that the Ge within the trench will not be single crystalline. The overall trench filling design is shown schematically in Figure 2.11. Multiple 2D GCLG structures, now to be called seeds, are arrayed with their channels opening into a common trench. The goal of the device is to nucleate multiple Ge crystallites within the trench, and then grow epitaxially until the crystallites coalesce and eventually fill the trench. This is creates a design tradeoff between duration of Ge deposition, and number of grains in the trench. A large pitch between seeds yields large grains, but also increases growth time.

This growth technique is very different than typical fabrication techniques. Typically, Ge optoelectronic devices are fabricated in a planar fashion. Either a planar film is deposited and then etched into the appropriate device geometry, or a trench is etched in SiO_2 and Ge is selectively deposited on a Si substrate. In both conventional techniques, the growth time is set by the desired film thickness. However, in the design shown in Figure 2.11, the growth time is set by the distance between seeds, and the height of the trench. It is not sufficient to grow Ge thick enough to emerge from the channel, but the Ge has to grow vertically enough to fill the trench, and laterally enough to coalesce with the neighboring seeds.

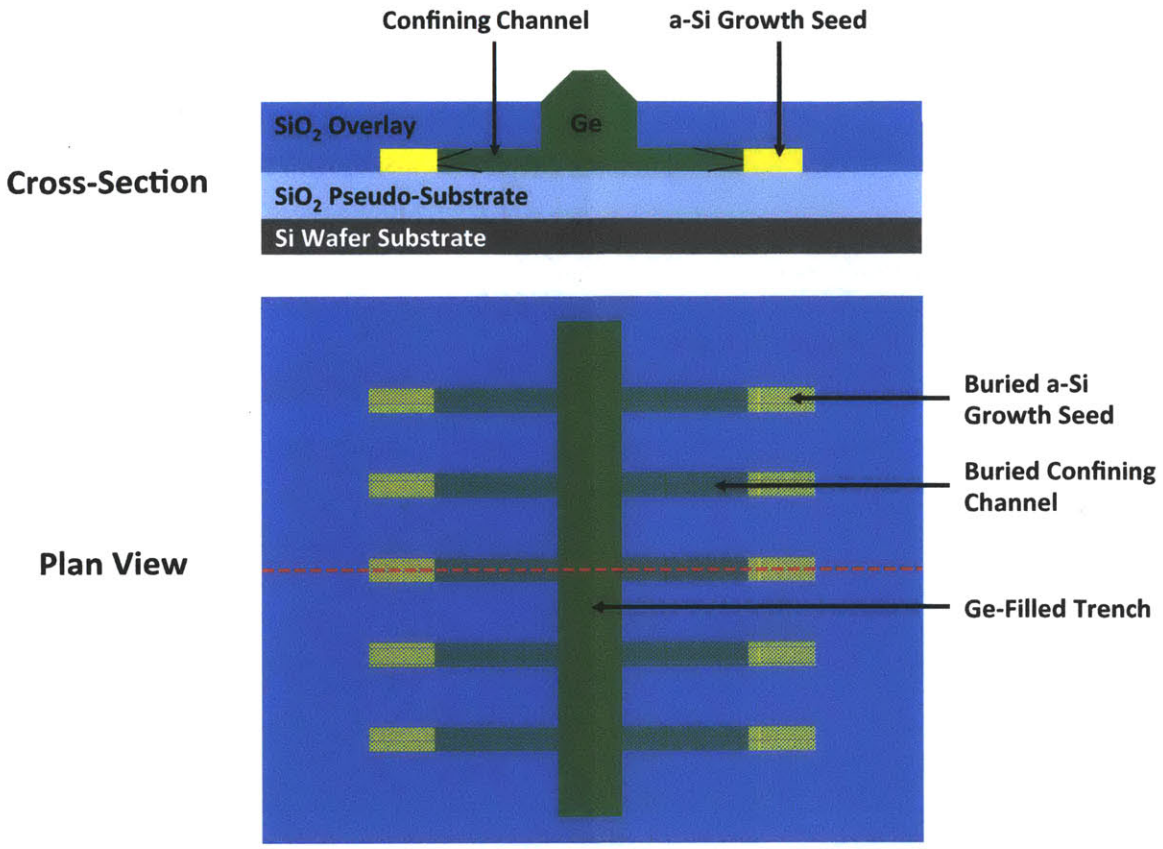

Figure 2.11 Design of a Ge-filled trench with large grains. The red dotted line in the plan view image indicates the plane through which the cross-section view is shown. This design has aligned seeds in order to minimize growth time and for ease of fabrication.

The two designs presented, one in Figure 2.11 and the other in Figure 2.12, show a similar design to achieve the same goal. Both approaches utilize an array of 2D GCLG seeds to nucleate Ge crystallites within the trench, and then grow epitaxially until the crystallites coalesce, filling the trench. However, the approach in Figure 2.11 has aligned seeds while the design shown in Figure 2.12 has staggered seeds. The approach with aligned seeds minimizes Ge growth time, and allows for ease of fabrication. The growth time is minimized because the seeds are closer to each other and therefore less total growth is required before the Ge grains coalesce.

It is also easier to fabricate because a lower tolerance is required in aligning mask levels, while performing the lithography. The fabrication technique will be described in Section 2.2.2 below.

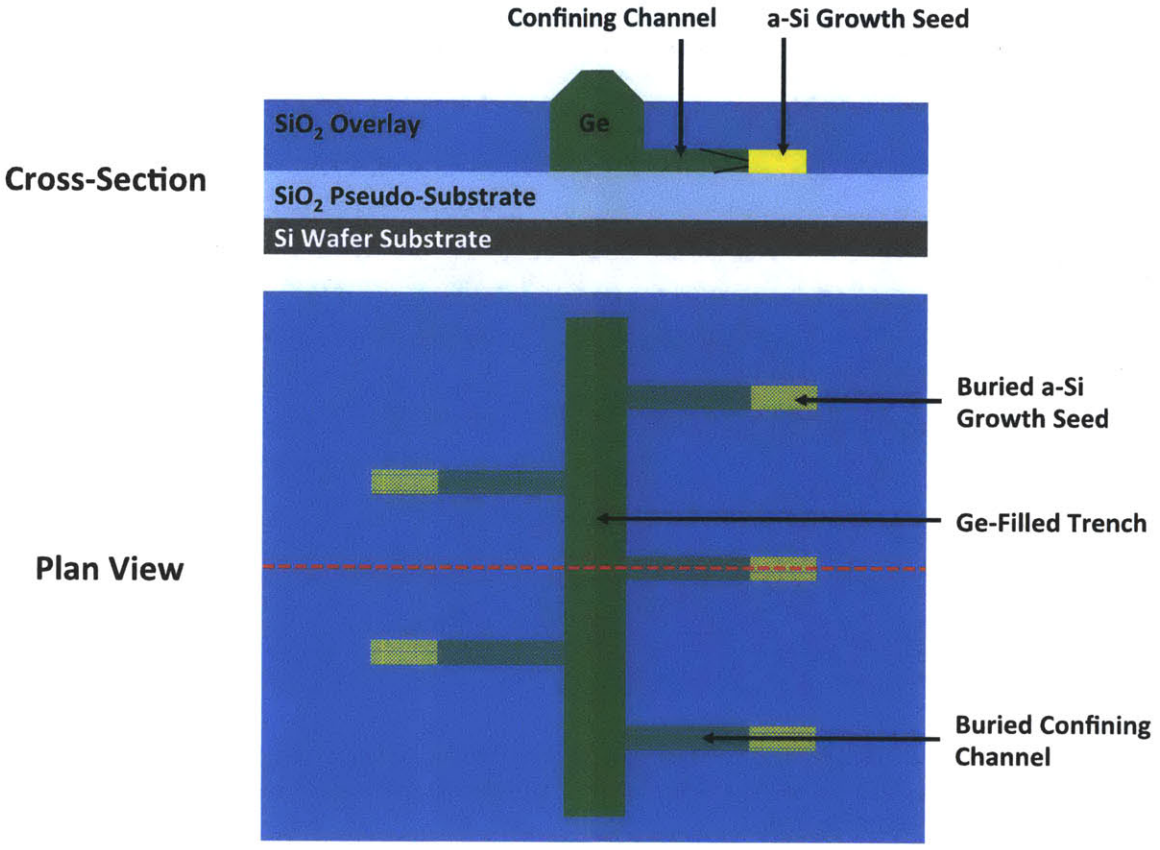

Figure 2.12 Modified design that reduces the number of grain boundaries, and increases the grain size. The red dotted line in the plan view image indicates the plane through which the cross-section view is shown. This design has staggered seeds in order to minimize grain boundaries.

While the aligned seed approach is easier to fabricate, and allows for shorter Ge growth times, the staggered seed approach minimizes the number of grains in the trench, and therefore maximizes the overall crystal quality within the trench. With the ideal seed and channel design, it can be assumed that one single grain emerges from each seed. Therefore, since the aligned seed design has twice as many seeds as the staggered seed design, it will also have twice as many grains. This yields twice as many grain boundaries. Grain boundaries can act as carrier

recombination sites, optical and electrical scattering sites, and create high resistance, current blocking barriers [14], [15]. Therefore, less grain boundaries are beneficial for device quality. These two devices both offer a technique to fill lithographically defined trenches with large grain germanium. The exact number of grains, their physical location, and their size, can be predicted by the number of seeds and the spacing between them. However, design tradeoffs exist between ease of fabrication, crystal quality, and Ge growth time. A larger seed pitch yields larger Ge grains at coalescence, and therefore enhanced overall crystal quality within the trench and better device properties. However, it also means a longer growth time, which can be very significant at low temperatures. For example, depending on germane overpressure, growth times can be between 18 – 20 hours for approximately 1.5 μm of growth, yielding a growth rate of approximately 75 nm/hr. Therefore, a seed pitch of several microns or greater is not realistic for a practical application.

2.2.2 Fabrication Method

An overview of the fabrication method is shown schematically in Figure 2.13. All steps are CMOS compatible, and occur at or below 450 °C. To start, a layer of SiO_2 is deposited on a Si wafer using a plasma-enhanced CVD (PECVD) deposition technique. The thickness of the SiO_2 is unimportant, because it merely serves as the amorphous pseudo-substrate. The purpose of this layer is to emulate the amorphous dielectric in the back end of the interconnect stack in a CMOS process flow. Once the oxide pseudo-substrate is formed, a thin layer of a-Si is deposited using PECVD, as indicated in (a) of Figure 2.13. The thickness of this layer determines the height of the 2D GCLG channel, and is therefore carefully controlled. For these devices, this layer was 50 nm thick.

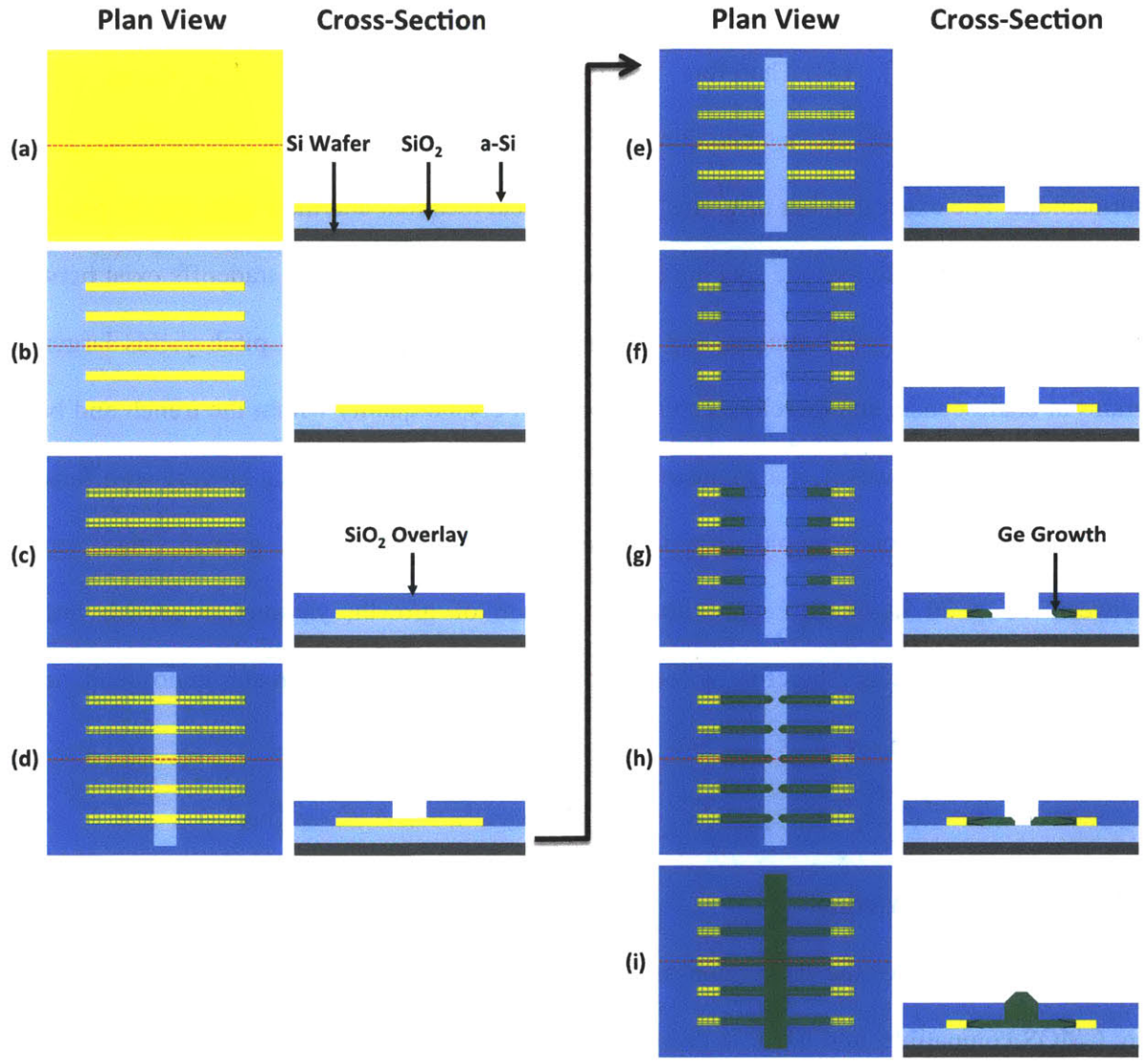

Figure 2.13 Process flow for utilizing 2D GCLG for trench filling. Each step shows a plan view perspective on left. The right side shows a cross-section view through the red dotted line. Part (a) shows a thin blanket a-Si film on an oxide film on a Si wafer substrate. Next, the a-Si is patterned into thin lines, shown by (b). In (c), an SiO_2 overlay is deposited on top of the a-Si lines. In (d) a trench is opened up in the SiO_2 overlay by RIE dry etching, exposing the a-Si. In (e), the exposed a-Si is dry etched. In (f), an undercut etch of the a-Si is performed with TMAH. In (g), Ge deposition begins in UHVCVD. In (h), the Ge emerges from the channels as a single crystal. Finally, in (i), the Ge crystallites coalesce and growth continues until the trench has been filled.

Next, as indicated in (b) of Figure 2.13, the a-Si is patterned into thin lines. The widths of these lines determine the width of the 2D GCLG channels and were patterned to be between 100

and 300 nm. The stepper that was used to pattern these lines only has a reliable resolution limit of around 1 μm, and therefore a double exposure technique with a sub-micron offset was employed to reduce the line widths. After the double exposure of the photoresist, the remaining photoresist lines were further thinned by dry etching in an oxygen plasma. With this approach, it was possible to pattern lines as narrow as 100 nm with a stepper that is only capable of exposing 1 μm features.

After the a-Si lines are patterned, a SiO_2 overlay was deposited using PECVD, as shown in (c). The thickness of this layer determines the height of the trench that is to be filled. This thickness was varied between 200 and 500 nm. Next, the trench itself is opened up. This is done by first selectively dry etching through the oxide, as indicated in (d), and then selectively etching the a-Si, as indicated in (e). At this point, the a-Si lines are intersecting the trench sidewalls. Next, the growth channel is defined. This is done with a tetramethylammonium hydroxide (TMAH) wet etch at 80 °C. TMAH selectively etches Si, and not SiO_2 and therefore undercut etches the Si lines, while leaving the SiO_2 overlay intact, as shown by (f). This etch defines the channel depth, and therefore was timed in order to finalize the dimensions of the growth channel, effectively defining N_G of the 2D GCLG structure. Finally, the wafers are ready for Ge growth.

