Silicon’s extremely high refractive index and low loss at telecommunications frequencies make it possible for silicon waveguides to confine optical modes to sub-diffraction-limited areas. We can gain control of photons on the nanometer scale, and can force strong interactions with nonlinear waveguide claddings. Taking advantage of these capabilities, it is possible to integrate femtosecond-scale nonlinear optical functionality into silicon chip-scale devices.

Under the AFOSR YIP program, we worked to build practical silicon-organic hybrid nonlinear optical devices, including low-voltage modulators, and to build mid-infrared integrated photonic devices in silicon. In addition, a number of side projects on biosensing and optomechanics bore significant fruit in terms of high-profile papers. This program produced a number of world-record results, as well as 13 papers in peer-reviewed journals.
Award Number FA9550-08-1-0101
AFOSR YIP
Title: Nanophotonic Devices In Silicon for Nonlinear Optics
PI: Michael Hochberg

Silicon’s extremely high refractive index and low loss at telecommunications frequencies make it possible for silicon waveguides to confine optical modes to sub-diffraction-limited areas. We can gain control of photons on the nanometer scale, and can force strong interactions with nonlinear waveguide claddings. Taking advantage of these capabilities, it is possible to integrate femtosecond-scale nonlinear optical functionality into silicon chip-scale devices.

Under the AFOSR YIP program, we worked to build practical silicon-organic hybrid nonlinear optical devices, including low-voltage modulators, and to build mid-infrared integrated photonic devices in silicon. In addition, a number of side projects on biosensing and optomechanics bore significant fruit in terms of high-profile papers. This program produced a number of world-record results, as well as 13 papers in peer-reviewed journals.

Key Technical outcomes:
- Demonstration of the first silicon-organic hybrid modulators operating at RF speeds
- Demonstration of the first sub-volt silicon-organic EO-modulator operating at RF speeds
- Demonstration of a variety of low-loss passive optical components, such as strip-to-slot converters, at world-record performance
- Demonstration of world’s lowest loss slot waveguides, made in a DOD-trusted foundry (BAE Systems)
- Design study showing tradeoffs between loss, bandwidth and voltage for EO silicon hybrid organic modulators using traveling wave electrodes. This is turning out to be a very highly-cited design study.
- Design study on analog links using the above modulators.
- Demonstration of the first silicon waveguides for the mid-infrared. This was one of the initial results in what is turning out to be an active sub-field. We are now working with BAE to explore DOD relevant applications of this technology.
- Demonstration of the world’s lowest-voltage EO modulator, at .25Vπi
- First demonstration of ultra-low-noise displacement sensing in optomechanical waveguides on-chip. Published in Nature.

Key Non-technical Outcomes:
- This work resulted in the formation of Portage Bay Photonics, which is negotiating a final license agreement with UW for the relevant patent portfolio. PBP has successfully won two STTR programs related to transitioning this technology into DOD platforms.
The results generated under this YIP program resulted in follow-on funding through the PECASE program, which is currently underway and producing a stream of world-leading results.

A postdoctoral scholar funded in this program won a prestigious faculty position at the University of Aachen.

This work has resulted in an extremely close collaboration with BAE Systems and with Intel, both of which are likely transition partners.

The initial results from this program helped seed the creation of OPSIS, which is a world-leading silicon photonic foundry service. Without this YIP funding, these results would not have been possible.

The work in this program was published in high prestige journals, including Nature Photonics, Nature, and as a cover article in JPCC.

**Publications associated with this program:**


Witzens and M. Hochberg. Optical detection of target molecule induced aggregation of nanoparticles by means of high-Q resonators. Optics Express 2011; 19: 7034-7061


J. Witzens, T. Baehr-Jones and M. Hochberg. Design of transmission line driven slot waveguide Mach-Zehnder interferometers and application to analog optical links. Optics Express 2010; 18, 16: 16902-16928


T. Baehr-Jones, M. Hochberg, A. Scherer , Photodetection in silicon beyond the band edge with surface states, Optics Express 16, 1659-1668 (2008)


**Project details**

There are two major thematic areas where significant progress was made as part of this grant. The first was in silicon-organic hybrid device engineering and design, the second was in mid-infrared photonics technology. While there were a number of results that benefitted from this program’s support outside these areas (most notably work on biophotonics and on optomechanics) these results were largely peripheral to the major goals of the work.

**Silicon-Organic Hybrid Devices**

In the area of silicon-organic modulator devices, we pursued a program where we progressed from theoretical studies on the design of traveling-wave EO modulators to actually demonstrating some world-leading devices. Our initial work on this topic began at Caltech, during my time as a graduate student. In 2005, we demonstrated the world’s first slot waveguide EO modulators, and we developed a theoretical paper in 2007 showing that silicon/organic hybrid devices provide a path to millivolt scale EO modulators.

