Large-Scale Integration of Graphene Optoelectronic Devices in Photonic Integrated Circuits

by

Jordan Goldstein

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

© Massachusetts Institute of Technology 2016. All rights reserved.
Large-Scale Integration of Graphene Optoelectronic Devices in Photonic Integrated Circuits

by

Jordan Goldstein

Submitted to the Department of Electrical Engineering and Computer Science on May 11, 2016, in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science

Abstract

Graphene is a 2D material recognized for its extremely high mobility and novel optoelectronic properties. In this thesis, we argue in favor of integrating graphene as an active optoelectronic material alongside optical waveguides on the back-end of CMOS ICs for photonic links in access network and computing applications. We describe a simple fabrication process which can accommodate both graphene modulators and photodetectors on almost any waveguide platform. We use this process to fabricate such devices on silicon waveguides and provide preliminary measurements. Finally, we discuss further research opportunities to improve graphene modulators and detectors to the point of being a competitive technology.

Thesis Supervisor: Dirk R. Englund
Title: Assistant Professor
Acknowledgments

First and foremost I would like to thank Prof. Dirk Englund for his unending dedication to and interest in the progression of my research. Despite his busy schedule he always somehow manages to have time to sit down and discuss roadblocks or other questions that I may have, as well as to brainstorm future directions for research. In the former capacity he has been a source of motivation; in the latter, inspiration. I would also like to thank Ren-Jye Shiue and Dmitri Efetov for their willingness to pass on the baton of knowledge and skill in working with graphene. I am thankful for our conversations on diverse topics such as device characterization, the electronic behavior of graphene, clean graphene transfer, assorted processing steps, etc. These individuals have been instrumental in accelerating my research. I would like to thank Cheng Peng for innumerable conversations on our research and the tools and methods we are both using, for helping me with characterization, and for unfailing emotional support.

I would like to thank Marek Hempel, Wenjing Fang, Yuxuan Lin, and Lili Yu for helpful conversations on working with graphene. I would also like to thank Mark Mondol, Jim Daley, and the staff of MTL for maintaining the fabrication equipment which I use on a day-to-day basis. I would like to thank Janice Balzer for being an extremely efficient administrative assistant and helping me and my colleagues purchase important items from various sources.

I would like to thank Qin Zhang for her undying emotional support as well as for being a positive influence in my life and for my productivity. I would also like to thank my coworkers and friends in the Englund group and beyond for their companionship in graduate school. Finally, I would like to thank my family for raising and supporting me up to this moment, for being a constant supply of affection and for always being willing to talk on the phone when I am in need of conversation.
2.3 Device descriptions .............................................................. 38
2.4 System description ............................................................. 40

3 Preliminary results ................................................................. 41
  3.1 Electrolyte-gated proof of concept (earlier device design) ............ 41
  3.2 ALD-gated detection (present device generation) ....................... 42

4 Future Prospects ................................................................. 45
  4.1 Future efforts within this project ............................................ 45
     4.1.1 Contact resistance ..................................................... 45
     4.1.2 Graphene quality .................................................... 46
     4.1.3 Lithography .......................................................... 47
     4.1.4 Upper graphene layer adhesion ..................................... 47
  4.2 Future projects and technologies ......................................... 48
     4.2.1 Boron nitride encapsulation ........................................ 48
     4.2.2 Chalcogenide glass waveguides ..................................... 49
     4.2.3 Optical resonators .................................................. 49
  4.3 Conclusion ................................................................. 50

A Optimal graphene detector length ........................................... 53
List of Figures

1-1 Hierarchy of an optical network. Higher tiers here correspond to larger length scales, which are accessed by geographically smaller sub-networks of lower tier. From [45].

1-2 A wider hierarchy of data networks, ranging from telecom to on-chip networks. Optical technology continues to gradually work its way down to the smallest length scales as technology improves. From [23].

1-3 Cost of a network data link over a wide range of length scales and for different bandwidths, taking into account the cost of equipment, design, power, etc. From [5].

1-4 On the left is the honeycomb lattice of graphene. The carbon atoms occupying crystallographically nonequivalent lattice sites are colored different (blue and yellow). The real lattice vectors are labeled $\mathbf{a}_i$ and the nearest neighbor vectors $\delta_i$. On the right is its reciprocal lattice and first Brillouin zone, with the reciprocal lattice vectors $\mathbf{b}_i$ and high-symmetry points labeled. From [7].
1-5  a. This figure depicts the tight-binding bandstructure of graphene. One of the two Dirac cones is highlighted. From [7].  b. Interband optical absorption in graphene requires the photon energy to be more than twice the absolute value of the Fermi level. From [19].  c. An example of a basic graphene optical modulator. Between the graphene and the doped silicon waveguide is a layer of ALD dielectric. The conductive waveguide is used to electrostatically modulate the graphene doping, changing its optical absorption. From [30].  d. An example of a waveguide-integrated graphene PTE photodetector. The asymmetric gold contacts give rise to an symmetric Seebeck coefficient profile, leading to a nonzero PTE thermovoltage. From [15].

2-1  A simplified depiction of the device fabrication process. Here a modulator is shown.

2-2  A detailed depiction of the graphene transfer process. The cleaning steps involves an hour-long etch in Transene CE-100 FeCl₃-based copper etchant, two water rinses, a 15-minute clean in 20:1:1 DI water:HCl:H₂O₂, and two more water rinses.

2-3  Cross-sections of the three types of devices used on the chip. a. Modulator/detector utilizing an asymmetric top gate. b. Modulator/detector utilizing an asymmetric top gate as well as a gold contact close to the optical mode for Fermi level bending. c. Simple modulator. Color conventions are same as in Fig. 2-1.

2-4  Top down views of devices A (left) and C (right). Device B is omitted because its top-down appearance is no different from that of A at this zoom level. Legend: Gray, waveguide; red, lower gold; light brown, lower graphene; blue, upper gold; dark green, upper graphene.
3-1  Left: Orange: Photoresponse of first device prior to electrolyte application as a function of source-drain voltage. Blue: Photoresponse of second device after application of electrolyte on the first device. Upper right: Optical image of devices. The middle row of devices was characterized, with light coupled in from the left; used devices are highlighted. Lower right: Optical image showing contact/waveguide overlap due to poor lithographic alignment. The device measurements and plot here were performed by Dmitri Efetov.