Immediately before Ge growth, the wafers are cleaned in a standard RCA clean, ending with a quick dip in hydrofluoric acid (HF). This is in order to passivate the exposed Si by occupying the dangling bonds in Si with H, inhibiting the formation of a native oxide. Then, the wafers are loaded into an ultra high vacuum chemical vapor deposition (UHVCVD) reactor. The a-Si and SiO_2 films deposited by PECVD have hydrogen incorporation into the films. To remove the H, the wafers were annealed at 450 °C in the UHVCVD reactor tube at a base pressure of 1.8×10^{-8} mBar for two hours to allow the hydrogen to sufficiently outgas from the films. Finally,

Ge growth was initiated with germane (GeH_4) flow at 7.5 sccm at a chamber temperature of 450 °C and a growth pressure of 3.4×10^{-3} mBar. No carrier gasses were used. The Ge selectively deposits on the a-Si, and not on the SiO_2, as indicated in (g). Eventually, the Ge emerges from the channels as a single crystal crystallite, as shown by (h). These crystallites are used to seed epitaxial growth until, eventually, the neighboring crystallites coalesce and eventually fill the trench, as shown in (i). The total growth time is 18 hours. Once the trench is eventually filled, the Ge within the trench is non-planar and highly faceted. If the Ge is to be used for an optoelectronic device where low-loss modal propagation is required, such as a modulator, the device must be chemically and mechanically polished (CMP) before further processing.

2.3 Experimental Results

The technique of arrayed seeds was shown to be a viable technique to fill a lithographically defined oxide trench with large grain Ge. There is no fundamental limit to the length of the trench that can be filled with this technique, as it is the seed pitch that determines the growth time, not the overall trench length. An example of a trench that has been filled using this technique is shown in Figure 2.14. The plan view SEM image shown in this figure shows a trench that is greater than 50 μm long, and is completely filled with Ge. The vertical lines in the image are caused by the a-Si seed lines. The SiO_2 overlay deposition process is conformal and therefore expands the shape of the a-Si seed lines. The vertical lines are SiO_2, but indicate the location of the growth seeds below.

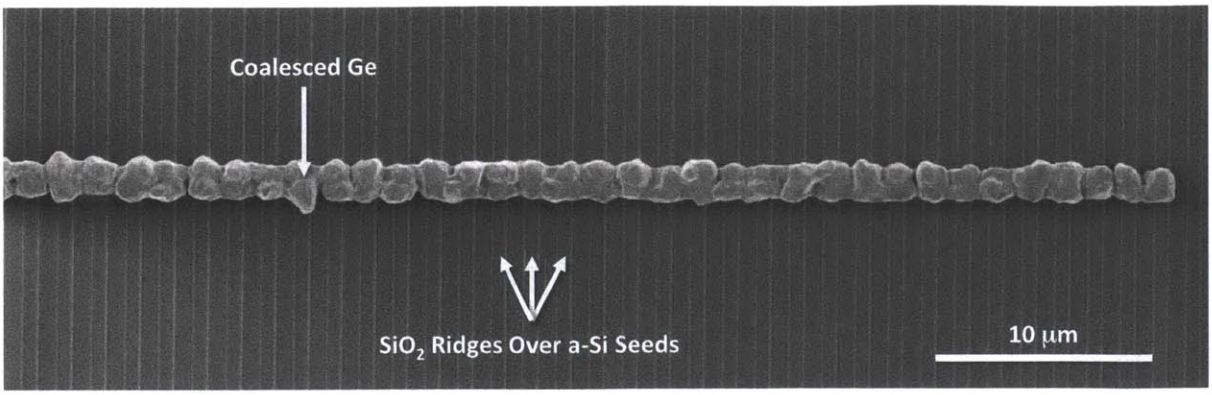

Figure 2.14 Plan-view SEM image of trench filling after Ge deposition. The vertical lines are from the a-Si seeds. Ge deposition occurred at 450 °C with a chamber pressure of 3.4 x 10^{-3} mBar for 18 hours.

The Ge that has emerged from the trench is clearly non-planar. However, the Ge crystallites have coalesced with their neighboring grains, and therefore the trench has been successfully filled. There is significant vertical overgrowth of Ge out of the trench. This is because the pitch between seeds is 1.5 μm and the trench height is only 0.5 μm in this device. The growth process continues until the Ge grains coalesce with the neighboring grains, and therefore significant trench overgrowth occurs before the lateral growth is large enough for coalescence.

An additional image of a trench filled using the aligned seed device is shown in Figure 2.15. The places at which the grains coalesced are marked with a dotted line. The zoomed in image here shows the clear coalescence of the grains, to fully fill the trenches. The top of the Ge is very non-planar. However, the topographical features of the Ge are located above the surface of the trench. If these devices are to be used in as the active region in actual optoelectronic devices, then a CMP step would be required in order to planarize the Ge.

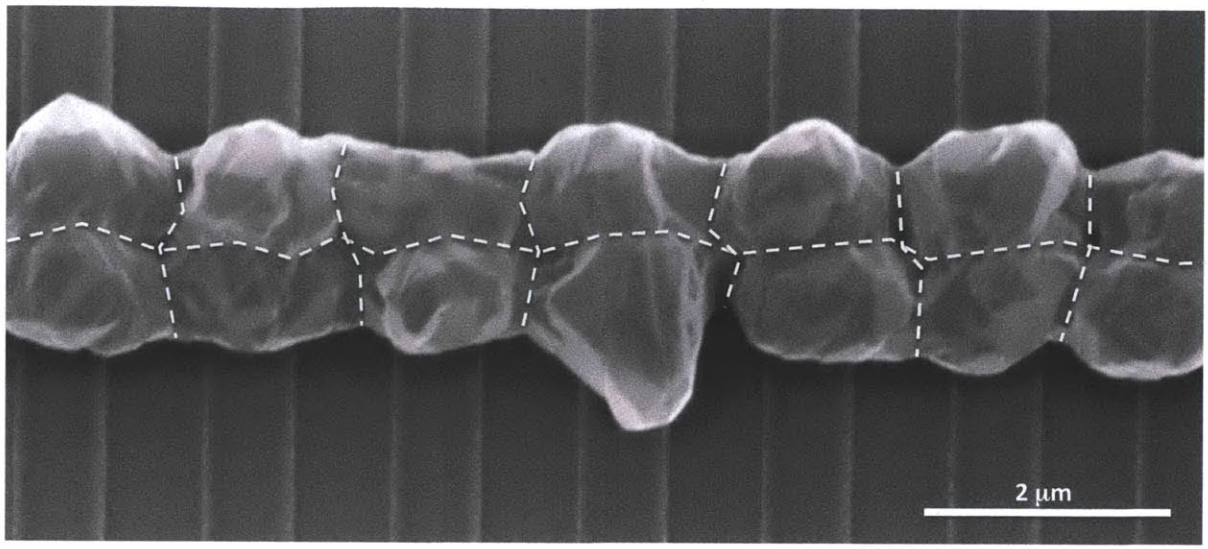

Figure 2.15 Plan view SEM image of trenches filled using the aligned seed technique. Approximate grain boundary lines have been drawn in. The dotted lines indicate the place at which the grains coalesced.

The aligned seed design and the staggered seed design were both fabricated. However, the staggered seeds proved more difficult to fabricate, given the limitations of the lithography equipment that was available. In the aligned seed design, as long as the trench intersects the seeds, the channel will open into the trench, effectively seeding Ge growth. However, with the aligned design, accurate alignment between the trench and the seed mask levels must be achieved. An example of a fabricated staggered seed design is shown in Figure 2.16. Here, it is clearly seen that there was a misalignment between the trench and the seed mask levels. The trench is offset vertically by over a micron from its target position. This mask misalignment made it impossible to verify whether the staggered seed design works as well as the aligned seed design. With better lithography equipment, it is hypothesized that the staggered seed design will still work, as long as the Ge growth step is long enough. This hypothesis is further supported by the amount of lateral growth shown in the staggered seed design. Since only half of the intended seeds intersected the trench, the effective pitch between seeds was doubled. Therefore, the Ge

grains needed to grow twice as large in order to coalesce. However, as long as the lateral growth is half of the effective seed pitch, the grains would have coalesced if both sets of seeds were exposed to the trench. In Figure 2.16, the grains occupy approximately half of the lateral space between exposed seeds. Therefore, if the bottom seeds also intersected the trench, then the lateral Ge grain density would have doubled, and therefore the grains would likely coalesce, thereby filling the trench.

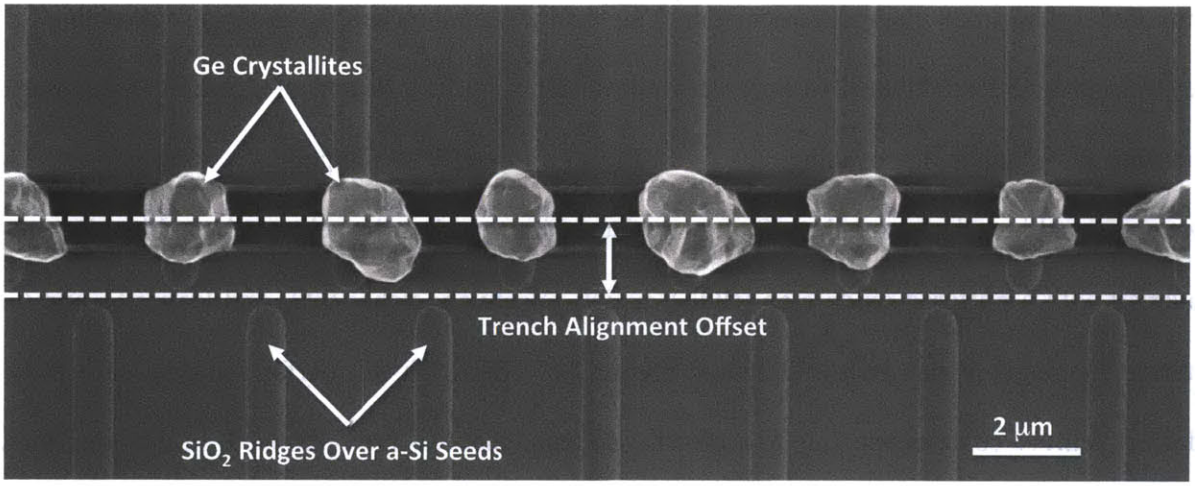

Figure 2.16 Plan view SEM of the staggered seed design. These devices suffered from a misalignment between the trench and the seed mask levels. There is a clear vertical misalignment, which meant that both arrays of seeds did not intersect the trench opening. Only the top seeds intersected the trench opening and therefore contributed to seeding the trench with Ge crystallites.

2.4 Problems With Method

The staggered and aligned seed designs are a first iteration proof of concept, proving that it is possible to fill a lithographically defined trench with large grain Ge, on an amorphous substrate. However, there are still large problems with the design. The first problem is due to the misalignment of the trench and the seeds in the staggered seed design. This problem can be

remedied with a trench mask level offset, or with the use of more modern equipment. The alignment in a stepper tool should be much less than 1 μm, and therefore this is not necessarily a design problem. However, one process step clearly limits the reproducibility of the design. That step is the TMAH undercut etch. The undercut etch is a timed wet etch. TMAH is selective, and will etch Si while leaving SiO_2 mainly un-etched. However, the only way to define the channel depth, is by timing the undercut etch such that the correct amount of Si is removed.

The problem of the TMAH undercut etch arises due to the geometrical limitations imposed by the small channel dimensions. The channel being etched has a cross-section that can be as small as 50×100 nm^2. At these small dimensions, mass transport of etchant can become a limiting factor when determining etch rates. In order to verify this hypothesis, the oxide overlay was removed in order to see the a-Si/Ge interface. This was done by selectively etching the SiO_2 overlay with a buffered oxide etch (BOE), which selectively etches SiO_2 and leaves Si and Ge intact. The result of this etch is shown in Figure 2.17. This etch was performed on a device that did not achieve coalescence, due to a growth sequence that was not optimized. In this plan view SEM image, the interface between the a-Si and the Ge is clearly shown, and occurs at the leading etch of the Si seed. This boundary shows the depth of the undercut etch that was achieved by the TMAH etch. The sidewalls of the trench have receded and are now bowed due to the SiO_2 being etched from the sidewalls within the trench, at the same time as it was etched from the top surface.

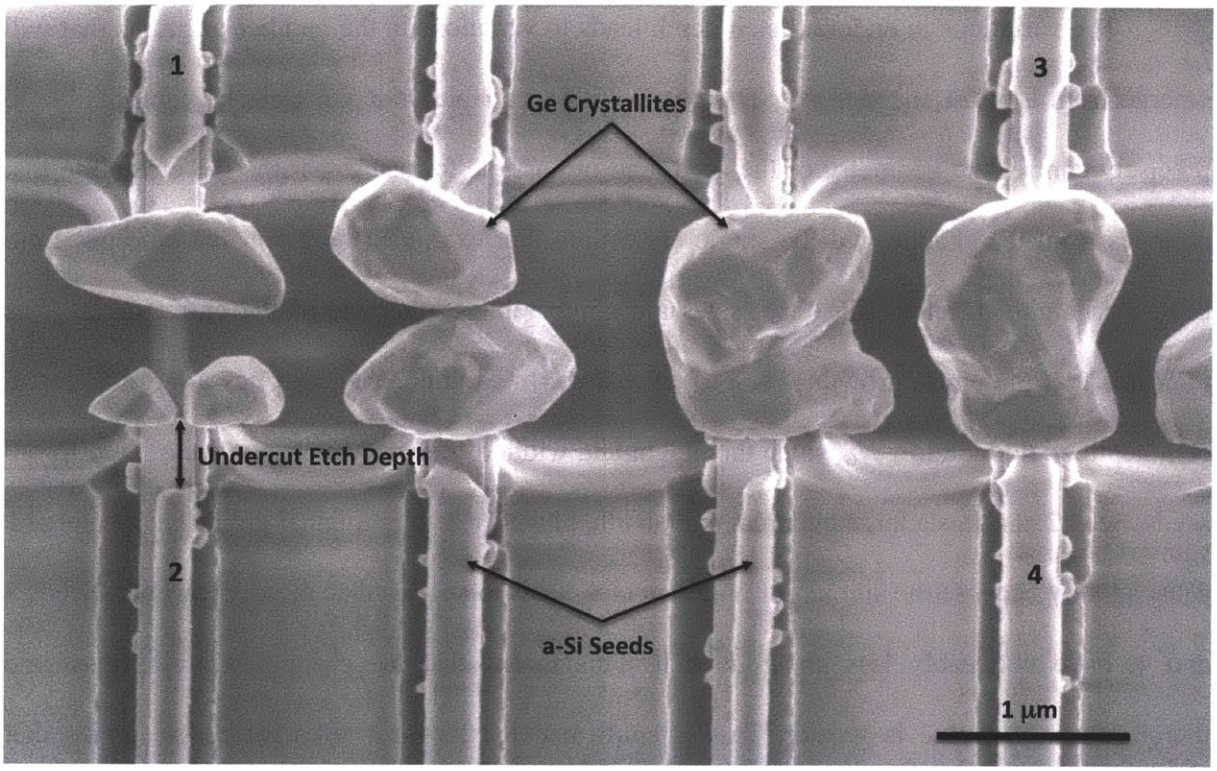

Figure 2.17 A plan view SEM image of a failed growth that has the top oxide removed by BOE. There is a clear non-uniformity in undercut etch of the a-Si. The seeds with a longer growth channel (larger undercut etch) show smaller grains with larger facets. Smaller undercut etches show larger grains with smaller facets. Large facets are a sign of single crystallinity.

The problem with the TMAH undercut etch is due to the non-uniformity in undercut etch rates. From Figure 2.17, it is clearly seen that the total undercut etch depth is not constant for all of the seeds. For example, the two seeds on the left, labeled 1 and 2, have larger undercut etch depths than the two seeds on the right, labeled 3 and 4. It is difficult to quantify the absolute length of the undercut etch, due to the Ge overgrowth over the trench sidewalls. Therefore, the edge of the channel is hidden under the Ge overgrowth. The only way to measure undercut etch depth precisely, would be to use a focused ion beam (FIB) to mill a cross-section in each growth, and then measure each seed individually.