Under this program, we pushed forward the state of the art in terms of the performance of all the components of the modulators: We worked with BAE to radically reduce the
waveguide losses of silicon slot guides, to the point where they are now competitive with lithium niobate. We worked with BAE and LETI to improve the coupling losses through both Y-junctions and through slot-to-ridge transformers. We published what is becoming a highly-cited design study on the details of how to build traveling-wave EO modulators in silicon-organic systems. Early in the program, we demonstrated low-speed modulators operating at a Vpi value of .25 Volts, which remains the lowest value in the literature to our knowledge. And finally we demonstrated the device described below, which represents the current state of the art for silicon organic hybrid devices: A modulator working at RF speeds, but with less than 1V drive voltage. Work continues, under the PECASE program, on improving these devices further.

Lowering the operating voltage of electrooptic modulators is desirable for a variety of applications, most notably in analog photonics and digital data communication. In particular for digital systems such as CPUs, it is desirable to develop modulators that are both temperature-insensitive and compatible with typically sub-2V CMOS electronics; however, drive voltages in silicon-based MZIs currently exceed 1.8V. We have shown an MZI modulator based on an electrooptic polymer-clad silicon slot waveguide, with a halfwave voltage of only 0.69V (corresponding to a 0.62 V-cm modulation figure of merit), and a bandwidth of 500 MHz. We also show that there are paths to significantly improve both the bandwidth and drive voltage.

I. INTRODUCTION

The first silicon MZI modulator was based on carrier-depletion in a p-n structure [1], with bandwidth around 1 GHz and a modulation figure of merit of 8 V-cm. Improvements in design have led to silicon MZI modulators that exhibit modulation figures of merit of 4 V-cm at 30 GHz [2], and more recently 1.4 V-cm at 12 GHz [3]. But in the latter case, the authors identify a fundamental tradeoff between carrier concentration and modulator performance: that figure of merit can only be obtained with around 19 dB/cm of intrinsic absorption loss. It may never be possible to build a practical MZI modulator in silicon based on carrier depletion with a halfwave voltage of less than 1 Volt. A similar limitation has been encountered with Lithium Niobate based modulators; at speeds near 20 GHz, halfwave voltages are typically near 2.7 V or higher, while even at speeds as low as 1 GHz, halfwave voltages are still 1.2V or more [4],[5]. Typical Lithium Niobate device lengths for these halfwave voltages are 5 cm or longer. With forward-biased diode based silicon MZIs [6], the halfwave voltage can be lowered to 1.8V, but the actual drive voltages for 10 GBit/s RF bitrates are around 7.6V [7]. At lower bitrates, such modulators can achieve lower drive voltages of 1.8V [8]. Resonant enhancement [9],[10], and electroabsorption modulators [11] can reduce drive voltages, but introduce other limitations, and would be unsuitable for use in analog links [12] or as phase modulators [13], [14].

Slot waveguides were first proposed as a means to focus the propagating optical mode outside the silicon [15]. Later, it was demonstrated that by electrically contacting both arms of a slot waveguide, and coating the waveguide with an electrooptic polymer [16] cladding, a particularly responsive modulator could be created [17]; in this device, the
performance of the organic active material was enhanced by the RF and optical mode confinement of the slot waveguide. Slot-waveguide polymer MZI modulators have since been demonstrated with halfwave voltages of 0.25V [18], though only at quasi-DC speeds. Another approach combines an electrooptic polymer-clad slot waveguide with a photonic crystal [19]-[22]. We have recently demonstrated that slot-waveguide polymer MZI modulators can work at RF speeds [23].

Here we demonstrate, for the first time, a silicon-organic modulator with a low absolute drive voltage at RF speeds. A 9 mm MZI modulator results in a halfwave voltage of 0.69V with a bandwidth of 500 MHz; this corresponds to a modulator figure of merit of 0.62 V-cm. The modulator reported here is the first silicon photonic device to demonstrate an absolute operating voltage compatible with the low operating voltage of modern CPUs: The next-best results are the forward-biased diode based modulators at 1.8V at 8 GBit/s [8] and 6.5V at 10 GBit/s [7]. Additional work will be required to make the slot-waveguide modulator competitive to the forward-biased diode approach in speed, though we feel there is a clear path towards doing so [22]; in fact, a 40 Gbit/s polymer based modulator has been recently demonstrated although with high $V_\pi$ voltage [24]. It is also not clear that the modulators based on the competing approach could necessarily have their drive voltages decreased substantially, simply by lowering the operating speeds. Insertion loss will be one limitation – in both of the results mentioned above [7,8], on-chip losses are already around 13 dB, which will limit the length of the device. A greater limitation is the device drive impedance; both of the forward-biased diode modulators have impedances around 50Ω, and increasing the device length to decrease the drive voltage will lower this impedance. Driving a modulator with an impedance substantially lower than 50Ω will likely create difficulties, and may not be practical in many cases.