3-2  Left: Photocurrent response of device of type a. with a source-drain bias of $V_{SD} = 20 \text{ mV}$. Right: Image of device; damage is from removed wirebonds.
List of Tables

1.1 Comparison of optical modulator technologies. “Electronics-compatible?” refers to compatibility with CMOS electronics either on the FEOL or BEOL of a CMOS process. Data from [41], [8], [29], [2], [11], [30], [31], [36], and [51]. *Negligible for telecom applications. . . . . . . . . . . 27

1.2 Comparison of photodetector technologies. Data from [33], [9], [26], [43], [44], [15], [39], [57]. *Dark current is proportional to bias; these numbers represent the range for typical device operation. . . . . . . . . . . 30
Chapter 1

Introduction

Optical networks are a rather old technology (since the 1980s) which has gradually trickled down to smaller and smaller length scales, moving closer to consumers, especially with the increasing availability of fiber internet [45]. As link bandwidth requirements increase for networks of all sizes—from worldwide to intra-IC—optical technology will become necessary for smaller and smaller network scales, and will become more advanced to support the higher bandwidth requirements. As such, there is a need for compact, efficient and highly-integrated optoelectronic technologies so that photonic links can be inexpensive enough for widespread use and can meet power consumption and energy dissipation requirements. In this thesis we will argue in favor of graphene as an active optoelectronic material, presenting preliminary device details and measurements to motivate further research and optimization of graphene processing and device technology. Chapter 1 provides a general overview of optical networks for various length scales and applications, describes the relevant properties of graphene, compares integrated photonics technologies and argues in favor of graphene/CMOS photonic/electronic integration for access network, data center, inter-IC and intra-IC applications. Chapter 2 discusses the fabrication process and devices used for our graphene-enabled optical network experiments. Chapter 3 provides some preliminary measurements of these devices. Finally, chapter 4 discusses future work within and outside the scope of this project and concludes.
1.1 Photonics in Telecommunications

Optical telecommunications (telecom) is by no means a novel technology. In the 1980s, telecom companies began to retire electronic cable links in favor of fiber optics, starting with inter-city links and gradually introducing the technology to shorter length scales. The motivation for adopting fiber optic technology is two-fold: The signals suffer from less degradation and loss, and the bandwidth can be much higher as the carrier is an optical signal with a frequency of hundreds of THz [45]. Optical telecom signals are typically modulated onto carriers in the vicinity of 1550nm for two reasons: First, this coincides with minimum combined absorption and scattering loss wavelength in silica glass [35]. Second, this coincides with the spectral gain region of erbium-doped fiber amplifiers (EDFAs), which are critical to compensate for fiber attenuation for propagation over distances of hundreds of km. Although early fiber telecom systems used only a single carrier wavelength per fiber, it is now standard to transmit many (perhaps even over 100) carrier wavelengths over a single fiber using wavelength-division multiplexing (WDM). Compared to a one-signal-per-fiber system, this has the advantage that all of the signals can be amplified simultaneously with a single EDFA. Since many EDFAs need to placed along a long optical link, WDM is less expensive per transmitted bit [45].

Optical networks are organized into a hierarchy of length scales. This is illustrated in Fig. 1-1. The backbone tier covers nationwide scales. The regional tier covers multiple metropolitan areas, and the metro-core tier covers a single metropolitan area. The access tier connects the metro-core tier directly to individual customers. Higher tiers serve more customers and thus must accommodate higher bandwidths. However, they also have fewer links per area, so more money can be invested in each link. This cost is consistent with the higher bandwidth requirement. The hierarchical architecture allows telecom companies to deploy higher-bandwidth technology in phases, beginning with the backbone network and working downwards as the technology matures [45].

The equipment at network nodes supporting WDM must perform multiple func-
Figure 1-1: Hierarchy of an optical network. Higher tiers here correspond to larger length scales, which are accessed by geographically smaller sub-networks of lower tier. From [45].

...tions; for instance, it should be able to reconfigurably pass on individual WDM wavelengths to connected nodes in the same tier, or “drop” them from the WDM to pass them to higher or lower tiers. Additionally, it should be able to restore distorted or attenuated optical signals, or convert WDM wavelengths, by detection and regeneration. Currently, these functions are performed by large, expensive, power-hungry routers [45]. For this reason, WDM access networks (for instance, WDM-to-the-home) are still relegated to the distant future. Even as of 2010, only 13% of North American telecom customers are using fiber access networks, and those have only one wavelength for downstream communication and one for upstream [20]. Hence, there is clearly a need for technology which can reduce the cost and size of WDM equipment, as this would give consumers access to higher-bandwidth telecom services.

There is a growing trend to replace bulk optical components with photonic integrated circuits (PICs) in WDM routers, in which many (if not all) of the optical
components needed for WDM–wavelength multiplexers and demultiplexers, waveguides, lasers, modulators, detectors, perhaps even electronic amplifiers and control circuitry or micro-electro-mechanical (MEMS) optical switches–are integrated onto a single small die or package. This monolithic integration promises to lower the cost and improve the reliability of network components, allowing the technology to enter the realm of access networks and provide faster telecom services to businesses and consumers [45].

1.2 Photonics in Computation

Telecom is not the only industry that stands to benefit from photonic integration—data centers also have a strong economic incentive to adopt such technology. Similar to telecom networks, data center networks feature a hierarchy of scales, and so-called optical interconnect technology has been trickling down from the largest scales gradually as technology progresses. These scales are illustrated in Fig. 1-2 [23].

<table>
<thead>
<tr>
<th>Network type</th>
<th>MAN &amp; WAN</th>
<th>LAN</th>
<th>System</th>
<th>Board</th>
<th>Chip</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Metro &amp; long haul</td>
<td>Campus, Enterprises</td>
<td>Intra-rack inter-rack</td>
<td>Chip-to-chip</td>
<td>On-chip</td>
</tr>
<tr>
<td>Distance</td>
<td>Multi-km</td>
<td>10 – 300 m</td>
<td>0.3 – 10 m</td>
<td>0.01 – 0.3 m</td>
<td>&lt;2 cm</td>
</tr>
<tr>
<td>Adoption of optical</td>
<td>Since 80s</td>
<td>Since 90s</td>
<td>Since late 00s</td>
<td>After 2012</td>
<td>After 2012</td>
</tr>
<tr>
<td>Type of Connectivity</td>
<td>All-optical</td>
<td>Point-to-point and All-optical</td>
<td>Point-to-point</td>
<td>Point-to-point &amp; all-optical</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1-2: A wider hierarchy of data networks, ranging from telecom to on-chip networks. Optical technology continues to gradually work its way down to the smallest length scales as technology improves. From [23].

the system level and below, optical networking is not an enabling technology to the
Figure 1-3: Cost of a network data link over a wide range of length scales and for different bandwidths, taking into account the cost of equipment, design, power, etc. From [5].

same extent that it is for telecom. In the realm of telecom, optical technology is required to make high-bandwidth long-haul (100-1000’s of kms) networks feasible with respect to equipment costs. On the other hand, its main benefit for data centers is its energy efficiency (and the associated energy cost) compared to copper interconnects. The cost of copper and optical links scale differently with interconnect length and frequency. Glass fiber loses close to 0.2 dB/km at any bandwidth [45]. On the other hand, copper loss increases with frequency due to the skin effect; Cat-7 cable used for 10 gigabit Ethernet is only rated for frequencies up to 600 MHz, which experience 55 dB of attenuation over only 100 meters [1]. Due to the power cost of transmitting and amplifying signals over such lossy links, for a given frequency, there is a crossover link length above which fiber is cheaper despite the higher cost of optical components. Furthermore, due to the frequency-dependence of the skin effect, this crossover length shrinks with increasing bandwidth [5]. This is depicted in Fig. 1-3.