The variations in undercut etch depth leads to variations in the number of grains that will emerge from a given channel, N_G. Therefore, with a variable undercut etch, it is impossible to precisely control the geometry of the seed, and impossible to insure that the Ge crystallite that emerges from the channel will be single crystalline. For example, the grain that emerges from Seed 1 is highly faceted. The facet sizes are comparable to the size of the entire grain. This indicates that the entire growth is single crystalline. In addition, this seed has a large undercut etch depth, yielding a low N_G, since the height and width of the growth channels are fixed for all seeds. Seed 2 also has a large undercut etch depth, however, two individual grains have emerged. This is likely because the amount of undercut etch leads to an N_G value of ~2. However, Seeds 3 and 4, have noticeably less undercut etch. This means that the channel depth is less, and therefore the N_G value is greater. This is confirmed by the facet sizes in the Ge grains that emerged from these channels. The facets are much smaller than the total grain, indicating that the growth is likely poly-crystalline and multiple grains emerged from the channel.

In addition to the lack of control of the total undercut etch depth, the remaining a-Si seed geometry after etching is also variable. Seed 2 shows a straight undercut etch that yielded an a-Si seed front parallel to the channel opening. This is exactly the geometry that was assumed in the model for predicting the number of grains that emerge from a 2D GCLG channel. However, the exposed front edge of Seed 1 and Seed 3 is clearly pointed. This means that the etch rate was not equal along the width of the seed, and therefore the undercut etch was not uniform, even within a single channel. This may be caused by capillary action, or mass transport limitations within such small channels. The front surface of the a-Si seed is the surface upon which Ge nucleates during growth. The area of this surface defines the number of seeds that are predicted to nucleate on the surface. The variation in seed geometry adds unpredictable variations to the model for N_G,

therefore making it impossible to design channels that consistently yield single crystal Ge emergence. The TMAH undercut etch is the step, that limits the reproducibility and reliability of this design. Therefore, the process flow needs to be redesigned to eliminate the timed undercut etch step.

2.5 Design Solution

The problem with the TMAH undercut etch step, is it relies on a timed etch in a small channel. This creates the potential for mass-transport limitations due to the small cross-sectional dimensions of the channels. The only reliable way to control the channel depth is through either a gas-phase selective etch of Si at high vacuum, or by using lithography. Processing in high vacuum increases the mean free path of particles, and therefore reduces the potential for mass-transport limitations within small channels. However, high vacuum selective Si etches are not a standard processing step, so the lithography approach was utilized.

The enhanced design incorporates a new material, Si_3N_4. The nitride is used to define the channel cross-sectional dimensions. In order to define the channel depth, the a-Si seed is lithographically patterned such that it is a set distance from the edge of the trench. This distance defines and sets the channel depth. The final device design, after fabrication, is shown in Figure 2.18. The a-Si lines are now patterned such that they are parallel to the trench, with the gap between them determining the channel depth. The channels are made from patterned silicon nitride lines, instead of a-Si lines. Hot phosphoric acid selectively etches silicon nitride without etching a-Si or SiO_2 [41], [42], and therefore the Si_3N_4 can be over-etched. This eliminates the need for a timed etch for channel depth formation.

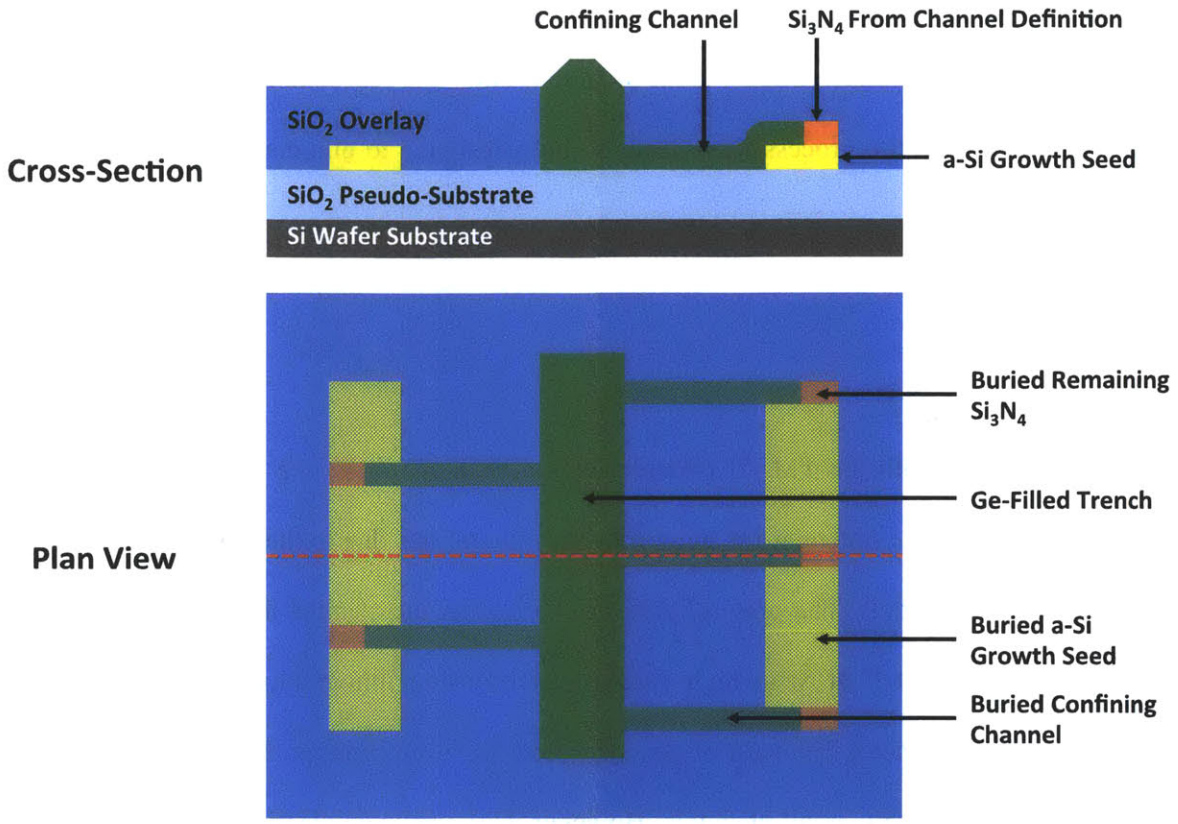

Figure 2.18 Schematic of the enhanced trench filling design utilizing silicon nitride for channel formation.

With the utilization of Si_3N_4 for channel formation, all critical channel dimensions are now controlled with either lithography or film thicknesses. Film thicknesses can be carefully controlled, and are reliable and repeatable. With the right equipment, patterning with lithography is very accurate and reproducible. Therefore, the yield of the new nitride design should be high, with reliable and reproducible trench filling.

A schematic of the process flow to fabricate the nitride-incorporated design is shown in Figure 2.19. The process for the enhanced design begins the same way as the original design. First, a SiO_2 film is deposited on a Si wafer using PECVD. This serves as the pseudo-substrate, emulating the wafer surface above the interconnect stack in a Si CMOS process flow. Next, an a-

Si film is deposited using PECVD. The thickness of the a-Si film corresponds with the designed channel height. Next, the a-Si is patterned into two parallel lines. The width of the lines is not important, but they are approximately the same length as the trench. The critical dimension is the space between the lines. The space is equal to two times the channel depth, plus the width of the trench. This is because these lines define the channel depth. The process, up to this point is depicted in part (a) of Figure 2.19.

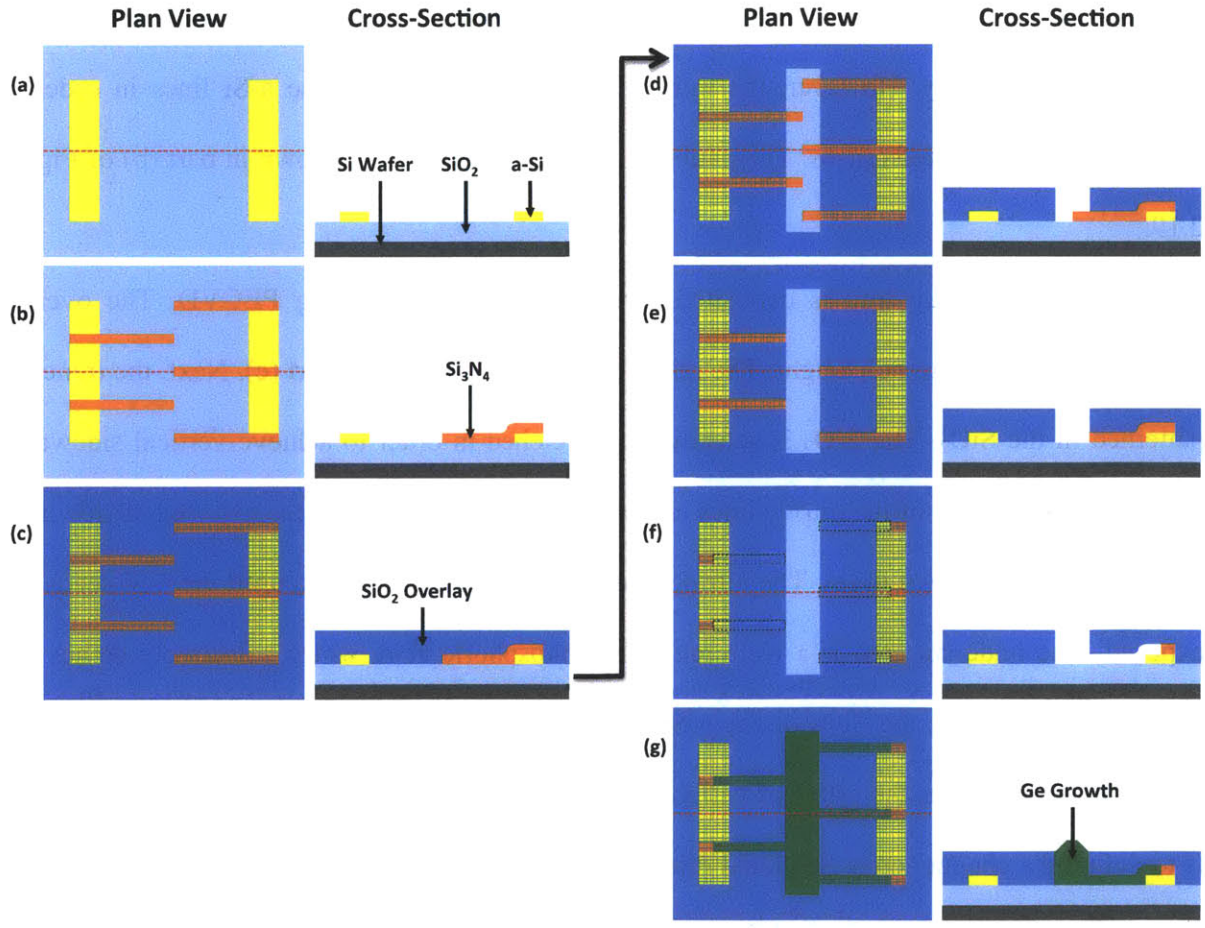

Figure 2.19 Process flow for fabricating the enhanced design with silicon nitride. The dotted red line in each plan view image represents the plane through which the cross-section view is shown.

After the a-Si lines are patterned, a Si_3N_4 film is deposited using PECVD. The thickness of the Si_3N_4 film is equivalent to the thickness of the a-Si film, and is designed to be the height of the 2D GCLG channel. The Si_3N_4 film is then patterned into thin lines. The width of the lines is equal to the width of the of the growth channels. The nitride lines are patterned in a staggered design in order to maximize grain size. This will require accurate alignment between mask levels, but this is easily achievable with deep ultraviolet (DUV) lithography equipment. One edge of the staggered nitride lines will be aligned with the center of the trench. The other end overlaps the a-Si seed lines. If perfect alignment between mask levels could be achieved, then the overlap would be unnecessary. However, the nitride is designed to overlap the a-Si lines in order to reduce the alignment tolerances. The patterning of the nitride lines is shown in part (b) of Figure 2.19.

A SiO_2 overlay film is deposited, over the nitride lines, using PECVD. The overlay thickness defines the trench height. The result of this step is shown in part (c). Next, the trench is patterned in the SiO_2. The SiO_2 is selectively dry etched in order to achieve vertical sidewalls. This exposes the ends of the nitride lines and is shown in part (d). Next, the exposed nitride lines are dry etched, as shown in (e). The nitride lines are then undercut etched until the a-Si seeds are exposed. The undercut etch is a wet etch in hot phosphoric acid, which selectively etches the Si_3N_4 without etching the SiO_2 or the a-Si. Due to the etch selectivity, the Si_3N_4 can be over-etched. This minimizes the effect of variable undercut etch rates. The purpose of the nitride lines is to define the channel cross-sectional dimensions. The total depth of the channel is determined by the space between the a-Si seed lines and the trench, which has been defined lithographically. Therefore, the Si_3N_4 is over-etched, ensuring the a-Si seed is exposed to the open trench. This is shown in part (f). Finally, the Ge is deposited in a UHVCVD reactor at 450 °C. The Ge will still

randomly nucleate on the a-Si seed at the end of the growth channel. With the correct channel design, a single crystal Ge crystallite will emerge from each channel, and therefore seed epitaxial growth within the channel. The crystallites will continue to grow epitaxially until the grains coalesce with the neighboring grains, effectively filling the trench, as shown in part (g).

The enhanced design for trench filling, which utilizes nitride to define the growth channels, eliminates the problems that were observed in the original design. It is an optimized design that uses the fundamentals of the 2D GCLG technique, and applies them to the goal of filling a trench with large grain Ge. This trench is lithographically defined, and all processing is compatible with Si CMOS back end integration. The design effectively creates a trench, filled with large grain Ge, which can be used as the active region in Ge-based optoelectronic devices.

Chapter 3: Evanescent Optical Power Coupling

For complete integration of photonics in the back-end, three significant barriers must be overcome: (1) how to fabricate high quality Ge for active regions of optoelectronic devices while adhering to back-end processing constraints, (2) how to couple optical power to these devices, and (3) how to metalize and electrically contact these devices. The first step, namely how to fabricate high quality Ge, has been demonstrated in Chapter 2. This chapter explores the next step and optimizes techniques to couple optical power to these devices.

One of the benefits of using optical interconnects is the low transmission loss and hence the potential for low energy consumption. Since this is a potential benefit that photonic interconnects possess over metal interconnects, it is imperative to optimize this design aspect. With modern fabrication capabilities, it is possible to create high index contrast waveguides with very low losses [43–46]. With waveguides already achieving very low propagation losses, it is critical to control losses associated with coupling from waveguides to devices.

There are two main ways in which light is coupled from waveguides to devices – butt-coupling and evanescent coupling [47]. Butt-coupling is when the waveguide directly abuts the active region of the devices. In this case, there is a sharp change in index of refraction, which can cause reflections and therefore coupling losses. The cross-sectional geometry of the waveguide and the device being coupled to can be designed such that the effective index of the guided mode is equal in each component, and therefore coupling losses are reduced. Although this technique may be effective in reducing coupling losses, it places strict limitations on the cross-sectional geometry of the devices in order to match the effective indices and the modal overlap of the

guided modes in each component. On the other hand, evanescent coupling between waveguides and devices can yield almost perfect coupling with coupling efficiencies between tapered waveguides greater than 99.5% [48].

Evanescent coupling is described by coupled mode theory and is a technique that can be used to transfer power between two waveguides, or a waveguide and a device, that are in close proximity [49]. Evanescent coupling occurs when the evanescent tail of a guided mode interacts with the core of a nearby waveguide. One drawback of evanescent coupling is that it occurs over a finite propagation distance. On the other hand, butt-coupling is abrupt and therefore inherently compact. Therefore, the use of evanescent coupling, in lieu of butt-coupling, means that additional space must be used for coupling. However, a coupler that utilizes evanescent coupling can be designed such that it achieves nearly perfect coupling efficiencies with minimal reflections or scattering losses. This means that when comparing butt-coupling to evanescent coupling, there is a tradeoff between size and coupling efficiency.