II. Device Fabrication

Silicon waveguides were fabricated with two steps of self-aligned photolithography with a 193 nm stepper and dry etching. The initial substrate was SOI, with a 220 nm thick silicon layer on a 2000 nm oxide layer, on a silicon handle with resistivity of 10Ω-cm. The wafer was diced, and further fabrication and testing occurred on individual dies. A blanket implant was performed, with a target concentration of $3 \times 10^{17}$ cm$^{-3}$ phosphorous across the entire chip. A masked implant with final concentration target $10^{20}$ cm$^{-3}$ phosphorous was done for pad contacts, with contact photolithography, followed by metallization with a 10 µm clearance from the slot waveguide to the metal. The 10 µm clearance was used to ease alignment tolerances, and could be significantly decreased in future devices. The metal layer consisted of 2 µm of aluminum topped with 10 nm of gold. Final device layout, illustrated in Figure 1, consisted of a 9 mm MZI, formed with a strip-loaded slot waveguide and two 10 µm wide metal contacts, optically coupled via grating couplers [25]. The slot size was 200nm, and the two sides of the slot waveguide had widths 230 nm and height 220 nm, while the strip-loaded height was 68 nm. Butt-coupling was used between ridge waveguide and strip-loaded slot. Figure 2 shows the geometry and mode pattern of the contacted strip-loaded slot waveguide.
Fig. 1. Device Layout: a, A rendering of a portion of the MZI is shown, illustrating the optical path. b, The layout of the device is shown. Note that after traveling through the MZI, the optical signal returns back to the side of the MZI in a ridge waveguide. Several optical micrographs and an SEM micrograph of key features are also displayed. It should be noted that in two of the optical micrographs, portions of adjacent, unrelated devices are shown.

Fig. 2. Device Layout and Modal Structure: a, The contacted strip-loaded slot structure is shown, as well as the optical mode. A contour plot of $|E_x|$ is shown, normalized for 1 mW of propagating power. b, The RF mode is shown, with a contour plot of the voltage that would be seen at high frequencies, but below the bandwidth limit.

A path length difference of 80 µm for the two MZI arms is used; this allowed the phase shift to be measured by device transmission, as shown in Section 3.1, and enabled setting the MZI bias point by tuning the wavelength. The device dimensions are approximately 9 mm x 200 µm excluding contact pads, resulting in a total area of 1.8 mm². The device layout was driven by the need to manually probe between the waveguides, which requires
large pads. Considering only the metal leads and silicon waveguide and excluding the probe pads, the total area is 0.8 mm$^2$.

The device was spin-coated (1500 rpm) with an electrooptic polymer that consisted of PMMA doped with 14 wt% of AJLZ53 chromophore [26] in chlorobenzene, leaving a film of around 2 µm. The device was then baked at 75 °C under vacuum. This electrooptic polymer typically exhibits a refractive index around 1.53, a relative dielectric constant of 3.2 to 3.5 at RF frequencies, and optical loss on the order of 1 dB/cm for wavelengths near 1550 nm. The device was poled in push-pull configuration, with the three pads biased at −20V, ground and 20V, with a poling temperature of around 110 °C.

III. DEVICE PERFORMANCE CHARACTERIZATION

Measurements of the device were initially taken with an Agilent 81980A tunable laser, and an Agilent 81636B fast power sensor. Testing was typically performed with 13 dBm of laser power. The on-chip insertion loss of the device was measured to be, at the peak transmission, around 30 dB. This figure does not include losses due to the grating couplers, which were used to couple light on and off the chip. Due to simulations we attribute approximately 8 dB to the two Y-junctions and ridge to slot mode converters, leaving 22 dB of loss from the 9 mm of strip-loaded slot waveguide, and just over 1 cm of the connecting ridge waveguide. The $3 \times 10^{17}$ cm$^3$ phosphorous dopant should introduce 8 dB/cm bulk silicon loss at 1550 nm [27]. For a strip-loaded slot, the waveguide losses due to this implant concentration should then be approximately 4 dB/cm; the losses in a ridge waveguide are higher at around 8 dB/cm, due to differences in the optical mode patterns. In the ideal case, then, one should expect a contribution of around 4 dB of loss from the nearly cm long strip-loaded slot waveguide, and 8 dB from the connecting ridge waveguide, based solely on the implants, for a total of 12 dB of excess loss. We have measured the undoped strip-loaded slot waveguides to have losses of approximately 6.5 dB/cm [28], while the ridge waveguides have losses of approximately 4 dB/cm. Combining the intrinsic losses of both waveguide types with the excess losses from the implants, the final loss budget is in close agreement with the measured value.

There is a clear path towards reducing this optical loss. First, the intrinsic waveguide loss of the strip-loaded slot waveguide can be reduced; we have separately demonstrated losses that are better than 2 dB/cm [29]. Moreover, the nearly 12 dB of loss from the ridge waveguide could be largely eliminated in an eventual device. Since this ridge waveguide is not actually part of the modulator, the implants could simply be masked to not fall on the ridge waveguide. Moreover, the intrinsic losses from a ridge waveguide in silicon can be greatly reduced, with values of 0.8 dB/cm[30] achieved in the literature.