Consequently, there is an energy-cost incentive to bring optical interconnects to smaller and smaller length scales, especially given that data center bandwidth requirements have been increasing rapidly over the last several years and will continue
to do so [23]. For some high-bandwidth applications, it may even become necessary to bring optical links directly to the integrated circuits themselves, perhaps even using PIC technology alongside electronics for intra-IC interconnects. This necessity is due to limits to IC power dissipation. D. A. B. Miller found in 2009 that in order to keep up with CPU bandwidth projections in the International Technology Roadmap for Semiconductors (ITRS) while meeting power dissipation limits, chip-to-chip links will have to use less than 50 fJ/bit, 1-2 orders of magnitude behind state-of-the-art electronic links but possible with sufficiently advanced yet feasible optical interconnect technology. His findings for intra-chip links were similar [34]. Even if CPU bandwidths do not grow as projected in the ITRS, on- and off-chip optical interconnect technology could still yield 75% energy savings in the American information and communications technology sector according to a U.S. Dept. of Energy-sponsored report [52].

1.3 Basic Properties of Graphene

To compare graphene-based optoelectronics with existing technologies, it is necessary to understand graphene’s fundamental optoelectronic properties. Graphene is a two-dimensional allotrope of carbon in which the carbon atoms form a honeycomb lattice, shown in Fig. 1-4. Tight-binding calculations predict that graphene has the bandstructure shown in Fig. 1-5a. The bandstructure features Dirac cones at the K and K' high-symmetry points where the conduction and valence bands form cone shapes that intersect at a single point. In intrinsic graphene, the Fermi level coincides with the intersection of the two Dirac cones at the Dirac point. The phrases Dirac cone and Dirac point are used because an electron near one of the K points obeys the 2D Dirac equation for a massless fermion, with graphene’s two lattice sites taking the role of the spin degree of freedom in the Dirac equation. The slope of the Dirac cones is given by $h v_F$, where $v_F \approx 1 \times 10^6$ m/s is the Fermi velocity. In graphene, the Fermi level with respect to the Dirac point, $E_F$, is related to the electron or hole density $n$ by $E_F = \pm h v_F \sqrt{\pi n}$; positive for electron-doping and negative for hole-doping [7].
As with any crystal, graphene’s bandstructure is closely tied to its optical properties. Since photons carry very little momentum compared to electrons, electrons optically excited from the valence to conduction band maintain roughly the same wavevector. Due to Pauli blocking, optical excitation requires an electron in the valence band to excite and a vacancy in the conduction band which is exactly $\hbar \omega$ up in energy and at roughly the same wavevector. Hence, interband optical absorption in graphene occurs when the photon energy is greater than twice the absolute value of the Fermi level. This is depicted in Fig. 1-5b. Given this condition is met, graphene’s optical absorption is constant down to visible wavelengths; for light normally incident on suspended graphene, 2.3% is absorbed. Equivalently, graphene has a constant optical conductivity of $\sigma = \pi e^2/2h$ [19].

This Fermi-level dependent optical absorption of graphene can be used to build optical modulators. Graphene has a very low density of states near the Dirac point as a result of its 2D nature as well as it small Fermi surface. As a result, its Fermi level can be tuned electrostatically by fabricating a field-effect transistor (FET). This is similar to a MOSFET, except with a sheet of graphene instead of an induced electron
channel. By applying a voltage to a capacitor consisting of a sheet of graphene separated from another conductor by some dielectric, the graphene’s Fermi level can be shifted substantially. This process is often referred to as electrostatic doping or just doping, not to be confused with the impurity doping used in silicon. Fermi levels as high as 1 eV have been achieved by doping graphene through a solid electrolyte; the electrolyte here gives rise to an extremely high capacitance through the formation of an electric double layer on the surface of the graphene [19]. Alternatively, atomic layer deposition (ALD) can be used to fabricate a thin gate dielectric; such techniques can be used to dope graphene up to and above a Fermi level of 0.4 eV, which is roughly the threshold of Pauli blocking for telecom wavelengths. Hence, by modulating the Fermi level of graphene in the path of light, one can modulate an optical signal. This effect can be enhanced by laying the graphene on top of a waveguide [56] or
a ring resonator to increase the light-graphene interaction time [38]. Furthermore, graphene’s potentially extremely high mobility (roughly 80000 cm$^2$/Vs for state-of-the-art chemical vapor deposition-grown (CVD) samples encapsulated in hexagonal boron nitride (h-BN), measured at room temperature [4]) allows Fermi level modulation at frequencies consistent with modern telecom bandwidths [38].

Graphene has an additional set of unique electronic properties which allow it to behave as a photodetector. When graphene absorbs photons, the photo-excited carriers thermalize to a Fermi-Dirac distribution of elevated temperature within tens of femtoseconds [16]. The carrier energy distribution then cools somewhat due to to intra-band optical phonon scattering, but once the photo-excited carriers are below the optical phonon energy the cooling process (due to acoustic phonon scattering) slows to few-nanosecond time scales. This relatively long time scale for hot electron cooling, combined with a graphene’s low density states, means that an optically excited sheet of graphene will sustain an elevated electronic temperature profile which can be significant higher than room temperature. For instance, a sheet of typical CVD-grown graphene laying on a single-mode silicon waveguide excited with 1 mW of telecom-wavelength light can sustain an electronic temperature on the order of 100 K above room temperature [48].

To convert this raised electronic temperature into an electric signal, the so-called photothermoelectric (PTE) effect is used. Since a material’s Seebeck coefficient is strongly dependent on its Fermi level, the Seebeck coefficient of graphene can be tuned by electrostatic doping. Hence, a graphene P-N junction, likely made by electrostatic doping, can also act as a thermocouple capable of reading out an elevated electronic temperature at the junction due to the different Seebeck coefficients for P- and N-type graphene. In this way, optical excitation of graphene P-N junction can be read out electronically. There is also a photocurrent contribution due to charge separation at the junction (the photovoltaic effect) but this is typically much smaller than the photothermoelectric contribution [48]. Finally, the photoconductive and bolometric effects give rise to a light intensity-dependent resistance, which can be read out under bias. However, biasing graphene detectors is undesirable in practice due to the low
signal visibility within the high dark current [13].

1.4 Comparison of Waveguide-Integrated Optical Modulator Technologies

Here we will compare various common waveguide-integrated optical modulator technologies, including graphene, using common figures of merit. In particular, we focus on technologies which show promise for chip-integration alongside electronics for short-link applications (such as access networks or within data centers), which we will argue is the main market for graphene PIC technology. The typical figures of merit of each technology are shown in Table 1.1.