3.1 Vertical Couplers

The approach of integrating photonic interconnects into the back end is beneficial because it allows for dense 3D integration. With this approach, the photonic devices are no longer limited to the substrate and therefore multiple photonic planes can be realized. This eases the drive for small device footprint since substrate real estate is no longer as valuable for photonic devices. Although the devices can be fabricated on multiple planes, there needs to be a technique to couple light from one photonic plane to another. Therefore, this requires a vertical coupling technique. One way to vertically couple light between waveguides on different planes is with overlapping tapers.

3.1.1 Background and Design

A design for vertical coupling between waveguides has been presented by Sun et al. in [48]. In this method, two waveguides are oppositely tapered and separated by a layer of oxide, which serves as a universal cladding for each Si waveguide. Along the propagation direction, the lower waveguide tapers linearly to a narrow tip, while the upper waveguide tapers linearly from a narrow tip to a single-mode waveguide. This design is showed schematically in Figure 3.1.

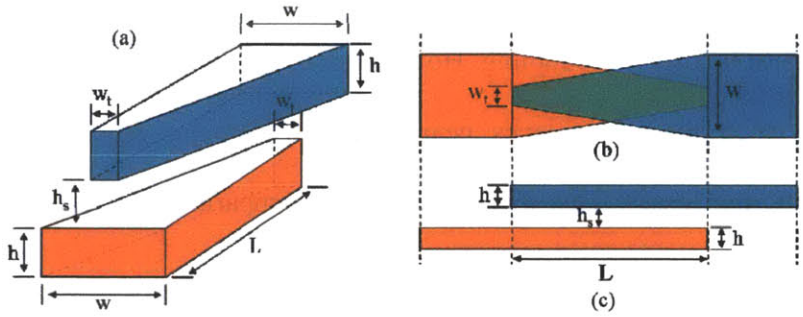

Figure 3.1 Schematic of the vertical coupler design. Three views are shown in which (a) is a 3D view, (b) is a top-down view, and (c) is a side view. For this study, the separation height (h_s) and the wavegide height (h) were both kept constant at 200 nm. The taper overlap length is given by L, and the taper tip width is given as w_t. In order to maintain single mode operation, the waveguide width, w, was kept at 500 nm. The light is injected into the lower waveguide and then coupled out of the upper waveguide. From [48].

The tapered vertical coupler takes advantage of evanescent coupling from the lower to the upper waveguide. The tapers serve as a way to vary the effective index along the propagation direction of the waveguide so that at one point along the overlapping taper, the effective index in each waveguide will be equal. As the lower waveguide tapers to a narrow tip, its effective index continues to decrease. This is because the mode expands and therefore more of the mode is exposed to the low-index cladding. As the mode expands, the confinement factor in the lower

waveguide decreases, and more of the mode begins to interact with the upper waveguide. While the effective index of the lower waveguide decreases along the propagation direction, the opposite occurs within the upper waveguide. The effective index of the upper waveguide increases along the propagation direction and the confinement factor also increases. With the correct design of the coupler, there is one point within the overlapping tapers at which the effective index of each waveguide is equal. At this point, the optical impedance of the waveguides is matched and therefore this is the point at which strong coupling occurs.

3.1.2 Performance and Design Tradeoffs

The vertical coupler design, shown in Figure 3.1, was modeled for various specific geometries. The purpose of the model was to determine the robustness of the taper design. The goal was to test coupling between two different materials, different taper geometries, and to learn more about the coupling mechanisms and trade-offs. In order to simulate the coupling in the taper design, a beam propagation method was used. The simulation software is commercially available software called BeamPROP™ from RSoft Design Group. Within the model, the lower waveguide was considered to be crystalline silicon with a refractive index (n) of 3.5. The upper waveguide was considered to be amorphous silicon with a potentially variable refractive index. Since the properties of a-Si can change depending on the processing conditions, the refractive index was varied from a minimum index of 3.3 to 3.7 [13], [38], [50]. The waveguide dimensions were kept constant with a w = 500 nm width and an h = 200 nm height, which is representative of a single-mode waveguide for light with a wavelength of λ = 1.55 μm. In addition, the vertical coupling height, h_s, was kept constant at 200 nm. For initial analysis, the taper tip width, w_t, was kept constant at 100 nm. With this initial geometry, the effect of taper

overlap length, L, and the refractive index of a-Si was studied. The results of the simulation are shown in Figure 3.2.

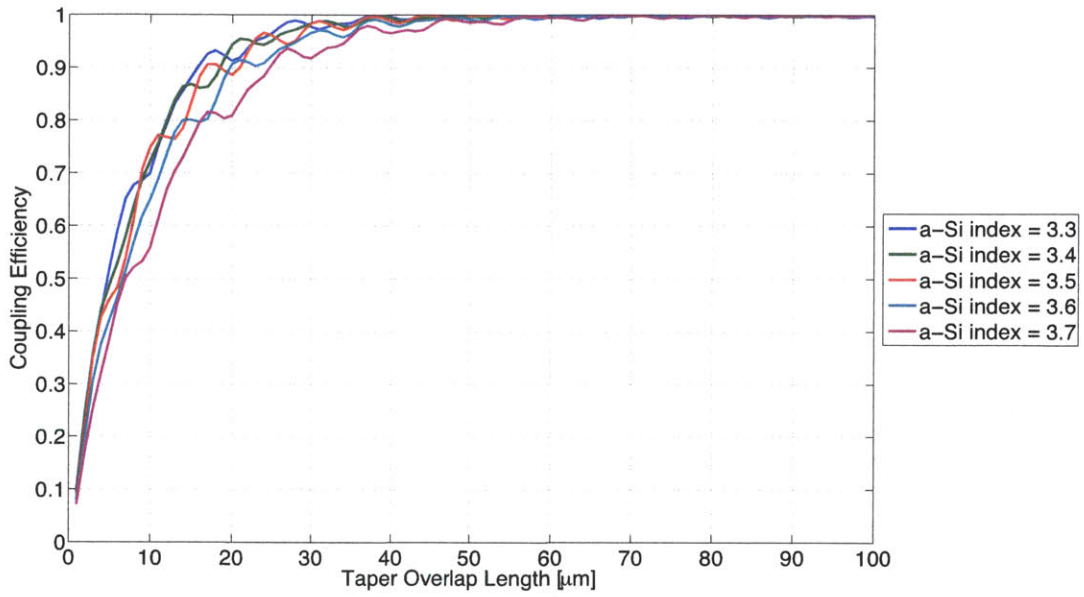

Figure 3.2 Coupling efficiency for varying taper lengths. For taper overlap lengths greater than 60 μm nearly perfect coupling is achieved. Nearly perfect coupling is achieved for a-Si waveguides that have an index of refraction greater than and less than that of crystalline Si. For these simulations, the taper tip width is kept constant at 100 nm. Upon inspection, there is a second order effect that is creating some periodicity in the plot.

It is evident from Figure 3.2 that very high coupling efficiencies can be achieved with the vertical coupler design. Nearly perfect coupling can be achieved for all ranges of refractive indices of a-Si at taper overlap lengths greater than 60 μm. This is due to the dual taper design. If the taper tip width is narrow enough, then the effective index of the taper will vary from the effective index of the waveguide, to the refractive index of the cladding. Therefore, there is always some point within the tapers where the effective index is equal and therefore the optical impedance matching condition is met, leading to efficient power coupling. By increasing the refractive index of the a-Si, this only increases the effective index of the mode within the straight waveguide. Within the taper, the effective index will still decrease along the propagation

direction until it matches the effective index of the lower waveguide taper. On the other hand, if the refractive index of a-Si is decreased, then the effective index of the upper waveguide is decreased. However, the taper in the lower waveguide reduces the effective index in the lower waveguide enough so that there is still a point in which the effective index in each waveguide is equal. Therefore, the vertical coupler design that implements two tapers allows for a robust design to impedance match waveguides of different refractive indices.

The plot in Figure 3.2 shows a general trend of increasing coupling efficiency with increasing taper length. This is due to the changing slope of the taper. If the waveguide tapers very quickly, then the taper strays from the adiabatic taper condition. This may excite higher order modes and initiate scattering losses. However, if the taper is long, then the cross-section of the taper will evolve more slowly and it will act more like an adiabatic taper. In an adiabatic taper the mode evolves slowly, stays in the fundamental mode, and experiences no losses.

When planning to design real devices, processing constraints become a concern. A large limitation to fabrication is due to lithography tolerances. Very small dimensions become more difficult and more expensive to fabricate. A 100 nm taper tip width is possible to fabricate, but it is worth exploring larger taper tip dimensions. A wider taper tip width would be easier and cheaper to fabricate. Therefore, after the general trends were established for waveguide refractive index and taper length, the effect of taper tip width was examined. The taper tip width was varied from a perfect point with w_t = 0 nm, to a relatively blunt taper with a w_t = 300 nm. Both waveguides were kept at a constant cross sectional width of 500 nm. The refractive index of the a-Si was fixed at 3.6. The results of the simulation are shown in Figure 3.3. Once again, some similar trends are observed as were seen in Figure 3.2. In general, the coupling efficiency increases with increasing taper overlap length. However, a very clear new trend is also observed.

There is a second order effect that can be seen in addition to the typical monotonic increase in coupling efficiency with increased taper overlap length. There is an additional periodicity whose amplitude increases with increased taper tip width.

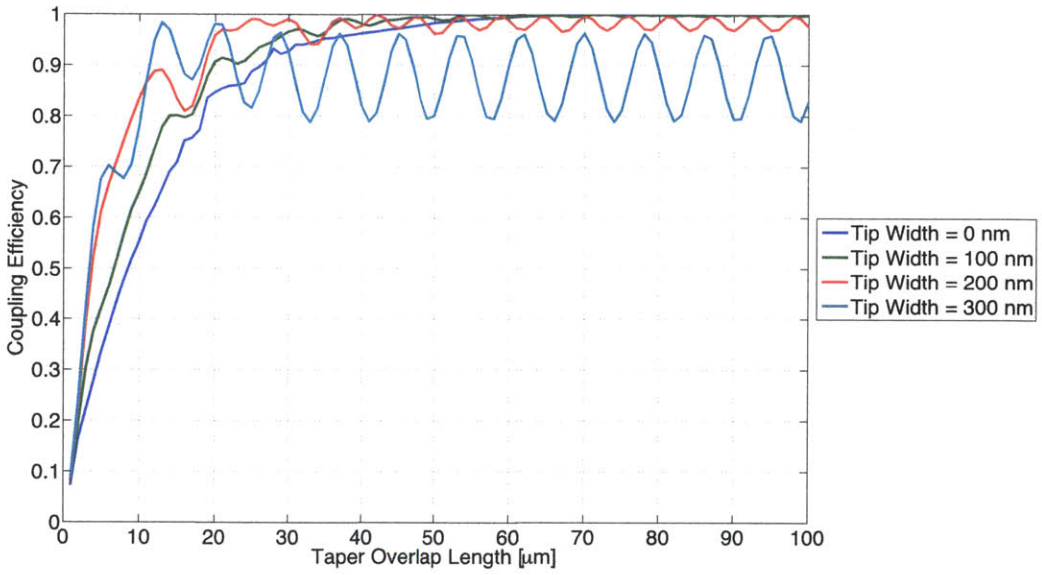

Figure 3.3 Coupling efficiency at different taper tip widths. The refractive index of the a-Si is fixed at 3.6. This figure shows a clear second order periodicity. The amplitude of the periodicity scales with the taper tip width. The largest periodic disturbance is when the taper tip width is 300 nm. For the taper tip width of 300 nm, there is a local maximum at a taper overlap length of 20 μm and a local minimum at 25 μm.

When the taper comes to a perfect point (w_t = 0 nm), there is no periodicity. Instead a monotonic increase in coupling efficiency is observed with increased taper overlap length. In addition, the coupling efficiency nears 100% as L > 60 μm. When the taper tip width increases, especially for w_t = 200 nm and w_t = 300 nm, a noticeable second order periodicity is observed. The oscillations have a period of ~5 μm and ~8 μm for w_t = 200 nm and w_t = 300 nm respectively.

There are two main effects that would yield a periodic coupling efficiency as a function of taper overlap length. The first explanation is the potential formation of a Fabry-Perot cavity, and the second is due to beating of the mode between the two waveguide tapers. The period of the oscillations that would form from a Fabry-Perot cavity (L_{FP}) are given by Equation 9.

$$L_{FP} = \frac{\lambda}{2 \cdot n_{eff}} \tag{9}$$

In Equation 9, L_{FP} gives the period of the oscillations that would form from a Fabry-Perot cavity. This equation is derived from the definition of a resonant cavity. A cavity will become resonant if an integral number of half-wavelengths will fit within the cavity. Therefore, the $\lambda/2$ term gives the length of one half of a wavelength. The n_{eff} term is the effective index of the mode within the tapered coupler. It is a weighted average of the refractive indices that the propagating mode interacts with. This term is added in the denominator because the free-space wavelength (λ) will be reduced by a factor of n_{eff} while in the tapered structure. A local minimum in the plot in Figure 3.3 would be caused by a resonant cavity formation. This is because the light would become trapped in the cavity and not propagate through the coupler. It is difficult to calculate L_{FP} exactly because n_{eff} is changing throughout the tapered structure as the mode evolves. However, there are some definite bounds that can be imposed. The effective index must at least be within the bounds of the refractive indices of the materials that it interacts with. Therefore, n_{eff} is bounded to be within the refractive index of the a-Si and the refractive index of SiO_2, 3.6 and 1.46 respectively. When these values of n_{eff} are used in Equation 9, along with the value of λ = 1.55 μm, a range of 0.22 μm < L_{FP} < 0.53 μm is established. In reality, stricter limitations can be placed on n_{eff}. The maximum effective index occurs when the mode is strictly confined to a single waveguide, and not in the taper. The effective index in the single-mode waveguides is about 2.5 and therefore the Fabry-Perot period length is further limited to 0.22 μm < L_{FP} < 0.31

μm. The period of oscillation given by the Fabry-Perot cavity is much less than the period of oscillation seen by the graph in Figure 3.3, and hence the oscillations observed are not due to the formation of a Fabry-Perot cavity.

In order to verify that the oscillations are not due to the formation of a resonant cavity, a wavelength sweep was simulated in order to observe the effect of the free-space wavelength on the coupling efficiency. The results of the wavelength sweep are shown in Figure 3.4. If the tapered coupler forms a resonant cavity, then a clear free spectral range (FSR) would be observed. The expected FSR is given by Equation 10 where λ is the free-space wavelength, L is the taper overlap length, and n_{eff} is the effective index.

$$FSR = \frac{\lambda^2}{2 \cdot n_{eff} \cdot L} \tag{10}$$

In order to calculate the predicted FSR, the same central free-space wavelength was chosen to be λ = 1.55 μm. The cavity length was chosen to be a local minimum of Figure 3.3, where a resonant cavity would be formed. A clear local minimum occurs at L = 25 μm for at taper tip width of 300 nm. A local minimum in coupling efficiency would indicate a resonant cavity formation, which is why this length value was chosen. The conservative bounds of n_{eff} are used such that the maximum and minimum values are 3.6 and 1.46 respectively. Using these values, a range for the FSR is calculated to be between 13.3 nm and 32.9 nm. However, upon inspection of Figure 3.4, no FSR is evident. The FSR would seen as local sharp drops in coupling efficiency with a period equal to the FSR. Since there are no such drops, it can be concluded that the oscillations from Figure 3.3 are not formed from the creation of a resonant cavity.