By using graded implants [12], the implant losses on the strip-loaded slot portion, already only 4 dB, could be lowered further. Finally, the losses from the ridge to slot mode converter can also be substantially reduced; ridge to slot mode converters with losses of only 0.81 dB have been demonstrated, and we have shown such approaches can be adapted to our process [31]. Taking these factors together, an eventual 9 mm slot waveguide modulator loss budget might look roughly like this: there would be 2 dB intrinsic loss from the slot waveguide, 2 dB excess loss from the implants, and around 3 dB from mode converters, Y-junctions and connecting ridge waveguides, for a total of 7
dB. This value is comparable to the typical insertion losses seen in commercial modulators [32].

A. DC Halfwave Voltage

To characterize modulator performance at low speeds, optical transmission spectra with several different DC bias voltages were taken, showing a fringe spacing of around 6.9 nm due to the arm length imbalance. A balanced MZI could maximize the optical bandwidth and make the device athermal, however an unbalanced device makes it possible to measure the $V_{\pi}$ by observing the transmission spectrum shift. Figure 3 shows the device transmission spectrum under two different bias voltages, as well as the phase shift deduced from the location of one of the peaks. The slight increase in insertion loss at shorter wavelengths is due to the bandwidth of the grating coupler. The DC device resistance was high, typically in excess of 100 kΩ.

![Fig. 3. DC device performance: (a) Device transmission excluding off-chip coupling losses as a function of wavelength is shown for several bias voltages. (b, c) A plot of phase shift in radians as a function of bias voltage for two different measurement sets, determined by the migration of the minimum near 1550 nm as a function of bias voltage. The slopes indicate halfwave voltages of 0.72±0.06 V and 0.65±0.03 V, respectively. The]
device transmission plots correspond to the data points of the first measurement set.

To measure the halfwave voltage with device transmission, we used the following methodology. To a constant factor, the transmission through an unbalanced MZI can be expressed as a function of wavelength as

\[
\frac{1}{2} \left( 1 + \cos \left( \frac{2}{\lambda} n(\lambda) L \pm \frac{V}{V_p} \right) \right)
\]

(1)

where \( \lambda \) is the wavelength, \( n(\lambda) \) is the effective index, which is generally a function of wavelength, \( \Delta L \) is the length difference between the arms, \( V \) is the bias voltage, and finally \( V\pi \) is the halfwave voltage. The sign in front of the bias term will be determined by the polarity of the device poling, in combination with the sign of the \( \chi_{xxx}^2 \) tensor component, where \( x \) in this case is the direction perpendicular to the direction of propagation, and parallel to the waveguide substrate. One can expand the argument to the cosine in Eq. (1) to first order in \( \lambda \) as

\[
\frac{2}{\lambda} \left( n(\lambda) \right) L \pm \frac{V}{V_p} \n_g \lambda_0
\]

(2)

where \( \lambda_0 \) for convenience is chosen to be a wavelength where for no bias, the device has a minimum in transmission, \( n_g \) is the group index of the optical mode in the waveguide. This amounts to neglecting, among other things, group velocity dispersion. The minimum in the transmission spectra will be found where Eq. (2) is equal to \( \pi \) radians. It then is immediately clear that the shift in the location of a transmission minimum will have a linear relationship to both the bias voltage and the induced phase shift. The phase shift can be expressed most conveniently in terms of the fringe-to-fringe spacing \( \Delta \lambda \).

\[
\pm \frac{V}{V_p} = 2 \left( n_g \lambda_0 \right) \frac{L}{V} = 2 \left( n_g \lambda_0 \right)
\]

(3)

To determine the location of a minimum in the transmission spectra, we selected the very lowest power level in a given region, and then averaged all points within 3 dB of this level. The variance of this distribution was then used as an uncertainty in the measurement. A halfwave voltage of 0.69±0.04V was derived as the average of the two measurement sets. When combined with a modulation figure of merit \( \gamma \) (defined in [33]) for the waveguide of 0.21 \( \mu m^{-1} \), this suggests an \( r_{33} \) value for the poled polymer of around 54 pm/V. This is slightly lower than the peak performance for the material of 65 pm/V, which we confirmed from a measurement of a single layer film on ITO.
B. High-speed performance

To characterize the device at RF speeds, the output from the device was connected to a New Focus 1647 avalanche photodetector with 1 GHz bandwidth, and S-parameters S21 and S11 were taken with an Agilent E8361C network analyzer. The wavelength was chosen to bias the device at the 3dB point. Figure 4 shows the predicted S21 value for the DC halfwave voltage, along with the measured S21 value. The modulation bandwidth, typically defined as the 6 dB point for an RF S21 [1],[23], is around 500 MHz. It was measured with -10 dBm RF power, to ensure that the modulator remained in the linear regime.

The bandwidth limitation seen is almost certainly not due to the electrooptic polymer; similar polymers have shown bandwidths of 165 GHz [34]. Nor is it likely that carrier transit times are the source of the limitation; the silicon acts only as a transparent conductor. Instead, the bandwidth limitation is due to RF parasitics. The equivalent circuit for both arms of the device is shown in Figure 4. The resistance due to the strip-loaded arms, $R_{sl}$, should be around $8\Omega$. From a finite element simulation, $C_{sl}$ should be approximately $1.8\,\text{pF}$. $C_s$ should be $3.1\,\text{pF}$, while $R_s$ can be calculated as $6\Omega$.