1.4.1 Silicon carrier depletion modulators

Carrier depletion modulators feature a waveguide coincident with the depletion region(s) of one or more pn junctions. The pn junction can run lengthwise with respect to the waveguide, or the waveguide can pass through multiple interdigitated pn junctions. The refractive index of silicon changes when it is depleted of carriers due to the free carrier plasma effect. Hence, by modulating the reverse bias voltage of the pn junctions, the size of the depletion region(s) changes and the effective index of the waveguide changes. This yields phase modulation which can be converted to intensity modulation using a Mach-Zehnder interferometer (MZI). Due to the relatively small refractive index phase, these devices need to be quite long (millimeter-scale), so a traveling-wave architecture is used to achieve a high speed. This involves copropagating the RF modulation signal alongside the optical signal and requires that the waves have matched group velocities. Alternatively, the change in refractive index can be read out as a shift in the resonant frequency of a ring or photonic crystal resonator. Carrier depletion modulators are appealing because they can be integrated directly into a CMOS process with no additional materials. This makes them the predominant choice for active silicon photonics. Some typical figures of merit are
shown in Table 1.1. Efficiency is defined here as the voltage required for a $\pi$ phase shift over 1 mm, so lower numbers are better [41].

### 1.4.2 GeSi electroabsorption modulators

GeSi electroabsorption modulators rely on either the Franz-Keldysh effect or the quantum confined Stark effect (QCSE) to induce a change in optical absorption with an applied electric field. The Franz-Keldysh effect involves a electric field-dependent absorption edge in bulk materials, whereas the QCSE takes place in quantum wells. Compared to carrier depletion modulators, electroabsorption modulators have the advantage of being more efficient due to the fundamentally stronger modulation mechanism. However, GeSi must be grown epitaxially on silicon, and the requirement of quantum wells for QCSE modulators further complicates fabrication. Furthermore, Ge and GeSi are indirect bandgap materials, and parasitic indirect absorption can lead to high insertion loss in these devices. Nevertheless, this may be a promising modulator technology on silicon photonic platforms where compactness is a major concern. Typical figures of merit are shown in Table 1.1. Here we define the efficiency to be the operating peak-to-peak voltage swing multiplied by the device length required for a 3 dB of extinction [8].

### 1.4.3 Polymer electro-optic (EO) modulators

Certain organic molecules exhibit a strongly quadratic electric susceptibility which can be utilized to construct exceptionally fast electro-optic modulators based on the Pockels effect. These molecules, called chromophores, can be integrated into polymer “hosts” to form nonlinear optical materials. For the second order nonlinearity of the material to be high, the chromophores need to aligned in orientation so that their nonlinear polarizations add. This requires the polymer system to be annealed under a high electric field, which serves to align the chromophores [12]. Like carrier depletion modulators, polymer EO modulators are designed with MZI-traveling-wave or ring resonator architectures. One challenging aspect of these devices is reliability;
high optical intensities can destroy or reorient the chromophores. High temperatures can also cause the chromophores to reorient. Nevertheless, EO polymers have been demonstrated which can tolerate temperatures of up to 85° over decades with no degradation in performance. Another issue with EO polymer modulators is high loss due to optical absorption by the chromophores. EO polymer modulators have been demonstrated on a variety of waveguide platforms, including polymer waveguides, conductive sol-gel waveguides (in order to concentrate the applied electric field in the EO polymer core, with little voltage drop across the cladding), and silicon slot waveguides (the silicon waveguide/electrode-pair is used to apply an electric field across the EO polymer in the slot). However, integration with electronics would restrict these modulators to the back-end of a CMOS process (BEOL), as the materials are not CMOS compatible. In this case, such modulators could be used alongside a BEOL waveguide technology such as Si\(\text{N}_x\), SU8 or chalcogenide glass. Typical figures of merit are shown in Table 1.1. Efficiency here refers to the voltage for a \(\pi\) phase shift [29].

1.4.4 Graphene modulators

Graphene modulators rely on Pauli blocking to achieve modulation as discussed above. An illustrative example of a simple graphene-based modulator due to M. Liu et. al. is shown in Fig. 1-5c [30]. Here, the doped silicon waveguide allows electrostatic doping of the graphene through a layer of ALD. This changes the graphene's absorption, allowing optical modulation. Any transparent, conductive material can be used to dope the graphene. For instance, two sheets of graphene can be used to dope each other [51]. One of the main difficulties with graphene modulators is achieving high speeds. To modulate the graphene’s doping at reasonably low voltages, very high capacitances (on the order of 10 fF/µm\(^2\)) between the graphene and its gate are required. This decreases the RC-limited bandwidth such that straight-waveguide-based graphene modulators have topped out at a bandwidth of about 1 GHz [30][31][36]. Traveling-wave graphene modulators have not been demonstrated. Due to the capacitance issue, there has been a recent trend to further shrink graphene modulators;
### Table 1.1

Comparison of optical modulator technologies. “Electronics-compatible?” refers to compatibility with CMOS electronics either on the FEOL or BEOL of a CMOS process. Data from [41], [8], [29], [2], [11], [30], [31], [36], and [51]. *Negligible for telecom applications.

<table>
<thead>
<tr>
<th>Bandwidth (Gbps or GHz)</th>
<th>Si carrier depletion</th>
<th>GeSi Electroabs.</th>
<th>Polymer Electro-optic</th>
<th>Graphene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 – 50</td>
<td>10 – 40</td>
<td>Up to 150</td>
<td>≈ 1 (waveguide)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Up to 30 (ring)</td>
</tr>
<tr>
<td>Efficiency (V/mm)</td>
<td>10 – 30</td>
<td>0.01 – 0.08</td>
<td>Sol-gel: 20 – 300</td>
<td>0.01 – 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Si Slot: 0.3 – 11</td>
<td></td>
</tr>
<tr>
<td>Insertion loss (dB)</td>
<td>1 – 3/mm</td>
<td>5 – 15</td>
<td>5 – 20</td>
<td>0* – 4</td>
</tr>
<tr>
<td>Capacitance (pF/mm)</td>
<td>0.3 – 1</td>
<td>0.3 – 0.8</td>
<td>≈ 0.1</td>
<td>2 – 8</td>
</tr>
<tr>
<td>Compatible waveguide technologies</td>
<td>Si</td>
<td>Si</td>
<td>Passive waveguide platforms, BEOL</td>
<td>Passive waveguide platforms, BEOL</td>
</tr>
<tr>
<td>Electronics-compatible?</td>
<td>Yes</td>
<td>Probably</td>
<td>Maybe w/ BOEL waveguides</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1.1: Comparison of optical modulator technologies. “Electronics-compatible?” refers to compatibility with CMOS electronics either on the FEOL or BEOL of a CMOS process. Data from [41], [8], [29], [2], [11], [30], [31], [36], and [51]. *Negligible for telecom applications.

A ring-resonator-based modulator has been recently demonstrated with a bandwidth of 30 GHz [38], and photonic-crystal-based modulators have also been demonstrated [51]. Like EO polymer modulators, graphene is particularly well-suited for BEOL waveguide technologies, as it does not involve epitaxy. It is typically deposited using a wet transfer technique, which suffers from scalability and reliability issues; however, more sophisticated and reliable techniques are currently in development [24]. Typical figures of merit are shown in Table 1.1. Efficiency here refers to the voltage required for a 3 dB change in transmission.