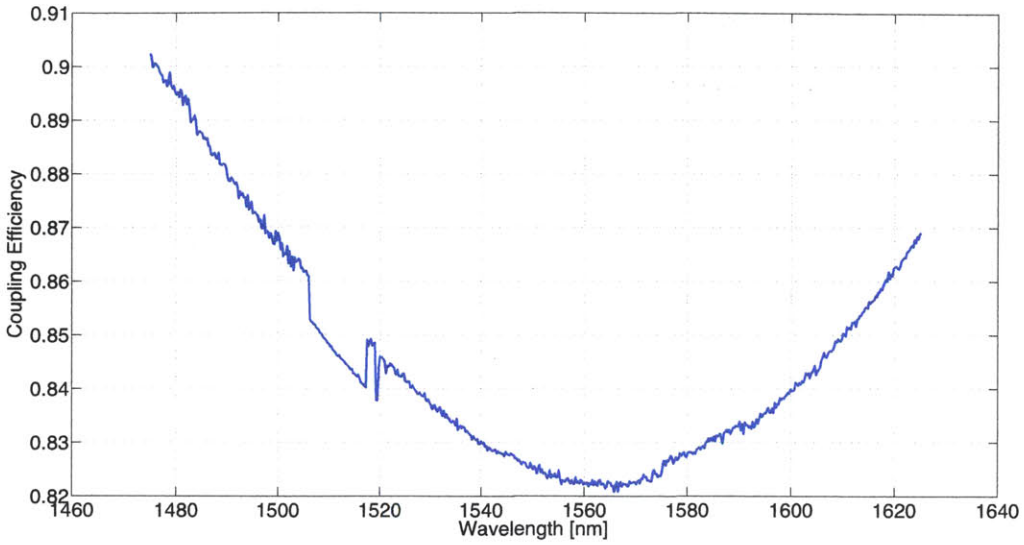

Figure 3.4 Free-space wavelength dependence of the coupling efficiency. Here, the geometry of the local minimum point of Figure 3.3 is taken at which the taper tip width is 300 nm and the taper overlap length is 25 μm. If the oscillations were due to a Fabry-Perot cavity, then there would be strong wavelength dependence. The expected free spectral range of this cavity would be between 13.3 nm and 32.9 nm. Since there are clearly no coupling efficiency drops with a periodicity between 13.3 nm and 32.9 nm, it can be concluded that the periodicity observed in Figure 3.3 can be attributed to modal beating, and not resonance in a Fabry-Perot cavity.

There may still be a resonant cavity formation, which would have a periodic effect on the coupling efficiency, however these effects are not seen on the beam propagation simulation. The increased tip width would also give yield to increased oscillation amplitudes. This is due to the increased reflections caused by a greater effective index jump at the taper ends. A taper that comes to a point has a smooth transition of the effective index from the taper to the cladding at the end of the taper. This eliminates all reflections and therefore eliminates the formation of a resonant cavity. To further investigate the presence of resonant cavity effects, a complete finite-difference time domain (FDTD) simulation would be the best technique.

Having eliminated the possibility that the oscillations in coupling efficiency are caused by the presence of a resonant cavity, the most likely explanation is due to a beating mode. Other

studies have shown a similar beating between two parallel waveguides with beating periods of 4 to 7 μm [51]. This period is very close to the period that is observed in this simulation. The beating of the mode between the two waveguides can be described by coupled mode theory [49], [52], [53]. The vertical coupler design is similar to a directional coupler in which the propagating mode will beat between the two adjacent waveguides. The period of the beating mode (L_{BM}) is given by Equation 11, where β_1 is the propagation constant in the lower crystalline waveguide taper, and β_2 is the propagation constant in the upper a-Si waveguide taper.

$$L_{BM} = \frac{2\pi}{|\beta_2 - \beta_1|} \tag{11}$$

It is difficult to solve for L_{BM} directly since the propagation constants of each taper change along the propagation direction. If $\beta_1 = \beta_2$, then the two tapers are said to be phase matched and efficient power transfer occurs between the two waveguide tapers [54]. In order to investigate the presence of an oscillating mode, individual geometry simulations were examined.

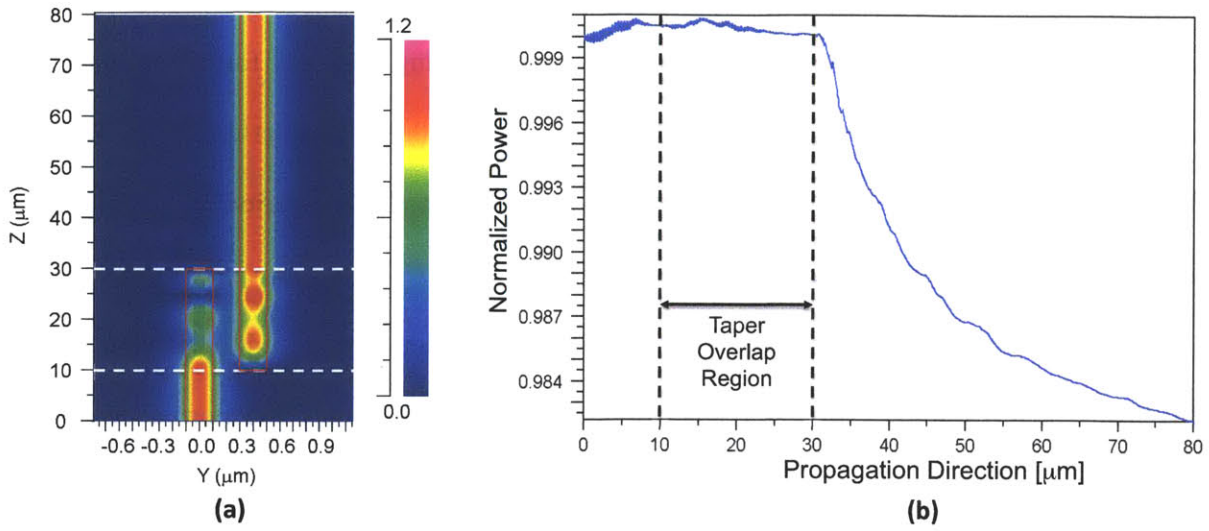

Figure 3.5 Simulation of the vertical coupler with a taper overlap of 20 μm and a taper tip width of 300 nm, a local maximum in Figure 3.3. Figure (a) shows the modal evolution, while figure (b) shows the total power in the waveguide, normalized to the input power. The red outline in (a) is demarcates the geometry of the waveguides and tapers from a cross-section side view perspective. The light is injected upwards from the point Y = 0, and Z = 0, and propagates in the positive Z direction. The white dotted lines in (a) show the limits of the tapers, while beyond the bounds are straight waveguides that allow for additional modal evolution. The equivalent taper overlap boundary is marked by black dotted lines in (b). In (a), it is seen that there is some modal beating between the two tapers, but at the point where the lower taper ends, the mode is in the upper waveguide. This is elucidated in figure (b) by the low coupling loss where the coupling efficiency is greater than 98%.

In Figure 3.5, the vertical taper is simulated at a local maximum of Figure 3.3. All of the geometry parameters are kept constant, and a taper overlap length of 20 μm is examined in further detail. This geometry yielded a local maximum in the periodic behavior for a taper tip width of 300 nm. The simulation results in (a) show the integrated power in the coupler in a longitudinal cross-section view. The red outline demarcates the geometry of the two waveguides. The lower crystalline-Si waveguide spans the vertical distance of -0.1 μm < Y < 0.1 μm and the upper a-Si waveguide spans the vertical range of 0.3 μm < Y < 0.5 μm in the simulation. The region bounded by the two dotted lines shows the region of the waveguides that is tapered and refers to the taper overlap length of 20 μm. The light is injected from the Z = 0 plane and

propagates in the +Z direction. From the simulation results in (a), it is evident that modal beating is occurring. In the overlapping section, the power is oscillating back and forth from the upper and lower waveguide tapers with 2.5 periods observed within this taper overlap length. With 2.5 periods occurring within the 20 μm taper, this yields an oscillation period of 8 μm, which matches exactly with the period observed in Figure 3.3.

The plot in (b) shows the integrated power in the coupler, normalized to the input power. The black dotted lines delineate the span of the taper overlap range. The normalized power remains constant at about 1 throughout the oscillations within the taper overlap region. This means that there are no inherent losses associated with the optical power beating between the two waveguides. At the end of the taper, there is a slight loss. This is caused by the reflections from the evanescent tail of the mode overlapping with the end of lower waveguide. The majority of the mode is within the upper a-Si waveguide at the taper end, but there is still some part of the evanescent tail that interacts with the sharp refractive index change at this point. This part of the evanescent tail experiences some reflections. However, the amount of reflection is minimal as the total power coupled is still greater than 98%.

In order to understand what causes the local minima in coupling efficiency, the geometry that yields relatively low coupling efficiencies is shown in greater detail in Figure 3.6. In this simulation, the same vertical taper geometries as modeled in Figure 3.5 are preserved. The only difference is the taper overlap length is increased to 25 μm. This condition yielded a local minimum in the plot in Figure 3.3.

Once again, the schematic in (a) shows the longitudinal cross section view of the power distribution in the two waveguides and tapers. The red line outlines the waveguides and tapers and the white dotted line marks the span of the taper overlap region. Once again, it is evident

that there is modal beating and the power oscillates between the two waveguides. The main difference between this case, and the case where the taper overlap was 20 μm is the location of the mode at the end of the taper. In the case where the coupling efficiency was high (Figure 3.5), the majority of the mode was in the upper waveguide at the end of the taper. This way, only the evanescent tail of the mode interacted with the sudden index change at the end of the taper. However, in the case where the coupling efficiency is low, the majority of the mode is in the lower waveguide at the end of the taper. In this case, the majority of the mode experiences the sudden effective index change and therefore a large part of the mode is susceptible to scattering.

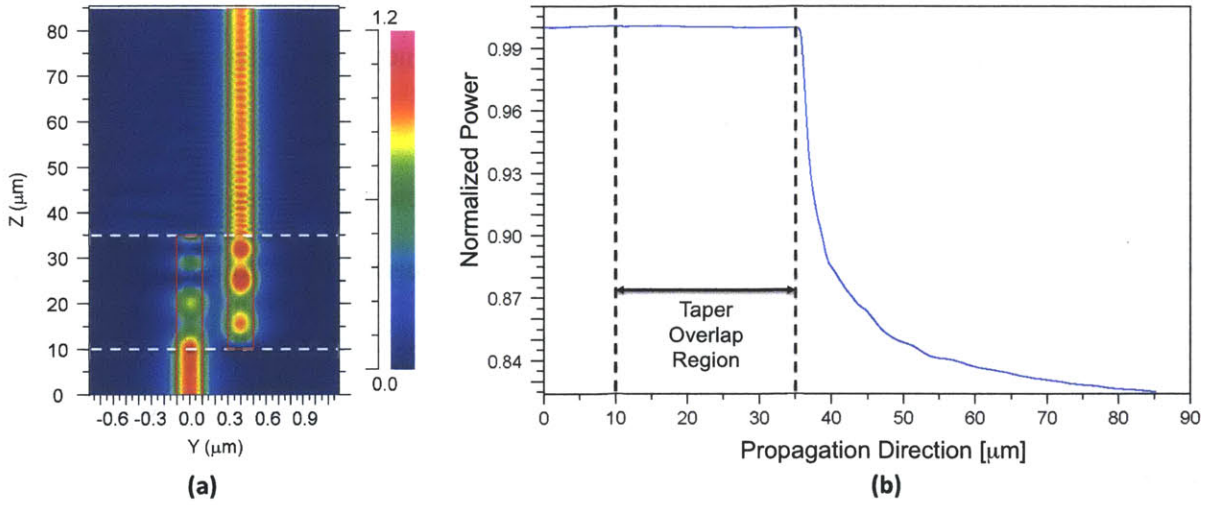

Figure 3.6 Simulation of the vertical coupler with a taper overlap length of 25 μm and a taper tip width of 300 nm. This was a local minimum in Figure 3.3. Figure (a) shows the modal evolution, while figure (b) shows the total power in the waveguide, normalized to the input power. The red outline in (a) is the outline of the waveguides and tapers from a side view perspective. The light is injected upwards from the origin, and propagates in the positive Z direction. The white dotted lines in (a) show the limits of the tapers, while beyond this is straight waveguides that allow for additional modal evolution. Here it is seen that there is some modal beating between the two tapers, but at the point where the lower taper ends, the mode is in the lower waveguide. Therefore, there is a sudden effective index change as the taper ends, which causes scattering. This is elucidated in figure (b) by the increased coupling loss where the coupling efficiency is less than 83%.

The plot in (b) of Figure 3.6 shows the power, normalized to the input power, in the coupler along the propagation direction. The black dotted line marks the span of the taper

overlap. Once again, the power in the coupler is constant at 100% along the taper overlap region. The power is conserved despite the mode oscillating between the two waveguides. However, there is significant losses at the end of the taper. The graph in (b) quantifies the result that can be inferred from the power map in (a). At the end of the taper, the majority of the mode is in the lower waveguide. When the taper ends, there is a sudden effective index change which induces scattering losses. The end result is a geometry which yields a coupling efficiency less than 83%.

This therefore proves that the periodicity in coupling efficiency is caused by modal beating between the two tapered waveguides. A wider taper tip will have a larger effective index change at the end of the taper, yielding a larger reflection and therefore larger magnitude oscillations. In contrast, a taper that ends in a point will have a negligible effective index change and hence not suffer from any reflections.

An overlapping tapered structure is capable of efficient power transfer between two waveguide devices. A general design of an input waveguide tapering down to a narrower width, while the output waveguide is tapering to a wider width, can yield nearly perfect coupling efficiency with minimal scattering losses. This is because the correct design will always yield a phase matching condition, which will exhibit efficient power transfer between the two waveguides. With finite taper tip widths, reflections become a concern, due to modal beating, but they can still be designed for efficient power transfer.

3.2 Electro-Absorption Modulators

In order to transmit data with a photonic interconnect, the light must be encoded with digital signals. In order to do this, the light must be modulated on and off at high frequencies. There are two main ways to encode data into an optical interconnect. One method is to directly modulate the laser to directly encode the data [55–59]. The other technique is to have a

continuous wave laser light source, and then use an external modulator. Most planar, on-chip integrated photonic links use a continuous wave light source with external modulation [60–64]. On-chip modulators are the current method of encoding data into CMOS silicon photonic optical interconnects. Therefore, the focus will be on external modulation with a continuous wave laser light source.

There are two main types of external modulators: refractive modulators and absorptive modulators [3]. Both modulation techniques take advantage of a change in the complex refractive index. Refractive modulators utilize a change in the real part of the refractive index (Δn) and are called electro-optic modulators. Absorptive modulators utilize a change in the imaginary part of the refractive index ($\Delta\alpha$) and are called electro-absorption modulators. Electro-optic modulators utilize refractive index shifts and therefore modulate the light using interference. Examples of these types of modulators are Mach-Zehnder interferometer (MZI) modulators [65–68] and ring resonator modulators [69–71]. Electro-absorption modulators utilize a shift in absorption coefficient, with applied electric field, to absorb light. Electro-absorption modulators are typically based on the quantum-confined Stark effect [72], [73], or the Franz-Keldysh effect [11], [35], [60], [74]. These two techniques are very similar with the Franz-Keldysh effect applying to bulk materials, while the quantum-confined Stark effect occurs in quantum wells. Therefore, the Franz-Keldysh is the limit of the quantum-confined Stark effect as the quantum well thickness increases.

Electro-optic effects are typically weak ($\Delta n \cong 1\times10^{-3}$) and therefore require long interaction lengths [3]. For a wavelength of 1.5 μm, if the refractive index is only changed in one arm of an MZI, then a device length of 750 μm is required for a $\lambda/2$ path-length change. Therefore to cause a phase change of π for deconstructive interference, MZI modulators require

path lengths in the hundreds of micron range, or even longer. This yields device footprints of about 10^3 – 10^4 μm^2 [75]. These devices also require large amounts of energy per bit for modulation. This is because carrier injection and depletion occurs over a very large volume and therefore device switching energy is in the range of 5×10^3 – 3×10^4 fJ/bit [66], [67]. In order to minimize this footprint, resonant cavities have been created to maintain high interaction lengths, while minimizing device footprint. This reduces the device footprints to a few micron radii for ring and disk resonators. Reducing the active area also reduces the energy consumption to about 50 – 85 fJ/bit [76], [77]. Although resonant cavity electro-optic modulators may be fast, low-power and have small device footprints, their working spectrum is inherently limited by the resonant wavelength of the cavity. Therefore, micro-ring and micro-disk electro-optic modulators have a working spectrum span of 0.1 – 0.2 nm, or even smaller. This imposes a large limitation because, due to thermo-optic effects. The index of refraction of rings is also very sensitive to temperature shifts. In order to insure reliable WDM operation, a ring resonator device would have to remain within 0.1 °C of the design temperature [78].