Here, there are four effects that can limit device bandwidth. First, at sufficiently high frequencies, 9 mm will no longer be a small fraction of a wavelength, and so distributed effects will likely be a limitation. But at 1 GHz, 9 mm is still under 1/20 of a wavelength, even in a dielectric material, so this is less likely to be an issue here. The second possibility is that the capacitive loading itself could become a limitation; at 1 GHz, the absolute value of the impedance of the two $C_{sl}$ capacitors is around $44\Omega$, and so even at 500 MHz, this effect is likely a limitation. Third, if $R_{sl}$ were sufficiently large, then an RC time constant between $R_s$ and $C_s$ would limit the bandwidth. And finally, $R_s$ has the potential to short circuit the modulator as the impedance across $C_s$ drops at higher speeds.
In fact, we believe that the latter three effects all play a role. Using the previously calculated values for the components in figure 4, a 6dB rolloff in S21 should be at 780 MHz. But the S11 would only be -0.8 dB in this case. Most likely Rs is larger than anticipated; at 75Ω, a 6dB rolloff at around 500 MHz is predicted with an S11 value of -3 dB, in agreement with the experimental data. In fact, measurements of test resistance structures on similarly prepared samples confirmed this view; these measurements indicated that the sheet resistance for the partially etched silicon was indeed more than 10x larger than would be predicted based solely on the implant dose. Etch-induced damage on the relatively thin strip-loaded region may explain this unexpectedly higher resistance [35]. Device speed could be improved by decreasing metal clearance, thus lowering Rs, or through a higher implant dose. High resistivity substrate wafers, which are available commercially from SOITEC, could also decrease Rs. For instance, using 1000 Ω-cm substrate could increase Rs to around 550Ω, and thus induce a negligible bandwidth limitation due to substrate coupling.

IV. CONCLUSION

To conclude, we have demonstrated a MZI modulator with a sub-1V halfwave voltage at RF bandwidth. This is to our knowledge the lowest halfwave voltage shown for any silicon-based MZI, and appears to be the lowest halfwave voltage shown in any material system for an RF electrooptic MZI modulator. Moreover, there is a clear path towards improved performance, from both better RF design [12] and more active polymers; in particular, polymers with r33 values of 300 pm/V have been demonstrated [36], which, if used in our device, could result in a halfwave voltage of 120 mV. Narrower slots would also improve performance [33]. We believe that slot-waveguide polymer modulators will play an important role in future silicon photonic systems.

REFERENCES


[29] R. Ding, Department of Electrical Engineering, University of Washington, Campus Box 352500, Seattle, WA 98195, Tom Baehr-Jones, Woo-Joong Kim, Bryan Boyko, Richard Bojko, Alexander Spott, Andrew Pomerene, Craig Hill, Wesley Reinhardt, and Michael Hochberg are preparing a manuscript to be called “Low-loss Asymmetric Strip-loaded Slot Waveguides in Silicon-on-Insulator”.


[31] Y. Liu, Department of Electrical Engineering, University of Washington, Campus Box 352500, Seattle, WA 98195, Tom Baehr-Jones, Andrew Pomerene and Michael Hochberg are preparing a manuscript to be called “Efficient Ridge to strip-loaded slot mode converter in silicon-on-insulator”


Mid-Infrared Photonics

Silicon waveguides are now widely used to guide radiation in the near-infrared, mainly in the wavelength range of 1.1 – 2.2 µm. While low-loss waveguides at longer wavelengths in silicon have been proposed, experimental realization has been elusive. Here we show that single-mode integrated silicon-on-sapphire waveguides can be used at mid-infrared wavelengths. We demonstrate waveguiding at 4.5 µm, or 2222.2 cm⁻¹, with losses of 4.3 ± 0.6 dB/cm. This result represents the first practical integrated waveguide system for the mid-infrared in silicon, and enables a range of new applications. In addition, we have more recently demonstrated similar performance for this waveguide system at 5.5 µm wavelengths, using a quantum cascade laser. This work is already generating a large number of citations, and extensive followup work in the academic community, especially in terms of nonlinear optics in the mid-infrared from faculty including, for instance, Alex Gaeta. Military applications for this technology include hyperspectral imagers, thermal cameras, and chip-scale spectroscopy systems.

1. Introduction

It has recently been shown that silicon waveguides can be used to construct all of the components of a photonic data transmission system on a single chip [1,2]. These components can be integrated together with CMOS electronics to create complex electronic-photonic integrated circuits [3]. High field confinement silicon waveguides enable exciting new applications, from chip-scale nonlinear optics [4] to biosensors [5] and light-force activated devices [6]. To date, most of the experiments in silicon waveguides have been at wavelengths in the near-infrared, ranging from 1.1-2.2 µm [7,8]. Here we show that single-mode integrated silicon-on-sapphire waveguides can be used at mid-infrared wavelengths, in particular at 4.5 µm, or 2222.2 cm⁻¹, with losses of 4.3 ± 0.6 dB/cm. This idea has been proposed in theoretical literature, but experimental realization has been elusive [9]. This result represents the first practical integrated waveguide system for the mid-infrared in silicon, and enables a range of new applications.