### 1.5 Comparison of Photodetector Technologies

There are fewer photodetector options for short-link applications than there are modulator options. Here we list some of the main contenders.
1.5.1 Germanium photodiodes

Germanium p-i-n photodiodes on silicon are the prevailing photodetector technology on silicon photonics platforms; the technology is already employed in commercial silicon photonics fabs [21]. Careful optimization of the growth process has enabled germanium epitaxy on silicon despite the lattice mismatch. Due to the single-crystal silicon requirement for germanium epitaxy as well as the fact that germanium (or SiGe) is already widely utilized in CMOS processes, it is natural to integrate germanium waveguides alongside transistors on the device layer of a CMOS process, although as of 2010 no CMOS foundries offer this capability. Typical figures of merit are shown in Table 1.2 [33].

1.5.2 Graphene photothermoelectric detectors

There are several photodetection mechanisms utilizing graphene, such as PTE (as discussed previously), the photovoltaic effect in graphene-silicon Schottky barriers, the bolometric effect, quantum dot-sensitized photodoping, etc. So far, PTE has proved to give the highest responsivity among technologies which have demonstrated bandwidths high enough for optical communications [26]. A few different device geometries can be used to generate a net PTE thermovoltage. One of the most common is shown in Fig. 1-5d. Here, the asymmetric gold contacts give rise to an asymmetric Seebeck coefficient profile due to the band-bending in the graphene induced by the gold-graphene junction. This asymmetry then allows the net thermovoltage, integrated across the device, to be nonzero; such devices have achieved responsivities of up to 78 mA/W at zero bias [44]. Much higher responsivities can theoretically be obtained by using electrostatic gates to dope a Fermi level “step function” in the graphene channel, but waveguide-integrated photodetectors based on this geometry have not been reported. One can also apply a bias voltage to increase the responsivity of a graphene photodetector; this makes use of the bolometric effect. However, this is not desirable for telecom or interconnect applications; graphene devices resistances on the order of 100 Ω and biasing can give rise to milliamps of dark current (and
the associated increase in noise) for only a modest increase in dark current, up to 0.36 A/W. All in all, further technological improvements (including those associated with graphene transfer as discussed in section 1.4.4) are required to make graphene photodetectors a competitive BEOL technology. Typical figures of merit are shown in Table 1.2 [33].

1.5.3 Black phosphorus photodetectors

One especially novel technology which is potentially promising for BEOL photodetection is black phosphorus (BP). BP has a thickness-dependent direct bandgap ranging from \( \approx 1.5 \text{ eV} \) for a monolayer to \( \approx 0.6 \text{ eV} \) in the bulk, and a hole mobility in the 1000's of \( \text{cm}^2/\text{Vs} \) [40]. BP's semiconducting nature means that it can be used to construct a photoconductor with low dark current. So far, only one demonstration of a waveguide-integrated BP photoconductor has been published; the authors achieve responsivities up to 0.63 A/W with a 3 GHz -3dB frequency for a 100-nm-thick BP absorbing layer, indicating they BP is a promising avenue for future research [57]. The BP in this particular study was mechanically exfoliated, as CVD BP is not yet available. Figures of merit from this paper for two devices of different thicknesses are shown in Table 1.2.

1.6 Discussion of technologies and applications

No photonic link technology is best for all applications, as different applications have different cost and technology constraints. In the realm of electronics and nanofabricated devices in general (especially those requiring sophisticated packaging, as is the case for PICs), a considerable portion of the module cost is due to packaging and assembly. Hence, it is desirable to integrate as much functionality as possible into a single package [47]. For instance, consider the potential future application of a high-bandwidth fiber-to-the-home home 4-channel coarse WDM transceiver. This would be a consumer product and would hence have to be as cheap as possible. One could certainly design such a device using separate chips for the optics and RF electron-
Table 1.2: Comparison of photodetector technologies. Data from [33], [9], [26], [43], [44], [15], [39], [57]. *Dark current is proportional to bias; these numbers represent the range for typical device operation.

<table>
<thead>
<tr>
<th></th>
<th>Germanium photodiode</th>
<th>Graphene PTE</th>
<th>Black phosphorus photoconductor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bandwidth</strong> (GHz, 3dB)</td>
<td>Up to 60</td>
<td>20 – 40</td>
<td>3</td>
</tr>
<tr>
<td><strong>Responsivity</strong> (A/W)</td>
<td>≈ 1</td>
<td>Zero bias: up to 0.08 Bias: up to 0.36</td>
<td>11.5 nm thk.: 0.019 100 nm thk.: 0.63</td>
</tr>
<tr>
<td><strong>Dark current</strong> (µA)</td>
<td>0.0002 – 3</td>
<td>Zero bias: 0 Bias: 500 – 10000*</td>
<td>11.5 nm thk.: 0.22 100 nm thk.: 560</td>
</tr>
<tr>
<td><strong>Compatible waveguide</strong></td>
<td>Si</td>
<td>Si (maybe with CMOS), BEOL</td>
<td>BEOL</td>
</tr>
<tr>
<td><strong>Electronics-compatible?</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Although this technology is not currently in production, one could fathom a few different solutions looking at Tables 1.1 and 1.2. The two clearest solutions are 1) CMOS-integrated-silicon photonics and 2) standard CMOS with BEOL waveguides and graphene actsives. Both are perfectly valid solutions. However, the former is likely to be more expensive for a few reasons. First, since the waveguides and transistors share the active silicon layer, the chip may have to be larger. Second of all, having silicon photonics and CMOS on the same device layer of a chip may increase the process complexity and hence prices. (It should be noted that silicon photonic functionality has been demonstrated on a standard CMOS process; however, a non-standard wavelength of 1180 nm had to be used to to enable photodetection with the SiGe already available in the process, and the inability to change the process to suit photonics resulted in performance sacrifices [50].) On the other hand, the graphene approach puts the electronics and photonics
on different layers and does not require any changes to the FEOL of a CMOS process, potentially resulting in a smaller chip and a cheaper process. This is undoubtedly sufficient justification to pursue research in graphene optoelectronics with the goals of improving device performance and demonstrating system-level functionality. (Note that in either solution, laser light for transmission could be provided from tiny, inexpensive off-the-shelf vertical cavity surface-emitting laser (VCSEL) dies which can be monolithically integrated onto the PICs.) A similar argument can be made for similar types of devices, such as transceivers for data centers.

For optical interconnects within ICs, the story is even simpler. In this case, CMOS integration is a requisite, and the comparison is again between graphene and silicon photonics. In this case, however, graphene is even more appealing. In the state-of-the-art CMOS processes used for server processors, highly complex processes (such as to make FinFETs) make chip costs per square millimeter considerably higher than in older process nodes. Thus, wasting precious silicon device layer space on large waveguides and optoelectronics is extremely expensive and unappealing. Furthermore, these processors are already nearing the maximum die size available with current photolithography technology, and adding photonics to the device layer would force semiconductor companies to put fewer transistors on their server CPUs. Hence, a BEOL photonics solution is extremely desirable, for which graphene is again the natural choice.