While electro-optic modulators show promising results and capabilities, they are not ideal for large scale integration into electronic-photonic integrated circuits (EPICs). Depending on the design, they are either limited by size, power consumption, temperature fluctuations, or operating spectrum. On the other hand, electro-absorption modulators (EAMs) show great promise for dense integration in EPICs. Since the observed electro-absorption effects are stronger than the electro-optic effects, with $\Delta\alpha \cong 10^2 - 10^3\ cm^{-1}$ [3], modulators based on the Franz-Keldysh effect are very compact with active areas in the range of 30 – 45 μm^2. They are also high-speed with bandwidths measured up to 12.5 GHz and predicted up to 30 GHz. In addition, energy consumption can be very low with reported values of 50 – 100 fJ/bit. Electro-absorption

modulators are not limited by narrow operation spectrum widths, like resonant cavity devices, and are operational over a 14 – 30 nm spectral range [11], [35]. This makes them useful for WDM applications. A summary of the device specifications is shown in Table 3.1.

Table 3.1 Direct comparison of different external modulation techniques. For bandwidth, the maximum of the cited references is shown.

	Si MZI*	Micro-Ring/Disk**	GeSi EAM***
Active Footprint (μm^2)	$10^3 - 10^4$	$20 - 10^3$	30 – 45
Bandwidth (GHz)	40	35	30
Energy Consumption (fJ/bit)	$5\times10^3 - 4\times10^4$	50 – 85	50 – 100
Working Spectral Width (nm)	~20	0.1 – 0.2	14 – 30
Insertion Loss (dB)	7 – 12	0.5 – 2	2.5 – 5

*The specifications from the Si MZI modulators came from [66], [67].
**The specifications for the micro-ring and micro-disk modulators came from [70], [76], [77], [79].
***The specifications for the GeSi Electro-absorption modulators came from [11], [35].

The cumulative specifications shown in Table 3.1 elucidate the benefits of utilizing a Ge based electro-absorption modulator. This type of design has the capability of combining small footprint, high bandwidth, low energy consumption, and modulation over a wide spectral range. Therefore, further investigation is focused on the integration of Ge based electro-absorption modulators. The only specification in which Ge-based electro-absorption modulators do not exhibit either equal or enhanced performance, is based on insertion losses. Micro-ring and micro-disk modulators have lower insertion losses than Ge EAMs. Therefore, any technique of reducing the total insertion loss of these modulators must be examined.

3.2.1 Background

The Ge-based electro-absorption modulator is an excellent candidate for external modulation of light in integrated optical interconnects. A design element that must be considered for all modulators is the ratio of extinction ratio to insertion losses. The ideal device has a high extinction ratio and low insertion losses. The extinction ratio is determined by device length and applied electric field. Fabricated devices have shown an extinction ratio of 8 dB [11]. However, a tradeoff exists between extinction ratio and insertion losses. A longer device yields an increased extinction ratio, however there will be increased indirect band absorption, hence an increased insertion loss. This tradeoff is embedded within the material properties, and can be optimized. However, one parameter that can independently reduce insertion loss, without lowering the extinction ratio, is to reduce coupling losses.

Insertion losses are composed of the combination of material absorption, and coupling loses. The coupling losses can be reduced with an efficient coupling mechanism, without affecting the extinction ratio. The designs for electro-absorption modulators cited in Table 3.1 suffer from coupling losses of 1.34 and 1.7 dB. This means that 27% and 32% of optical power is lost merely to scattering and inefficient butt-coupling. With energy efficiency as a potential benefit of photonic interconnects over metal interconnects, these losses must be minimized. A schematic of the existing coupling approach is shown in Figure 3.7. The approach takes advantage of a low-loss vertical coupler from a crystalline-Si waveguide to an a-Si waveguide. However, the coupling losses occur within the butt-coupling from the a-Si waveguide to the Ge modulator. At this point, there is a sudden change in effective index along the propagation direction, and therefore scattering occurs, effectively losing 32% of the optical power. If the

coupling losses were negligible, then the total insertion loss could be reduced from 3.7 dB to 2.0 dB. A 2 dB insertion loss is equivalent to some micro-ring modulators.

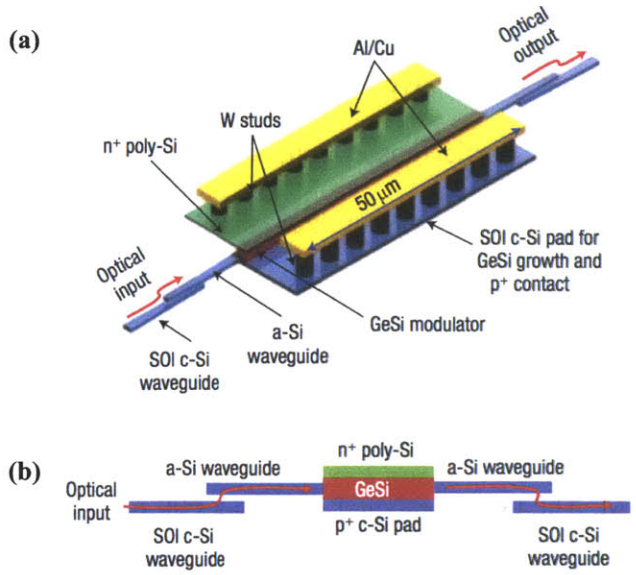

Figure 3.7 Schematic of current Ge electro-absorption modulators. A 3D view is shown in (a), while (b) shows the cross-sectional side-view. In this design, light is coupled vertically from a crystalline Si waveguide on an SOI platform to an a-Si waveguide. Then, the light is butt-coupled from the a-Si waveguide to the Ge-based modulator. In the end, it is butt-coupled out of the modulator, back into the a-Si waveguide, and then vertically coupled back into the crystalline Si waveguide. Coupling losses reside in the butt-coupling in and out of the Ge. From [11].

The butt-coupled technique is clearly not the ideal coupling approach. With this design, 32% of optical power is wasted merely on scattering from coupling losses. Therefore, a better approach would be the utilization of low-loss evanescent coupling.

3.2.2 Tapered Lateral Evanescently Coupled Modulator Design

The approach of designing vertical couplers by utilizing dual tapered waveguides yielded nearly perfect coupling efficiencies. Since this evanescent coupling design was proven efficient and robust, the same approach was taken for designing a coupler to the Ge electro-absorption

modulator (EAM). The EAM can be considered to be a waveguide while designing the coupling scheme. The Si waveguide and the Ge waveguide have different refractive indices. They may be different sizes and may not be collinear. For these devices, a vertical coupling scheme is not ideal. This is because the light would have to be coupled through either the doped p-type and n-type Si. This may induce free-carrier absorption, which would increase the insertion losses. Therefore, a lateral coupling scheme was utilized, similar to the design proposed in [80]. A schematic of the lateral, evanescently coupled EAM design is shown in Figure 3.8. In the figure, the yellow represents the Si input and output waveguides. The waveguide cross section is 500 nm width by 200 nm height, in order to keep it single-mode. The green is intrinsic Ge and is the active region in the EAM. The Ge is designed to have a 600 nm width and 400 nm height. The blue indicates p-type Si, and the red represents n-type Si. The p-type Si, and the n-type Si are designed to have 100 nm thicknesses. The combination of the p-type Si, Ge, and n-type Si forms a vertical heterojunction p-i-n diode.

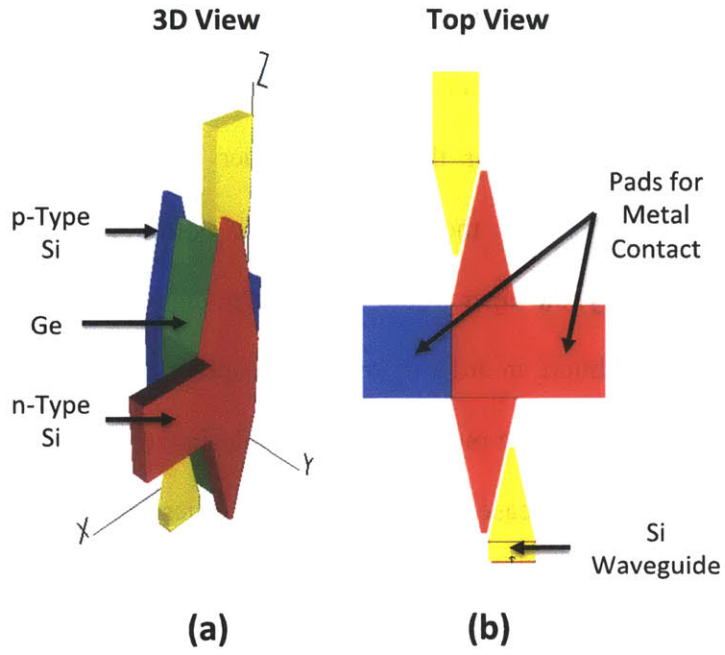

Figure 3.8 Design of a tapered lateral evanescently coupled electro-absorption modulator. The schematic in (a) shows the 3D view of the modulator design, while (b) shows the top view. The light is injected from the bottom Si waveguide, and travels upwards, in the +Z direction. The waveguides and the modulator are both tapered. The light then evanescently couples laterally from the waveguide to the modulator, and back to the second waveguide. This design eliminates the need for butt-coupling and therefore reduces insertion losses and minimizes reflections.

In the tapered EAM design, the light is evanescently coupled from the waveguide to the modulator, and back. This design eliminates the backscattering losses due to butt-coupling, as observed from the original design. Along the propagation direction, the input waveguide tapers down to a narrow width of 100 nm, causing the mode to expand into the cladding and therefore decreasing the effective index and the confinement factor. While the waveguide tapers to a narrow tip, the EAM tapers from a narrow 100 nm tip, to the nominal modulator width of 600 nm. Along the propagation direction of the modulator, the effective index slowly increases while the mode is being coupled into the Ge waveguide modulator. The centerline of the waveguides and the Ge EAM are all along the same height in order to maintain symmetry within the coupler.

This design requires the use of deep ultra-violet lithography in order to create feature sizes as small as 100 nm. This is in order to obtain efficient coupling over a short distance with minimal scattering. The modeling for the vertical taper showed increased scattering with increased taper tip widths. Therefore, the taper tips in the evanescently coupled modulator are designed to be 100 nm. Ideally, the taper would come to a point, but this is unrealistic in terms of fabrication limitations. In addition, in order to keep the coupling distance low, the horizontal gap between the waveguide taper and the modulator taper is designed to be 100 nm. A smaller gap between tapers results in an increased overlap of the evanescent tail of the mode in each waveguide. This increased overlap allows for more efficient evanescent coupling over short distances.

There is an optimal length for Ge-based electro-absorption modulators. Since they utilize the Franz-Keldysh effect, the operating wavelength is just below the direct band-edge. Therefore, when the diode is unbiased, the modulator is transparent and there is no direct-band absorption. When an electric field is created, via reverse biasing, the Ge exhibits absorption just below the direct band edge. This absorption depletes the optical power, effectively modulating the light. However, since Ge is an indirect gap material, there is still weak absorption in the range between the indirect and direct band edges. Therefore, although light in the wavelength range of 1.55 μm to 1.85 μm has lower energy than the direct bandgap of Ge, it will still be inefficiently absorbed. This must be accounted for when designing the modulator length because there is a tradeoff between extinction ratio and insertion losses, independent of coupling efficiency. In order to allow for optimization in modulator length a short taper section is ideal. The modulator length can be increased by adding a straight waveguide section between the tapered sections, but the minimum length of the modulator is defined as double the taper length. Using the vertical

coupler simulations in Figure 3.2 and Figure 3.3 as a guideline, a taper length of 25 μm was chosen. This is the taper overlap length at which the coupler began to yield high coupling efficiencies. Since the gap between the two tapered waveguides is much smaller for the modulator design, a shorter taper length is predicted. In addition, the EAM design fabricated in [11] was 50 μm long, and hence a 25 μm taper allows for nominally equivalent modulator lengths.

While the coupling in the vertical coupler, from Section 3.1 is in the vertical direction, the coupling in the modulator design is lateral. Despite the coupling direction, the mechanism is the same. In addition, the coupling out of the modulator is the same as the coupling into the modulator, only the direction is reversed.

3.2.3 Evanescently Coupled Modulator Performance

In order to model the coupling efficiency, the light propagation through the waveguide was modeled using a beam propagation method (BPM). The same commercial software was used as the vertical couplers, namely BeamPROP™ from RSoft Design Group. A simulation of the ideal design, which yields the highest coupling efficiency, is shown in Figure 3.9. The figure shown in part (a) is the total optical power distribution along the waveguide and modulator. The view is a top-down perspective of the waveguide and modulator coupling scheme with the general light propagation path indicated by the white arrows. Here, it is clear that the input and output waveguides are single-mode, yet there are some higher-order modes within the modulator. The higher-order modes are seen by the lateral oscillations of the optical power within the Ge EAM. Different modes may travel at different speeds in a waveguide, because they experience different effective indices. Therefore, some interference may be present between the modes,

causing oscillations in the power. Although some higher-order modes are present within the Ge EAM, the light in the output waveguide is still single-mode and therefore exhibits low loss transmission.

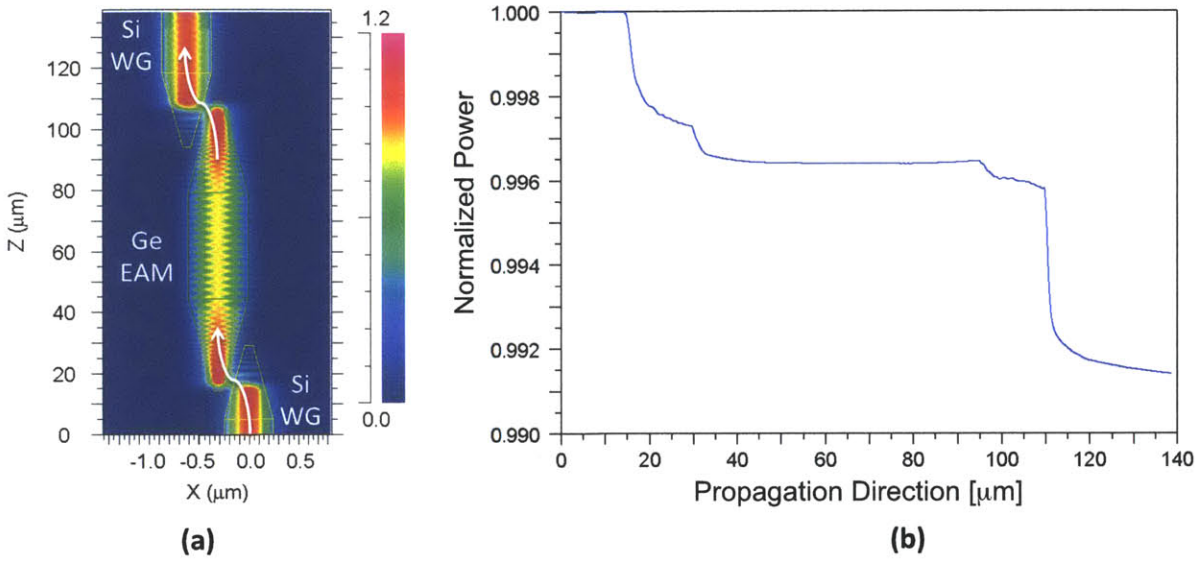

Figure 3.9 Simulation of a tapered lateral evanescently coupled electro-absorption modulator showing low coupling losses. The simulation in (a) shows a top-view of the modal evolution of the propagating light. The light is input from the point X = 0 and Z = 0 and propagates in the +Z direction. The white arrows outline the general path of the light. There are some lateral oscillations of the mode within the modulator, however it evident that the light that couples out of the modulator, into the second waveguide, is single-mode. The plot in (b) shows the power, normalized by input power, as a function of propagation distance. Since material absorption is neglected, all losses are due to coupling losses. Therefore, (b) shows that this design yields a greater than 99% coupling efficiency.

The total integrated power in the modulator coupling design is plotted in part (b) of Figure 3.9. The power has been normalized to the input power. Here, it is clear that the coupling efficiency of the entire design is 99%. This shows that although higher-order modes may exist within the Ge EAM, the light is coupled back out to the fundamental mode. There are two clear drops in intensity in the plot in part (b). These two drops in power are associated with the taper tips. A similar effect was observed in the vertical couplers that were modeled in Figure 3.5 and

Figure 3.6. This sharp drop is due to scattering losses from the end of the taper. Although most of the mode has coupled into the modulator at the first drop, and the output waveguide at the second drop, there is still an evanescent tail that interacts with the end of the original taper. This evanescent tail interacts with the sudden change in effective index, caused by a blunt taper tip, and hence is partially reflected, causing scattering losses. However, since most of the mode is already coupled, only a small amount of the evanescent tail of the mode is scattered, and the total coupling efficiency remains high. This optimal tapered modulator design yields nearly perfect coupling efficiency. It is capable of reducing the coupling losses of the original design from 1.7 dB to 0.04 dB.