When building telecommunications systems, it is vital to operate at near-infrared wavelengths in order to maintain compatibility with existing systems. Many other types of optical systems operate in the near-infrared as a matter of convenience: A wide variety of commercial optical components are readily available at 1310, 1480 and 1550 nm, largely as a spinoff of the extensive commercial work in telecommunications. But these wavelengths, while convenient for telecommunications, are obviously not suitable for all applications. It is often necessary to manipulate light with wavelengths in the 2-20 µm (mid-infrared) regime, for a variety of applications.
consist of thermal imaging (2.5 µm to 15 µm wavelengths) [10], chemical bond spectroscopy [11] (which often spans from the visible to 20 µm and beyond), astronomy [12], gas sensing [13], and military applications such as missile countermeasures [14].

Historically, the mid-infrared has been a problematic region for photonics. Coherent sources have been bulky and expensive [15], or have required cryogenic cooling [16], as did common mid-infrared detectors [17]. Of course, the lack of integrated optical waveguides at these long wavelengths meant that mid-infrared systems were implemented using free-space optics. Recently, the landscape has begun to change dramatically. Inexpensive, reliable, single-mode quantum cascade lasers are now available commercially all the way to 9 µm wavelengths, with powers from 10-100 mW, offering near room-temperature operation in most cases [18,19]. Single-mode fibers are now available at wavelengths out to 6 µm [20,21], as are mid-infrared photodetectors with bandwidths over 1 GHz [22]. As a result, building a single-mode optical system in the mid-infrared is now within the financial and technical reach of any modestly well-funded research group or small company. In fact, it is possible to build a full test system with decent performance for well under $100,000, which is comparable to the cost of many swept-wavelength near-infrared systems. This combination of new capabilities strongly suggests that mid-infrared photonics is poised to take off as a field in the immediate future.

But there are many missing pieces. High bandwidth modulators do not exist for these wavelengths. Neither are there low-loss splitters, tunable filters, or any of the rest of the building blocks of complex fiber-coupled photonic systems. Silicon waveguides provide an ideal platform for us to build all of these components. Furthermore, the silicon photonic platform offers the opportunity to integrate all of these components along with control electronics on the same substrate – opening the possibility of building, for instance, FTIR systems-on-a-chip with multi-centimeter path lengths but nanoliter sample volume. The first step on this path is to construct low-loss silicon waveguides.

2. Silicon-on-sapphire material system and waveguide design

We chose to build waveguides using the silicon-on-sapphire (SOS) [23] materials system, which is used in the electronics industry as an alternative to silicon-on-insulator (SOI). SOS is particularly desirable for this application because of the lack of a high index substrate, which eliminates the issue of substrate leakage. The resistivity of the silicon for the wafers used was specified by the manufacturer to be 100 Ω-cm, suggesting that optical loss due to free-carriers will be minimal [24]. Mid-infrared guides at longer wavelengths could also be built using free-standing silicon guides [25] or germanium-on-silicon. SOS, in particular, has the advantage of offering the ability to build high-confinement, fully etched waveguides from 1.1 µm all the way to around 6.2 µm [26] – over two octaves of bandwidth, including the telecommunications region, while maintaining electronics compatibility. While SOI has recently been used to build waveguides at 2.2 µm [8], substrate leakage will become an increasingly large problem at longer wavelengths. Another problem is that silica becomes extremely lossy at wavelengths longer than 4 µm.
Fig. 1. (a) Three-dimensional rendering of one of the mid-infrared waveguide devices, showing the optical mode propagating through the waveguide. The rendering includes single-mode fibers butt-coupled to the input and output facets. (b) A contour plot of the optical mode of the waveguide is shown, with the electric field magnitude corresponding to 1 Watt of average power. (c) The dispersion diagram for the silicon-on-sapphire waveguides, calculated using a Yee grid based eigensolver. (d) A false-color scanning electron micrograph of the cleaved endfacet of a waveguide. Silicon is shown in green, and sapphire in blue.

The SOS waveguides were simulated using a Yee-grid based eigensolver [27]. It was found that a 1.8x0.6 µm ridge waveguide offered a small mode (around 1.1 µm²), shown in Fig. 1, a tight bend radius (FDTD simulations predicted 10 µm with negligible losses), and allowed the silicon etch to stop on the sapphire, simplifying fabrication. The mode size is around 1/19 of a square wavelength in free space at 4.5 µm. At this wavelength, the TM0 mode was expected to either not guide or be weakly confined, since it is near the theoretically predicted cutoff. Higher order modes are predicted to not be supported. The devices are terminated in waveguides with widths of 8 µm, which were coupled into the 1.8 µm waveguides with a taper. This configuration was used to increase coupling efficiency. Predicted coupling losses to the 1.8 µm wide guide from the single-mode mid-infrared fibers that we used were on the order of 18 dB, including the loss from the taper.