It should be noted that we have neglected the possibility of pick-and-place die-bonding of InP-based detectors and modulators onto BEOL waveguides or silicon photonics via pick-and-place dies. Although this technology is being developed [42], the cost of picking and placing a large number of custom III-V dies is liable to increase costs considerably. Hence, this technology is likely more suitable for applications such as backbone/regional/metro telecom where cost is less of an issue and the high performance of III-V devices is required.
Chapter 2

Approach

Here, we will first discuss the overarching goal. Then we will discuss the fabrication process, highlighting important design constraints and parameters and how they affect the devices. Finally, we will motivate and describe the actual device designs.

2.1 Goal statement

The main goal of this project is to demonstrate the simplest kind of graphene photonic link—sending and receiving data over a single waveguide. An additional goal may be to perform the same function using two identical devices; that is, devices which can function either as a detector or a modulator. Since we intend to demonstrate the feasibility of graphene optoelectronics for large-scale integration, it is also desirable to achieve high yield, which for us is limited by the quality of the graphene transfer process at our disposal. We chose to use silicon waveguides for this project; in particular, we have chips from IBM with many parallel silicon waveguides spanning the chip crosswise. The 220 nm tall by 500 nm wide waveguides are designed to be single mode at a wavelength of 1550 nm. Besides SiNx edge couplers on either side of each waveguide, the chip is planarized, with SiO2 cladding to the side of and beneath the waveguides. The top side of the waveguides is exposed to allow strong interaction between the graphene and the optical mode. The entire chip is covered with roughly 5 nm of HfO2 atomic layer deposition (ALD) in order to electrically
isolate the graphene from the silicon waveguides. Although it is true that our eventual aspirations for graphene optoelectronics involve BEOL waveguide technologies such as SiN$_x$ instead of FEOL silicon, we chose to use silicon waveguides for this demonstration simply because we already had these chips at our disposal.

2.2 Process description

In order to build modulators it is necessary to have a gate dielectric and a top gate to dope the graphene channel with. Unless the gate dielectric is extremely thick (undesirable, as then the peak-to-peak voltage to drive the modulator would be very high), both will coincide with the optical mode and hence both must be relatively transparent. We choose to use another layer of graphene as the top-gate since there are very few other relatively telecom-transparent conductive materials which can be deposited at low temperatures. For the dielectric we use 65 nm of Al$_2$O$_3$ grown by ALD at 200°C which has already been shown to be capable of doping graphene into Pauli blocking for telecom frequencies [38]. Both the top and bottom graphene layers are contacted by Ti/Au electrodes. The ALD layer is thin enough to be penetrated by an ultrasonic wirebonder when contacting the lower gold pads. The process overall is shown in Fig. 2-1. We will now discuss the process in detail step by step.

2.2.1 Graphene transfer

Graphene transfer is the most technically difficult step in this process. Lack of care can lead to poor transfer quality, with many fissures, holes, or folded-up regions in the graphene film. The process is depicted in Fig. 2-2. We begin with CVD monolayer graphene grown on copper by our collaborator Marek Hempel in Prof. Jing Kong’s lab. The copper foil used for the CVD growth is chosen to be as flat as possible in order to avoid wrinkles and fissures in the final transferred graphene. A graphene transfer frame is assembled using PET film and Kapton polyimide tape as shown in Fig. 2-2. The bottom layer of PET allows PMMA to be spun on without creasing the CVD film, and the second layer of PET gives the film some mechanical integrity even
Figure 2-1: A simplified depiction of the device fabrication process. Here a modulator is shown.

after removing the bottom layer and protects the PMMA during the O₂ reactive ion etch (RIE). The Kapton tape, placed on all four sides of the rectangular CVD film, provides mechanical integrity and prevents PMMA contacting the bottom of the CVD film. Microchem 950K PMMA A6 poly(methyl methacrylate) e-beam resist is spun onto the frame at 3000 RPM for 60s and baked in an oven at 120°C for 30 minutes. The aforementioned process parameters make the PMMA/graphene film thin enough to conform to the substrate, but no so thin that it tears easily. The PET backside is removed and the frame is inverted into a PlasmaQuest RIE etcher and etched for 90 s in O₂:He to remove the unwanted graphene on the rear side of the copper. The PET/Kapton frame is cut off, and the Cu/graphene/PMMA film is cut into pieces of the desired size and etched in Transene CE-100 FeCl₃-based etchant until the copper
is removed (about an hour). Using a glass slide, the graphene is transferred into a dish of DI water, water again, and then a solution of 20:1:1 DI water:HCl:H₂O₂. This solution helps remove Fe⁺³ ions which may remain from the etching step [28]. The film is then transferred to DI water twice again to rinse. The desired substrate is subjected to oxygen plasma for 5 minutes to increase hydrophilicity and used to scoop out the graphene film. The assembly is gently blown dry with nitrogen, pushing water out from the edges of the film, and baked at 80°C for 15 minutes and then 150°C for two hours. The initial bake helps water gently evaporate from the edges of the film without violent boiling, and the second bake causes the PMMA to reflow, increasing conformity. Finally, the PMMA is removed with a 7-minute soak in acetone followed by an isopropanol (IPA) rinse and N₂ blow dry. For the first graphene transfer in the process, the chip is then be annealed at 350°C for 3 hours in a 1:1 Ar:H₂ environment to remove organics adsorbed on top of the graphene as well as to increase adhesion between the graphene and the substrate. For the second graphene layer, this is not
an option as the high annealing temperature will cause adsorbed organic residues beneath the gold and ALD to boil, causing mechanical damage.

2.2.2 Graphene patterning

In order to pattern the graphene, Microchem 950K PMMA A6 positive e-beam resist is spun onto the chip for 60s at 3000 RPM and baked on a hotplate at 180°C for 120s to cure. The devices are written in a Raith 150 with a beam accelerating voltage of 30 kV and a dose of 350µC/cm². Alignment is performed using features already present on the waveguide chip. The resist is then developed in 1:3 methyl isobutyl ketone (MIBK):IPA for 90s and then rinsed in IPA for 120s to flush out any remained developer. The graphene is then etched away through the resist mask in a PlasmaTherm RIE system using O₂:He gas for 90s. Finally, the resist is removed with a 7-minute soak in acetone and the chip is rinsed in IPA and blown dry with N₂.

2.2.3 Gold contact patterning

A re-entrant resist edge profile is required for high-quality liftoff, so we use a bilayer of two different PMMA molecular weights for this part of the process. First Microchem 495K PMMA A4 is spun on at 2000 RPM for 60s, followed by 950K PMMA A2 at 4000 RPM. The lower molecular weight PMMA develops away more strongly than the higher one, yielding an undercut. The Raith 150 is used to expose the chip with an accelerating voltage of 30 kV and a dose of 400µC/cm². Development is identical to the graphene patterning step. Prior to metal evaporation, a protective layer of PMMA is painted over the raised edge couplers and left to dry overnight. 5 nm of Ti followed by 100 nm of Au is then deposited with a Temescal FC2000 electron beam evaporator. The titanium improves adhesion between the gold and the substrate. Finally, liftoff is performed by vertically immersing the chip in boiling acetone. Agitation is typically required to remove the unwanted gold. The chip is then rinsed in IPA and blown dry with N₂.
2.2.4 Atomic layer deposition

Atomic layer deposition on graphene is difficult because the precursors have low adhesion to clean graphene. To get around this, prior to ALD, 1.5nm of Al is evaporated on the sample at a very low rate \(0.1 - 0.2\text{Å/s}\) using the Temescal FC2000 e-beam evaporator. Upon removal from vacuum the aluminum immediately oxidizes to form transparent \(\text{Al}_2\text{O}_3\). This forms a template that the ALD precursors can adhere to. To ensure complete oxidization of the aluminum layer, the sample is baked in air at 150°. 600 layers (roughly 65 nm) of \(\text{Al}_2\text{O}_3\) are deposited at 200°C in a Cambridge Nanotech ALD system using water and trimethylaluminum as precursors. Finally, the graphene and gold deposition and patterning steps described above are repeated to finish the chip.