In order to test the robustness of the tapered modulator design, the effect of taper overlap was examined. The taper overlap is defined in the top-view schematic of part (a) of Figure 3.10. The coupling efficiency, as a function of taper overlap length, is shown in the graph in part (b). Similar to the vertical coupling simulations, an oscillating coupling efficiency is observed. In order to more closely examine the cause of these oscillations, a local minimum was explored in (c), and a local maximum was explored in (d).

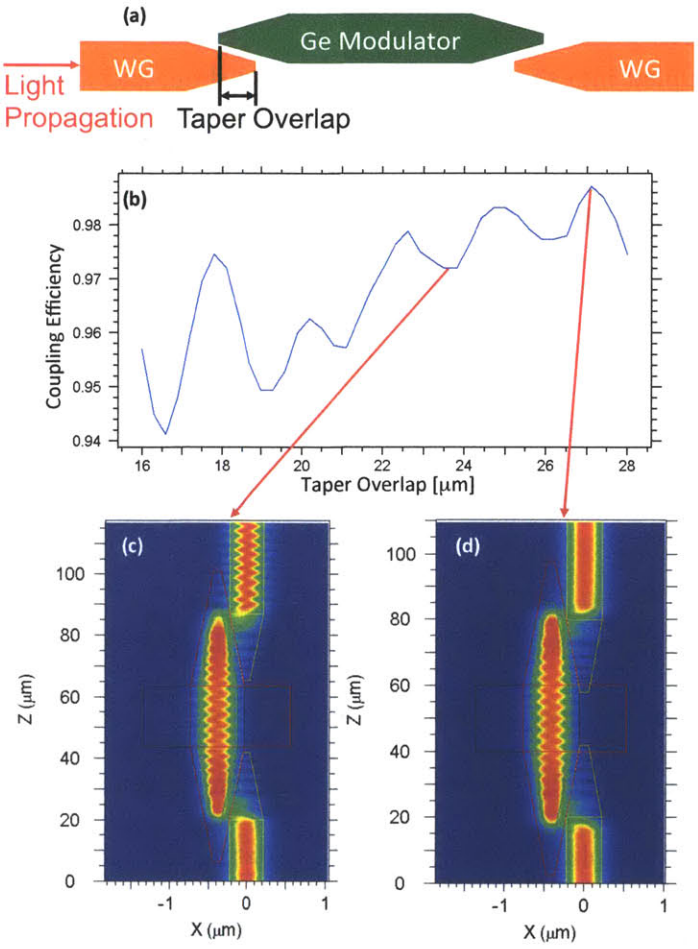

Figure 3.10 A study of the effect of taper overlap between the waveguide and the modulator. A top view of the device is shown in (a) and the taper overlap is defined. In (b), the coupling efficiency is examined as a function of taper overlap. A high coupling efficiency of greater than 98% can be attained with this design. The plot in (b) shows some periodicity, similar to Figure 3.3, and the reasoning is the same. The simulations in (c) and (d) show some mode beating between the waveguide and the modulator. The local minimum of (b) is shown in (c). The simulation in (c) shows some coupling into higher order modes in the output waveguide, which is lossy. However, (d) represents a local maximum in coupling efficiency in (b) and shows efficient single mode coupling to the output waveguide.

The simulations in (c) and (d) of Figure 3.10 show modal oscillations between the Ge EAM and the input and output waveguides. This is due to modal beating, as explained during the vertical coupler discussion. Therefore, a similar explanation can be utilized to explain the source of the periodicity. At the local maximum, as shown by (d), the mode has completely transferred

from the input waveguide, into the Ge EAM at the end of the waveguide taper. Therefore, there are minimal reflections and highly efficient evanescent coupling. Conversely, at the local minimum, as shown by (c), the mode has not completely coupled out of the modulator at the end of the taper. This is more evident at the coupling out of the modulator, into the output waveguide. It is more difficult to couple out of the modulator than into the modulator. This is because the Ge has a higher refractive index than Si (4.2 as compared to 3.5) and a larger cross-section. Therefore, it is more difficult to taper the Ge into a small enough cross-section to match the effective index of the Ge and the Si waveguide. The scattering caused by incomplete coupling is seen by the increased power in the Ge, at the end of the modulator taper, in (c) as compared to (d). The mode is beating back and forth between the two tapers and, for efficient coupling, the taper overlap must be designed such that the taper ends while the mode is in the output waveguide. However, since the taper tip is narrow, the reflections caused by effective index shift are small, and coupling efficiencies of greater than 94% are shown for a broad range of taper overlaps.

An additional effect is observed within the local minimum and local maximum in coupling efficiency, showed in parts (c) and (d) of Figure 3.10. The Ge has a large cross-section, and therefore allows for the propagation of multiple modes. However, the Si waveguide is designed to support a single mode. In the output coupling of figure (d), the light is coupled into the fundamental mode in the output waveguide. However, in figure (c), there is clear indication of higher-order modes in the output waveguide. This is indicated by the lateral oscillations in the mode, caused by modal interference. When multiple modes exist within a waveguide, there can be increased propagation losses, caused by increased interaction with waveguide sidewall roughness. In addition, there can be modal dispersion. This is caused by the variation in

propagation velocities of the different modes, since different modes experience a different effective index. Modal dispersion would cause pulse width spreading, and therefore reduce maximum bandwidths that can propagate within a single waveguide. Therefore, the coupling scheme shown in (d) is superior to (c). Not only does it exhibit increased coupling efficiency, but it also couples only a single mode to the output waveguide, which allows for higher bandwidths.

Evanescent coupling has proven to be an efficient, low loss coupling technique. With design optimization, overlapping tapers, that use evanescent coupling, can yield nearly perfect coupling efficiencies. This is due to the effective index matching that occurs when two waveguides are inversely tapered. If light needs to be coupled between waveguides with a large refractive index difference, then small taper dimensions may be required, which require deep ultra violet lithography in order to fabricate. If coupling losses are a significant concern, then evanescent coupling can be an optimal solution. Since evanescent coupling facilitates low-loss coupling in both lateral and vertical directions, it is very useful for the dense 3D integration scheme proposed for back-end photonic interconnects.

Chapter 4: Summary and Future Work

This thesis focuses on the integration of photonic interconnects with Si CMOS microprocessors. The ultimate goal is to realize electronic-photonic integrated circuits in which the photonic interconnects replace metals for global interconnects in an integrated circuit. The approach is to integrate photonic devices into the interconnect stack, thereby leaving the valuable substrate real estate to be used for dense transistor integration. This architectures imposes back-end processing limitations such as a maximum thermal budget of 450 °C, and a lack of a crystalline substrate.

In Chapter 1, a motivation was discovered for introducing electronic-photonic integrated circuits. It was found that photonic interconnects have the potential to increase the bandwidth density, while also decreasing the power consumption. Bandwidth density can be increased by implementation of wavelength division multiplexing. It was determined that power consumption benefits are only attained at longer interconnect distances. Therefore, the target application of photonics is to replace global interconnects in a Si microprocessor. In order to save valuable substrate real estate for Si transistors, the target photonic integration will be within the interconnect stack. Therefore, the challenge is to create high crystal quality Ge for use in back end optoelectronic devices.

In Chapter 2, techniques for depositing large grain Ge, at low temperature, were explored. It was shown that a two dimensional geometrically confined lateral growth is capable of creating large-grain Ge on an amorphous substrate while maintaining processing temperatures at or below 450 °C. This technique was then utilized, in a novel design, to fill a lithographically defined oxide trench with large-grain Ge. This was achieved by arraying the 2D GCLG structures, and aligning them such that the channel opens into the trench. The emerging crystalline Ge then

seeds further epitaxial growth until the entire trench has been filled. Then, the fabrication limitations with this design were pointed out. Namely, the inconsistent TMAH undercut etch of a-Si, within small channels, yields uncontrolled and unpredictable channel definition. Finally, a design solution was proposed to lithographically define all dimensions of the growth channel, eliminating the timed TMAH etch. A process flow was presented to realize the enhanced design.

In Chapter 3, designs for evanescently coupled devices were simulated and analyzed. Fist, vertical couplers with overlapping tapers were examined. The overlapping tapers ensure that there is a point along the coupler in which the effective index of both the input and output waveguides are matched. At this point, the optical impedance of each waveguide is matched and efficient coupling occurs. It was shown that coupling efficiencies greater than 99% were achievable with a long taper. The general idea of evanescent coupling from overlapping tapers was then applied to a coupling scheme for a Ge-based electro-absorption modulator. The coupling between the waveguide and modulator was evanescent, lateral coupling. With the correct design, the coupling losses were shown to be negligible. Therefore, the total insertion loss of the modulator was reduced from 3.7 dB to 2 dB, with coupling losses reduced to 0.04 dB. The remaining insertion loss was due to indirect band absorption, and is limited by material properties, not device design. Both coupling designs were shown to have a length-dependent coupling efficiency, which was attributed to modal beating between waveguides.

A significant amount of future work remains in order for this approach to be utilized in actual optoelectronic devices. For complete integration of photonics in the back-end, the Introduction mentions three significant barriers that must be overcome: (1) how to fabricate high quality Ge for active regions of optoelectronic devices while adhering to back-end processing constraints, (2) how to couple optical power to these devices, and (3) how to metalize and

electrically contact these devices. The first step is to fabricate the proposed trench filling design. These devices are already in the process of being fabricated, and results will hopefully show increased yield and control, thereby fulfilling the first requirement. The second requirement was addressed in simulations, showing that evanescent coupling is a technique capable of low-loss optical power transfer between devices. Therefore, the third barrier must be addressed. Metallization of the Ge-filled trenches can yield initial photodetectors. This can be achieved by depositing interdigitated metal contacts onto the Ge-filled trench, thereby creating a photoconductor, or a metal-semiconductor-metal (MSM) photodetector. These detector designs are easy to fabricate and have low capacitances and therefore can have very high bandwidths. These would be the first Ge based optoelectronic devices to be based on large grain, high quality, Ge, while still adhering to back-end processing limitations.

There is still a significant amount of work to be done in order to realize electronic-photonic integrated circuits with back-end photonics. However, this thesis lays the groundwork as to how high crystal quality Ge optoelectronic devices can be fabricated in the back-end. It also presents low loss coupling designs that can be used for dense 3D integration of these photonic devices.

References

[1] D. Ingerly, S. Agraharam, D. Becher, V. Chikarmane, K. Fischer, R. Grover, M. Goodner, S. Haight, J. He, T. Ibrahim, S. Joshi, H. Kothari, K. Lee, Y. Lin, C. Litteken, H. Liu, E. Mays, P. Moon, T. Mule, S. Nolen, N. Patel, S. Pradhan, J. Robinson, P. Ramanarayanan, S. Sattiraju, T. Schroeder, S. Williams, and P. Yashar, "Low-K Interconnect Stack with Thick Metal 9 Redistribution Layer and Cu Die Bump for 45nm High Volume Manufacturing," in *Interconnect Technology Conference, 2008. IITC 2008. International*, 2008, pp. 216 –218.

[2] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-power dissipation in a microprocessor," in *Proceedings of the 2004 international workshop on System level interconnect prediction*, New York, NY, USA, 2004, pp. 7–13.

[3] D. Miller, "Device Requirements for Optical Interconnects to Silicon Chips," *Proceedings of the IEEE*, vol. 97, no. 7, pp. 1166 –1185, Jul. 2009.

[4] "International Technology Roadmap for Semiconductors." 2005.

[5] D. J. Miller, P. M. Watts, and A. W. Moore, *Motivating Future Interconnects: A Differential Measurement Analysis of PCI Latency*. .

[6] L. Kimerling, L. Negro, S. Saini, Y. Yi, D. Ahn, S. Akiyama, D. Cannon, J. Liu, J. Sandland, D. Sparacin, J. Michel, K. Wada, and M. Watts, "Monolithic Silicon Microphotonics," in *Silicon Photonics*, vol. 94, Springer Berlin / Heidelberg, 2004, pp. 1999–1999.

[7] D. A. B. Miller and H. M. Ozaktas, "Limit to the Bit-Rate Capacity of Electrical Interconnects from the Aspect Ratio of the System Architecture," *Journal of Parallel and Distributed Computing*, vol. 41, no. 1, pp. 42–52, Feb. 1997.

[8] M. Haurylau, G. Chen, H. Chen, J. Zhang, N. A. Nelson, D. H. Albonesi, E. G. Friedman, and P. M. Fauchet, "On-Chip Optical Interconnect Roadmap: Challenges and Critical Directions," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1699 –1705, Dec. 2006.

[9] K.-H. Koo, H. Cho, P. Kapur, and K. C. Saraswat, "Performance Comparisons Between Carbon Nanotubes, Optical, and Cu for Future High-Performance On-Chip Interconnect Applications," *IEEE Transactions on Electron Devices*, vol. 54, no. 12, pp. 3206 –3215, Dec. 2007.

[10] R. E. Camacho-Aguilera, Y. Cai, N. Patel, J. T. Bessette, M. Romagnoli, L. C. Kimerling, and J. Michel, "An electrically pumped germanium laser," *Opt. Express*, vol. 20, no. 10, pp. 11316–11320, May 2012.

[11] J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L. C. Kimerling, and J. Michel, "Waveguide-integrated, ultralow-energy GeSi electro-absorption modulators," *Nature Photonics*, vol. 2, no. 7, pp. 433–437, May 2008.

[12] D. Ahn, C. Hong, J. Liu, W. Giziewicz, M. Beals, L. C. Kimerling, J. Michel, J. Chen, and F. X. K[illegible]rtner, "High performance, waveguide integrated Ge photodetectors," *Opt. Express*, vol. 15, no. 7, pp. 3916–3921, Apr. 2007.

[13] D. K. Sparacin, R. Sun, A. M. Agarwal, M. A. Beals, J. Michel, L. C. Kimerling, T. J. Conway, A. T. Pomerene, D. N. Carothers, M. J. Grove, D. M. Gill, M. S. Rasras, S. S. Patel, and A. E. White, "Low-Loss Amorphous Silicon Channel Waveguides for Integrated Photonics," in *3rd IEEE International Conference on Group IV Photonics, 2006*, 2006, pp. 255–257.

[14] R. K. Mueller, "Current Flow across Grain Boundaries in n-Type Germanium. II," *Journal of Applied Physics*, vol. 32, no. 4, pp. 640–645, Apr. 1961.

[15] W. E. Taylor, N. H. Odell, and H. Y. Fan, "Grain Boundary Barriers in Germanium," *Phys. Rev.*, vol. 88, no. 4, pp. 867–875, Nov. 1952.

[16] M. Imaizumi, T. Ito, M. Yamaguchi, and K. Kaneko, "Effect of grain size and dislocation density on the performance of thin film polycrystalline silicon solar cells," *Journal of Applied Physics*, vol. 81, no. 11, pp. 7635–7640, Jun. 1997.

[17] J.-S. Park, J. Bai, M. Curtin, B. Adekore, M. Carroll, and A. Lochtefeld, "Defect reduction of selective Ge epitaxy in trenches on Si(001) substrates using aspect ratio trapping," *Applied Physics Letters*, vol. 90, no. 5, pp. 052113–052113–3, Feb. 2007.

[18] T. A. Langdo, C. W. Leitz, M. T. Currie, E. A. Fitzgerald, A. Lochtefeld, and D. A. Antoniadis, "High quality Ge on Si by epitaxial necking," *Applied Physics Letters*, vol. 76, no. 25, pp. 3700–3702, Jun. 2000.

[19] J. Liu, R. Camacho-Aguilera, J. T. Bessette, X. Sun, X. Wang, Y. Cai, L. C. Kimerling, and J. Michel, "Ge-on-Si optoelectronics," *Thin Solid Films*, vol. 520, no. 8, pp. 3354–3360, Feb. 2012.