We designed a chip with a series of waveguide bends. Each waveguide started and stopped with an 8 µm wide region, which extended for 5 mm on the chip. These two regions were at 90 degrees, for the purpose of minimizing the noise due to light that might scatter directly from the source to the detector. Fig. 2 shows a more detailed picture of the layout of the chip, including the two planes on which the chip was cleaved. Both cleave planes were cleaved by hand orthogonal to the waveguides. The distance
from the cleave to the start of the taper was approximately 2.5 mm. Since it was identical for all waveguides, the insertion loss due to this region should be nearly identical for all measurements.

Some of the devices present on the chip were not used for waveguide loss measurements; the ones that were used are marked in Fig. 2. It was possible to identify which device was guiding light by observing the location of the fiber. The bend radii used were 40 µm, well in excess of the minimum bend radius of 10 µm suggested by FDTD simulations.

Fig. 2. A schematic of the chip layout is shown. A number of coiled waveguides with varying total length are used to establish waveguiding, and allow determination of the waveguide loss. The approximate location of the cleave planes is shown as well. The devices used for waveguide loss measurements are indicated with red arrows.

3. Fabrication

The waveguides were fabricated using standard semiconductor processing techniques. Epitaxial silicon-on-sapphire wafers (100 mm diameter) were used as starting material. Patterning was accomplished using an electron beam lithography system and HSQ resist on a wafer fragment. The resist was developed and the chip was etched using a CF₄ plasma in an Oxford Instruments parallel-plate RIE. Resist was first removed using just a wet resist remover, although some resist residue appeared to remain on the surface of the waveguides due to incomplete stripping. Further processing then involved cleaning the chip with a piranha etch, which appeared to improve waveguide loss dramatically, as shown in the following section. It is likely that the piranha etch removed some remnants of the resist. The chip was finally manually cleaved through the 8 µm wide waveguide segment, leaving an optical quality edge.
4. Measurements

The devices were tested using an Ekspla PL2241 Nd:YAG laser that drove an Ekspla PG501/DFG optical parametric generator/difference frequency generator (OPG/DFG). The OPG/DFG is capable of producing linearly polarized light from 2-9 µm, with 4.5 µm radiation used in our testing. We restricted our testing to this wavelength as the fiber cut off at longer wavelengths, and we were unable to achieve as great a dynamic range at other wavelengths, due to decreased emission power and beam stability from the OPG/DFG. We intend to explore the behavior of our waveguides over a range of wavelengths in future work. The laser was coupled through a polarizer into ZnSe lens with 12.7 mm focal length, and into a 9/125 µm single-mode optical fiber S009S17 from IR Photonics. The fiber, chip and detector were held on stages with piezoelectric actuated micrometers (Newport PZA12). The output of the chip was coupled directly to free space and then into a PVI-5 detector from Vigo Systems. Using a fiber for input allowed us to easily determine visually which device we were coupling radiation into, as both the fiber position and the device locations could be identified with an inspection scope.

Because the OPG/DFG provided 30 ps pulses of IR light at a repetition rate of 50 Hz with pulse energies around 150 µJ, we used a boxcar amplifier (Stanford Research SR250) to reject the signal during the times when the laser was not providing output. The signal to noise ratio was further enhanced by mechanically chopping the laser at 2 Hz, and using a Signal Recovery 7265 lock-in amplifier to detect the 2 Hz modulated signal. This resulted in an overall signal to noise ratio of 85 dB. Coupling from the free-space mode into the single-mode fiber was achieved with around 12 dB of insertion loss, leaving adequate dynamic range to perform measurements on the SOS waveguides.

In the process of testing the chip, we mechanically swept a single-mode fiber across one of the faces of one chip. We found a transmission pattern that closely matched the lithographically defined waveguide pattern, as shown in Fig. 3. By measuring a series of devices with different waveguide lengths on a chip, it was possible to determine the overall waveguide loss, and to separate this from the loss of the 8 µm waveguide regions and the insertion loss in getting the light onto and off of the chip. Devices that were tested ranged in total length from 1 mm to 1.4 cm, not including the input tapered regions, which were identical on all devices. The waveguide losses on the best chip were found to be 4.3 ± 0.6 dB/cm. A bend radius of 40 µm was used for the devices. We also included two control structures with waveguide widths of 1.2 µm (as opposed to the 1.8 µm devices, which were shown to guide), identical to other devices in all other respects. As predicted theoretically, the narrower devices did not show any transmission at all.
The light emerging from the OPG/DFG was linearly polarized with a TE polarization, that is, the optical field was parallel to the plane of the chip. To ensure that the polarization of the pulses was constant during measurements, a linear polarizer was placed in the optical path directly before the lens and the fiber, as shown in Fig. 3. Unfortunately, it was not possible to perfectly control the input polarization to the waveguides; a non-polarization maintaining fiber was used, due to the lack of a commercially available polarization-maintaining fiber at these wavelengths. We measured the polarization of the light emerging from the fiber, and found it to be TE polarized; it is likely that this was the polarization state at the output of the fiber when the
waveguides were tested. However, because the fiber was not polarization maintaining, it is impossible to be sure that this polarization state was uniformly maintained. It is likely that only the TE0 mode guides a meaningful amount of optical power, in light of the fact that the TM0 mode is so close to cutoff, and the relatively consistent results seen with initially TE polarized light entering the fiber. However, the possibility of some power propagating in the TM0 mode cannot be excluded completely; we would observe that in any case, single-mode waveguides are typically defined as supporting the TE0 and TM0 modes [28, 29]. We show the normalized optical power as a function of waveguide length for a series of devices measured on several different chips in Fig. 4.