2.3 Device descriptions

Three different types of devices were fabricated on shared waveguides. They are depicted in Fig. 2-3. Note that in all cases, although there may be overlapping contact pads, the contacts break out in the third dimension into individual pads. For ease of wirebonding, the pads are no smaller than 95 μm × 95 μm and are devoid of graphene in the middle 85 μm × 85 μm. In all cases, the graphene channel is 70 μm long, which is chosen to minimize the noise-equivalent power of the graphene detectors (see Appendix A).

The device in Fig. 2-3a. is a dual modulator/detector. The waveguide is 500 nm wide and the gold contacts are 750 nm from the waveguide. For modulator functionality, the voltage between the graphene channel and the topgate can be modulated to bring the graphene directly above the waveguide in and out of Pauli blocking. For detector functionality, the photocurrent coming from the source and drain contacts is monitored as light is sent through the waveguide. The topgate is used to dope an asymmetric Seebeck coefficient profile into the graphene (note that the topgate cuts off on the left side of the waveguide). For good photodetection bandwidth, it is important to apply a voltage to the topgate through a very high impedance to
eliminate any capacitive coupling between the topgate and the channel. There is a strong tradeoff between modulation and photodetection performance in the device. For photodetection, it is ideal for the topgate to cut off in the center of the waveguide so that the change in Seebeck coefficient occurs at peak of the electronic thermal distribution. However, this would yield a modulator with a high insertion loss, as half of the graphene in contact with the waveguide would be absorbing unconditionally. Device B is identical to device A with the exception of the left gold contact; here, the proximity between the gold contact and the waveguide provides additional shifting of the Fermi level in the graphene channel due to Fermi level pinning. Device C is a modulator, designed to be much faster than devices A and B as a result of its lower capacitance. Fig. 2-4 shows a top-down view of the actual devices, indicating placement of the contacts.
2.4 System description

Several different simple optoelectronic send-receive links were fabricated: Modulator A to detector A, modulator B to detector B, modulator C to detector A, and modulator C to detector B. The length of the waveguide between the two devices varies across the chip from just 10s of µm to 3 mm, with most being roughly 350 µm. The chip features 106 devices in total.
Chapter 3

Preliminary results

3.1 Electrolyte-gated proof of concept (earlier device design)

An earlier device design (based on device b. from Fig. 2-3 but with a different contact geometry from that in Fig. 2-4) was tested prior to the ALD step for send/receive capability. The devices and photocurrent are shown in Fig. 3-1. A laser operating at $\lambda = 1550$ nm and modulated between 0 and 5 mW was coupled into the middle row of devices from the right. The orange curve shows the photovoltage of the first device as a function of its source-drain voltage. In this state, no photoresponse was measurable on the second device due to light absorption in the first. To allow light to reach the second device, PEO:LiClO$_4$ electrolyte was manually deposited onto the first device to chemically dope the graphene channel beyond Pauli blocking. The photovoltage then measured on the second device is shown plotted in blue. This demonstrates that the photoresponse of one graphene device can be modulated affected by another. Comparing the two curves, the first device has an insertion loss of roughly 10 dB. This is probably associated with light absorption in metal due to waveguide-contact overlap resulting from poor lithographic alignment, as discussed further in Chapter 4.
3.2 ALD-gated detection (present device generation)

A laser operating at $\lambda = 1550$ nm and modulated between 0 and 5 mW was coupled into a device of type a. (see Figs. 2-3 and 2-4) to perform a top-gated photodetection measurement. The device’s resistance was measured to be $R_{SD} = 68 \Omega$. A source-drain bias voltage of $V_{SD} = 20$ mV was applied to increase the response of the device. The photocurrent was amplified by an SR570 transimpedance amplifier and measured with a SR830 lock-in amplifier as a function of top-gate voltage. The results and device image are shown in Fig. 3-2. The peaked, asymmetric photocurrent profile is consistent with the findings of Gabor et. al. for modulation of a top-gate covering part of the graphene channel, which they explain based on PTE [14]. We are not sure why the measured photocurrent is so low. One possibility is light absorption in the upper-layer graphene, which may have shifted onto the couplers during liftoff. Regardless, this measurement provides a proof-of-concept demonstration of photocurrent in a waveguide-integrated device with the doping profile provided by a partial top-gate instead of proximity to a gold contact.
Figure 3-1: Left: Orange: Photoresponse of first device prior to electrolyte application as a function of source-drain voltage. Blue: Photoresponse of second device after application of electrolyte on the first device. Upper right: Optical image of devices. The middle row of devices was characterized, with light coupled in from the left; used devices are highlighted. Lower right: Optical image showing contact/waveguide overlap due to poor lithographic alignment. The device measurements and plot here were performed by Dmitri Efetov.

Figure 3-2: Left: Photocurrent response of device of type a, with a source-drain bias of $V_{SD} = 20 \text{mV}$. Right: Image of device; damage is from removed wirebonds.
Chapter 4

Future Prospects

4.1 Future efforts within this project

4.1.1 Contact resistance

One of the major issues affecting graphene devices (and electronic devices in general) is contact resistance, as increased contact resistance lowers the RC cutoff frequency of devices in general and decreases the responsivity of detectors. If care is not taken to engineer and fabricate good contacts, resistances on the order of $10^4 \Omega \mu m$ can be expected, which can easily dominate the total resistance of a device. Hence, lowering the contact resistance to graphene devices has been an important area of research in the field. Perhaps the lowest reported contact resistance to date, $23 \Omega \mu m$ at room temperature with CVD graphene, has been achieved by a combination of polymer doping the graphene to increase its density of states at the Fermi level and patterning the graphene to increase the total edge length (graphene edges contribute significantly to conduction at contacts) [37]. Resistances as low as $100 \Omega \mu m$ to an unpatterned graphene edge have also been achieved with exfoliated boron nitride-encapsulated graphene [53]. Furthermore, a strong correlation has been found between the metal contact deposition pressure and the resulting contact resistance, with the contact resistance to exfoliated graphene varying roughly as the deposition pressure squared with no asymptote down to 40 nTorr [27]. Since the typical resistance per length of
a graphene channel is itself no less than a few hundred Ω μm, contact resistances of 100Ω μm are already low enough to yield close-to-optimal device performance. Within the scope of this project, contact resistances can be minimized by careful optimization of the metal deposition process. Beyond this, new process features would be required, such as boron nitride (CVD boron nitride is just becoming available) or chemical or plasma doping of the graphene at the contacts.