[20] K. Wada, H. C. Luan, D. R. C. Lim, and L. C. Kimerling, "On-chip interconnection beyond semiconductor roadmap: Silicon microphotonics," in *Proc. SPIE*, 2002, vol. 4870, pp. 437–443.

[21] M. Beals, J. Michel, J. F. Liu, D. H. Ahn, D. Sparacin, R. Sun, C. Y. Hong, L. C. Kimerling, A. Pomerene, D. Carothers, J. Beattie, A. Kopa, A. Apsel, M. S. Rasras, D. M. Gill, S. S. Patel, K. Y. Tu, Y. K. Chen, and A. E. White, "Process flow innovations for photonic device integration in CMOS," *Proceedings of SPIE*, vol. 6898, no. 1, pp. 689804–689804–14, Feb. 2008.

[22] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 Â¿m," *Electronics Letters*, vol. 15, no. 4, pp. 106 –108, 1979.

[23] J. Schmidtchen, A. Splett, B. Schuppert, K. Petermann, and G. Burbach, "Low loss singlemode optical waveguides with large cross-section in silicon-on-insulator," *Electronics Letters*, vol. 27, no. 16, pp. 1486 –1488, Aug. 1991.

[24] M. Lipson, "Guiding, modulating, and emitting light on Silicon-challenges and opportunities," *Journal of Lightwave Technology*, vol. 23, no. 12, pp. 4222 – 4238, Dec. 2005.

[25] M. L. Lee, E. A. Fitzgerald, M. T. Bulsara, M. T. Currie, and A. Lochtefeld, "Strained Si, SiGe, and Ge channels for high-mobility metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 97, no. 1, p. 011101, 2005.

[26] K. A. McComber, X. Duan, J. Liu, J. Michel, and L. C. Kimerling, "Single-Crystal Germanium Growth on Amorphous Silicon," *Advanced Functional Materials*, vol. 22, no. 5, pp. 1049–1057, Mar. 2012.

[27] S. K. Stanley, S. S. Coffee, and J. G. Ekerdt, "Interactions of germanium atoms with silica surfaces," *Applied Surface Science*, vol. 252, no. 4, pp. 878–882, Nov. 2005.

[28] C. Wild, N. Herres, and P. Koidl, "Texture formation in polycrystalline diamond films," *Journal of Applied Physics*, vol. 68, no. 3, pp. 973–978, Aug. 1990.

[29] Z. Gai, W. S. Yang, R. G. Zhao, and T. Sakurai, "Macroscopic and nanoscale faceting of germanium surfaces," *Phys. Rev. B*, vol. 59, no. 23, pp. 15230–15239, Jun. 1999.

[30] K. A. McComber, "Single-crystal germanium growth on amorphous silicon," Massachusetts Institute of Technology, 2011.

[31] J. Narayan and A. S. Nandedkar, "Atomic structure and energy of grain boundaries in silicon, germanium and diamond," *Philosophical Magazine Part B*, vol. 63, no. 5, pp. 1181–1192, 1991.

[32] G. Masini, L. Colace, and G. Assanto, "2.5 Gbit/s polycrystalline germanium-on-silicon photodetector operating from 1.3 to 1.55 μm," *Applied Physics Letters*, vol. 82, no. 15, pp. 2524–2526, Apr. 2003.

[33] S. Assefa, F. Xia, S. W. Bedell, Y. Zhang, T. Topuria, P. M. Rice, and Y. A. Vlasov, "CMOS-integrated high-speed MSM germanium waveguide photodetector," *Opt. Express*, vol. 18, no. 5, pp. 4986–4999, Mar. 2010.

[34] L. Chen and M. Lipson, "Ultra-low capacitance and high speed germaniumphotodetectors on silicon," *Opt. Express*, vol. 17, no. 10, pp. 7901–7906, May 2009.

[35] N.-N. Feng, D. Feng, S. Liao, X. Wang, P. Dong, H. Liang, C.-C. Kung, W. Qian, J. Fong, R. Shafiiha, Y. Luo, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "30GHz Ge electro-absorption modulator integrated with 3?m silicon-on-insulator waveguide," *Opt. Express*, vol. 19, no. 8, pp. 7062–7067, Apr. 2011.

[36] T. Yin, R. Cohen, M. M. Morse, G. Sarid, Y. Chetrit, D. Rubin, and M. J. Paniccia, "31 GHz Ge n-i-p waveguide photodetectors on Silicon-on-Insulator substrate," *Opt. Express*, vol. 15, no. 21, pp. 13965–13971, Oct. 2007.

[37] L. Vivien, J. Osmond, J.-M. F□d□li, D. Marris-Morini, P. Crozat, J.-F. Damlencourt, E. Cassan, Y. Lecunff, and S. Laval, "42 GHz p.i.n Germanium photodetector integrated in a silicon-on-insulator waveguide," *Opt. Express*, vol. 17, no. 8, pp. 6252–6257, Apr. 2009.

[38] A. H. Mahan, J. Carapella, B. P. Nelson, R. S. Crandall, and I. Balberg, "Deposition of device quality, low H content amorphous silicon," *Journal of Applied Physics*, vol. 69, no. 9, p. 6728, 1991.

[39] G. Ganguly and A. Matsuda, "Defect formation during growth of hydrogenated amorphous silicon," *Phys. Rev. B*, vol. 47, no. 7, pp. 3661–3670, Feb. 1993.

[40] Y. Q. Fu, J. K. Luo, S. B. Milne, A. J. Flewitt, and W. I. Milne, "Residual stress in amorphous and nanocrystalline Si films prepared by PECVD with hydrogen dilution," *Materials Science and Engineering: B*, vol. 124–125, no. 0, pp. 132–137, Dec. 2005.

[41] K. R. Williams, K. Gupta, and M. Wasilik, "Etch rates for micromachining processing-part II," *Journal of Microelectromechanical Systems*, vol. 12, no. 6, pp. 761–778, Dec. 2003.

[42] K. R. Williams and R. S. Muller, "Etch rates for micromachining processing," *Microelectromechanical Systems, Journal of*, vol. 5, no. 4, pp. 256 –269, Dec. 1996.

[43] P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. V. Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. V. Thourhout, and R. Baets, "Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography," *IEEE Photonics Technology Letters*, vol. 16, no. 5, pp. 1328 –1330, May 2004.

[44] F. Grillot, L. Vivien, S. Laval, D. Pascal, and E. Cassan, "Size influence on the propagation loss induced by sidewall roughness in ultrasmall SOI waveguides," *IEEE Photonics Technology Letters*, vol. 16, no. 7, pp. 1661 –1663, Jul. 2004.

[45] S. Lardenois, D. Pascal, L. Vivien, E. Cassan, S. Laval, R. Orobtchouk, M. Heitzmann, N. Bouzaida, and L. Mollard, "Low-loss submicrometer silicon-on-insulator rib waveguides and corner mirrors," *Opt. Lett.*, vol. 28, no. 13, pp. 1150–1152, Jul. 2003.

[46] R. Sun, K. McComber, J. Cheng, D. K. Sparacin, M. Beals, J. Michel, and L. C. Kimerling, "Transparent amorphous silicon channel waveguides with silicon nitride intercladding layer," *Applied Physics Letters*, vol. 94, no. 14, p. 141108, 2009.

[47] S. Assefa, F. Xia, W. M. J. Green, C. L. Schow, A. V. Rylyakov, and Y. A. Vlasov, "CMOS-Integrated Optical Receivers for On-Chip Interconnects," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 16, no. 5, pp. 1376 –1385, Oct. 2010.

[48] R. Sun, M. Beals, A. Pomerene, J. Cheng, C. Hong, L. Kimerling, and J. Michel, "Impedance matching vertical optical waveguide couplers for dense high index contrast circuits," *Opt. Express*, vol. 16, no. 16, pp. 11682–11690, 2008.

[49] A. Yariv, "Coupled-mode theory for guided-wave optics," *IEEE Journal of Quantum Electronics*, vol. 9, no. 9, pp. 919 – 933, Sep. 1973.

[50] R. Swanepoel, "Determination of surface roughness and optical constants of inhomogeneous amorphous silicon films," *Journal of Physics E: Scientific Instruments*, vol. 17, no. 10, pp. 896–903, Oct. 1984.

[51] K. Huang, S. Yang, and L. Tong, "Modeling of evanescent coupling between two parallel optical nanowires," *Appl. Opt.*, vol. 46, no. 9, pp. 1429–1434, Mar. 2007.

[52] D. Marcuse, "Directional couplers made of nonidentical asymmetric slabs. Part I: Synchronous couplers," *Journal of Lightwave Technology*, vol. 5, no. 1, pp. 113 – 118, Jan. 1987.

[53] K. A. Latunde-Dada and F. P. Payne, "Theory and Design of Adiabatically Tapered Multimode Interference Couplers," *J. Lightwave Technol.*, vol. 25, no. 3, pp. 834–839, Mar. 2007.

[54] B. E. A. Saleh and M. C. Teich, *Fundamentals of photonics*. Wiley-Interscience, 2007.

[55] O. Kjebon, R. Schatz, S. Lourdudoss, S. Nilsson, B. Stalnacke, and L. Backbom, "30 GHz direct modulation bandwidth in detuned loaded InGaAsP DBR lasers at 1.55 mu;m wavelength," *Electronics Letters*, vol. 33, no. 6, pp. 488 –489, Mar. 1997.

[56] D. Klotzkin, K.-C. Syao, P. Bhattacharya, C. Caneau, and R. Bhat, "Modulation characteristics of high speed ($f_{3\ dB}$=20 GHz) tunneling injection InP/InGaAsP 1.55 mu;m ridge waveguide lasers extracted from optical and electrical measurements," *Journal of Lightwave Technology*, vol. 15, no. 11, pp. 2141 –2146, Nov. 1997.

[57] S. K. Hwang, J. M. Liu, and J. K. White, "35-GHz intrinsic bandwidth for direct modulation in 1.3- mu;m semiconductor lasers subject to strong injection locking," *IEEE Photonics Technology Letters*, vol. 16, no. 4, pp. 972 –974, Apr. 2004.

[58] K. Sato, S. Kuwahara, and Y. Miyamoto, "Chirp Characteristics of 40-Gb/s Directly Modulated Distributed-Feedback Laser Diodes," *J. Lightwave Technol.*, vol. 23, no. 11, p. 3790, Nov. 2005.

[59] X. Zhao, D. Parekh, E. K. Lau, H.-K. Sung, M. C. Wu, W. Hofmann, M. C. Amann, and C. J. Chang-Hasnain, "Novel cascaded injection-locked 1.55-?m VCSELs with 66 GHz modulation bandwidth," *Opt. Express*, vol. 15, no. 22, pp. 14810–14816, Oct. 2007.

[60] J. Liu, D. Pan, S. Jongthammanurak, K. Wada, L. C. Kimerling, and J. Michel, "Design of monolithically integrated GeSi electro-absorption modulators and photodetectors on a SOI platform," *Opt. Express*, vol. 15, no. 2, pp. 623–628, Jan. 2007.

[61] I. A. Young, B. Block, M. Reshotko, and P. Chang, "Integration of nano-photonic devices for CMOS chip-to-chip optical I/O," in *2010 Conference on Lasers and Electro-Optics (CLEO) and Quantum Electronics and Laser Science Conference (QELS)*, 2010, pp. 1–2.

[62] M. A. Taubenblatt, "Optical Interconnects for High-Performance Computing," *J. Lightwave Technol.*, vol. 30, no. 4, pp. 448–457, Feb. 2012.

[63] A. Mekis, S. Gloeckner, G. Masini, A. Narasimha, T. Pinguet, S. Sahni, and P. De Dobbelaere, "A Grating-Coupler-Enabled CMOS Photonics Platform," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 3, pp. 597–608, May 2011.

[64] A. Narasimha, B. Analui, E. Balmater, A. Clark, T. Gal, D. Guckenberger, S. Gutierrez, M. Harrison, R. Ingram, R. Koumans, D. Kucharski, K. Leap, Y. Liang, A. Mekis, S. Mirsaidi, M. Peterson, T. Pham, T. Pinguet, D. Rines, V. Sadagopan, T. J. Sleboda, D. Song, Y. Wang, B. Welch, J. Witzens, S. Abdalla, S. Gloeckner, and P. De Dobbelaere, "A 40-Gb/s QSFP Optoelectronic Transceiver in a 0.13 #x003BC;m CMOS Silicon-on-Insulator Technology," in *Conference on Optical Fiber communication/National Fiber Optic Engineers Conference, 2008. OFC/NFOEC 2008*, 2008, pp. 1 –3.

[65] D. G. Girton, S. L. Kwiatkowski, G. F. Lipscomb, and R. S. Lytel, "20 GHz electro-optic polymer Mach–Zehnder modulator," *Applied Physics Letters*, vol. 58, no. 16, pp. 1730–1732, Apr. 1991.

[66] W. M. Green, M. J. Rooks, L. Sekaric, and Y. A. Vlasov, "Ultra-compact, low RF power, 10 Gb/s siliconMach-Zehnder modulator," *Opt. Express*, vol. 15, no. 25, pp. 17106–17113, Dec. 2007.

[67] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. D. Keil, and T. Franck, "High speed silicon Mach-Zehnder modulator," *Optics Express*, vol. 13, no. 8, p. 3129, 2005.

[68] K. Tsuzuki, T. Ishibashi, T. Ito, S. Oku, Y. Shibata, R. Iga, Y. Kondo, and Y. Tohmori, "40 Gbit/s n-i-n InP Mach-Zehnder modulator with a pi; voltage of 2.2 V," *Electronics Letters*, vol. 39, no. 20, pp. 1464 – 1466, Oct. 2003.

[69] Y. Li, L. Zhang, M. Song, B. Zhang, J.-Y. Yang, R. G. Beausoleil, A. E. Willner, and P. D. Dapkus, "Coupled-ring-resonator-based silicon modulator for enhanced performance," *Opt. Express*, vol. 16, no. 17, pp. 13342–13348, Aug. 2008.

[70] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Opt. Express*, vol. 15, no. 2, pp. 430–436, Jan. 2007.

[71] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature*, vol. 435, no. 7040, pp. 325–327, May 2005.

[72] Y.-H. Kuo, Y. K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller, and J. S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," *Nature*, vol. 437, no. 7063, pp. 1334–1336, Oct. 2005.

[73] J. E. Roth, O. Fidaner, R. K. Schaevitz, Y.-H. Kuo, T. I. Kamins, J. S. Harris, and D. A. B. Miller, "Optical modulator on silicon employing germanium quantum wells," *Opt. Express*, vol. 15, no. 9, pp. 5851–5859, Apr. 2007.

[74] N.-N. Feng, S. Liao, D. Feng, X. Wang, P. Dong, H. Liang, C.-C. Kung, W. Qian, Y. Liu, J. Fong, R. Shafiiha, Y. Luo, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "Design and fabrication of 3?m silicon-on-insulator waveguide integrated Ge electro-absorption modulator," *Opt. Express*, vol. 19, no. 9, pp. 8715–8720, Apr. 2011.

[75] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," *Nature Photonics*, vol. 4, no. 8, pp. 518–526, 2010.

[76] P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiiha, C.-C. Kung, W. Qian, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Low Vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator," *Opt. Express*, vol. 17, no. 25, pp. 22484–22490, Dec. 2009.

[77] M. R. Watts, D. C. Trotter, R. W. Young, and A. L. Lentine, "Ultralow power silicon microdisk modulators and switches," in *2008 5th IEEE International Conference on Group IV Photonics*, 2008, pp. 4 –6.

[78] V. Raghunathan, J. Hu, W. N. Ye, J. Michel, and L. C. Kimerling, "Athermal Silicon Ring Resonators," in *Integrated Photonics Research, Silicon and Nanophotonics*, 2010, p. IMC5.

[79] D. M. Gill, M. Rasras, K.-Y. Tu, Y.-K. Chen, A. E. White, S. S. Patel, D. Carothers, A. Pomerene, R. Kamocsai, C. Hill, and J. Beattie, "Internal Bandwidth Equalization in a

CMOS-Compatible Si-Ring Modulator," *IEEE Photonics Technology Letters*, vol. 21, no. 4, pp. 200 –202, Feb. 2009.

[80] J. Liu, L. C. Kimerling, and others, "GeSi photodetectors and electro-absorption modulators for Si electronic-photonic integrated circuits," Massachusetts Institute of Technology, 2007.