For each device, generally eight measurements were taken at around five-second intervals. These measurements were averaged to produce the final value, while the uncertainty was determined by observing the root mean squared deviation from the mean. Typically, the standard deviation was on the order of 5%. This estimation of error is shown on the final plots of the insertion loss versus waveguide length in Fig. 4. The position of the detector was also optimized for each device measured, in order to maximize the amount of light captured from the output of the waveguide. It was found that the optimal position for the detector moved as expected as waveguides on different portions of the chip were tested.

We tested the best-performing chip multiple times, in one instance testing the shortest devices first, and in another testing the longer devices first; this was an attempt to expose any systemic errors that our test setup might have had. We found that the waveguide losses were in agreement, with the second measurement predicting losses of $4.9 \pm 0.6$ dB/cm. A second chip that had been piranha cleaned also exhibited waveguide losses of $4.7 \pm 0.6$ dB/cm that were in statistical agreement with the best chip. A chip that was not piranha etched exhibited substantially higher losses of $9.6 \pm 1.2$ dB/cm.

Fig. 4. (a), (b) The transmitted optical power as a function of waveguide length is shown for two different sets of measurements of the best chip, with losses of $4.3 \pm 0.6$ dB/cm and $4.9 \pm 0.6$ dB/cm achieved. (c), (d) The
transmitted optical power as a function of waveguide length is shown for two other chips showing waveguide losses of $4.7 \pm 0.6$ dB/cm and $9.6 \pm 1.2$ dB/cm, respectively; in the latter case, the chip was not subjected to the piranha etch. For all data series, the fluctuations due to changes in the OPG/DFG power over a period of around 40 seconds are shown in the error bars. A best fit line is also shown, from which the waveguide loss is derived.

As is not surprising, the uncertainty for each waveguide loss measurement exceeded what would be expected solely on the basis of individual device measurement uncertainties. This is due to the fact that there are additional uncertainties beyond the fluctuation on OPG/DFG power; these would include the coupling loss for each endfacet, the random fabrication defects on each device, and a number of other factors. Crucially, however, the final uncertainties due to the linear regression, estimated based on the RMS error between the data points and the fit, are easily small enough for the data to be considered statistically meaningful. It is worth noting here that the linear dependence of the insertion loss on device length suggests that there cannot be a large level of nonlinear loss in the waveguides, as might be caused by a multi-photon absorption mechanism.

To determine the noise floor, the fiber was moved laterally out of alignment with a waveguide, until the power observed on the detector was flat as the fiber moved; the average value for the power reading was then noted. This value turned out to be around 20 dB lower than the typical peak transmission observed for the shortest waveguide bend on the best-performing chips. This should be very adequate for a conclusive measurement of the low level of loss that we obtained. This dynamic range is also evident in Fig. 3. The pulse energy at the detector for the peak power level was approximately 100 pJ after the device insertion losses. In the case of the chip with higher losses, shown in Fig. 4 (d), the shortest device exhibits optical power only 14 dB from the noise floor. As a result, the longest waveguide bend measurement is close to the noise floor, and has a larger uncertainty. But for the other chips shown, more than 10 dB remain between the response of even the longest device and the noise floor. Based on this, the total insertion loss from the fiber to the detector for the shortest devices is 50 dB. The theoretical loss that would be present for coupling from the fiber to a waveguide, and back to a fiber is around 36 dB. The likely discrepancy is due to higher losses coupling to a detector than a fiber, and additional losses from imperfections on the device endfacets. Small alignment errors also likely play a role.

Lower waveguide losses should be possible with improved processing. Based on the material losses of silicon and sapphire, an ultimate loss of around 0.1 dB/cm should be possible at 4.5 µm. We suspect the presently observed losses are due to surface roughness, though it is also possible that surface states could participate in the loss mechanisms [30].

4. Conclusions

With low-loss waveguides, it will be possible to build integrated mid-infrared lasers and detectors using techniques such as wafer bonding [31] and selective-area growth, and to construct a wide variety of further devices within the silicon platform. It should also prove possible to build high-confinement integrated nonlinear optical devices, such as
integrated OPOs and difference frequency generators, because these wavelengths are so long that more than several photons are required to reach the silicon bandgap energy. Another interesting opportunity emerges from the limits of lithography: A 20 nm trench [32] represents a small fraction of a wavelength at 1.55 µm, but represents a significantly smaller fraction for light at 4.5 µm; mid-infrared waveguides may be an ideal ‘playground’ for exploring ideas in ultra-subwavelength photonics. In the long run, we anticipate integrated mid-infrared optical systems, which will be much smaller and cheaper than contemporary systems.

References and links