4.1.2 Graphene quality

Impurities, often trapped between graphene and its substrate, are an important problem with graphene devices. Electric fields caused by trapped metal ions or organic molecules can cause random variation in the doping of graphene over sub-micron length scales. This gives rise to electron or hole “puddles”, localized regions of increased electron or hole density which can occur even for undoped graphene [32]. These puddles can also be caused by dangling bonds or charge traps on the graphene’s substrate [10]. J. Martin et. al. have directly measured such variations in the Dirac point of exfoliated graphene using a scanning single-electron transistor and have found standard deviations of 50 meV [32]. Stronger variation can occur for wet-transferred CVD graphene due to PMMA residue or trapped Fe$^{+3}$ ions left over from the copper etching. (This puddling can be easily measured with a gate-dependent resistance measurement; the puddling strength determines the width of the resistance peak, the “Dirac peak” at the Dirac point.) Puddling is problematic for PTE-based detectors since it smooths out the relationship between conductance and Fermi level, which decreases the Seebeck coefficient [48]. It is also problematic for modulators because it increases the gate voltage swing required to bring a sheet of graphene from fully transparent to maximally absorbing for a given wavelength of light. One potential remedy for the puddling problem is to transfer graphene using electrochemical delamination (so called “bubble-transfer”) instead of copper etching. Electrochemical delamination involves delamination of graphene from the CVD growth metal due to bubble formation when a voltage is applied between the metal and an electrolyte bath. A comparative study showed graphene transferred using a FeCl$_3$ etchant has
considerable metal ion contamination, compared to almost none for electrochemically delaminated graphene [54]. However, my coworkers report not seeing a difference in the Dirac peak between FeCl$_3$-etched and electrochemically delaminated graphene. More experimentation will be required (perhaps with different CVD graphene sources) to determine how to shrink the Dirac peak for better device performance.

4.1.3 Lithography

Lithographic misalignment has been a major source of trouble throughout this project. In particular, the machine we have been using for e-beam lithography (Raith 150) has been giving us poor write field stitching, with stitching errors of around 500 – 1000 nm despite our best efforts to the contrary. This could potentially be remedied by moving to the Elionix F125, another e-beam lithography tool available at MIT with better field stitching. However, the Elionix only operates with an e-beam accelerating voltage of 125 kV, so further process development would be necessary to account for that.

4.1.4 Upper graphene layer adhesion

High-temperature annealing of the lower graphene layer at 350$^\circ$C promotes adhesion with the substrate, allowing the graphene to survive mechanical agitation such as that which occurs during metal liftoff. However, it is not possible to anneal the upper graphene layer in this way because it will cause residual organics from the lower graphene transfer and graphene/gold lithography to boil and puncture the ALD. It may also cause the lower graphene-metal contact resistance to increase. As a result, shifting of the upper graphene layer can occur during the second liftoff, decreasing yield and potentially increasing the coupling loss into the chip due to graphene shifting onto the edge couplers. Phare et. al. report remedying this problem with a 10-minute air bake at 170$^\circ$C after removing the graphene transfer PMMA, which we have not tried. An alternative approach may be to investigate cleaner processing techniques, such that the upper graphene layer can be annealed (albeit perhaps at reduced temperatures) due to the absence of organic residues. For instance, we could
attempt to use solvents stronger than acetone for resist stripping and liftoff (such as N-methyl-2-pyrrolidone), or we could use a PMMA/PMGI (polymethylglutarimide) dual layer process in which the polymer layer in contact with the chip can be removed with a solution much stronger than the e-beam developer. Such techniques are worth investigating regardless in order to potentially shrink the width of the Dirac peak for improved device performance.

4.2 Future projects and technologies

4.2.1 Boron nitride encapsulation

Hexagonal boron nitride (h-BN) is a layered dielectric material of the same structure as graphene with boron and nitrogen at the two crystal lattice site. The different atomic species in h-BN give rise to a very large bandgap of $\approx 5.5$ eV. Furthermore, h-BN can be grown by CVD, giving it potential for large-scale integration [49]. Besides its utility for building low-resistance contacts, h-BN has a few other uses with respect to graphene devices. Primarily, graphene encapsulated in h-BN has improved electrical properties with respect to graphene on other substrates, such as silica. Due to h-BN’s atomically smooth surface and lack of dangling bonds and charge traps, it eliminates puddling associated with substrates as well as electron-charge impurity scattering. It also has high phonon energies which reduce the electron-phonon scattering in graphene, increasing PTE performance and lowering resistivity. Exfoliated graphene-on-h-BN samples have shown mobilities an order of magnitude higher, and carrier density inhomogeneity (a measure of puddling) an order of magnitude lower, than similar exfoliated graphene-on-$\text{SiO}_2$ samples [10]. Hence, h-BN encapsulation has the potential to yield considerable improvement in the performance of graphene optoelectronic devices, especially in the RC-limited cutoff frequency. In fact, the highest-responsivity detector listed in Table 1.2 relies on h-BN encapsulated graphene [44]. However, this device utilized all exfoliated graphene and h-BN. Such results must be replicated with CVD graphene and h-BN in order to demonstrate
scalability. This has yet to be reported.

4.2.2 Chalcogenide glass waveguides

Chalcogenide glasses are a promising material family for BEOL waveguides due to their low processing temperature and potentially quite high refractive index, up to 2.7 at $\lambda = 1550$ nm for As$_{40}$Se$_{60}$ [58]. These high refractive indices allow the design of waveguides which can accommodate small turning radii, critical for compact BEOL photonics. Demonstrating high-performance graphene modulators and detectors integrated with chalcogenide glass waveguides would constitute a major step towards BEOL integration of graphene interconnects.

4.2.3 Optical resonators

As mentioned in chapter 1, optical resonators will likely be necessary to implement graphene modulators with a cutoff frequency high enough for data links. As for detectors, while not strictly necessary, resonant structures can help decrease footprint and NEP. Good resonators rely on high index contrast waveguides. For instance, for a given bending radius, loss in curved waveguides decreases exponentially with the index contrast. As the index contrast decreases, the bending radius required for a 90-degree turn of a given loss increases roughly proportionally [46]. Hence, compact, hi-Q ring resonators require high index contrast waveguides. For instance, while silicon microring resonators with radii of 1.5 $\mu$m have been shown [55], SiN$_x$ resonators must have radii around or above 20 $\mu$m [18]. The situation is similar with photonic crystals; a TE-mode photonic bandgap is not even possible in a triangular lattice of holes in a dielectric film if the index contrast is small than about 1.4 [22]. Hence, high index BEOL waveguide materials such as As$_x$Se$_{1-x}$ are ideal to make compact, high-performance resonators for graphene detectors and modulators.

Another important consideration regarding resonators is tuning. Due to variation in fabrication conditions as well as laser source wavelengths, it is important to be able to tune the resonance of a resonator to line up with the source wavelength. For silicon
waveguides this is most commonly achieved with heaters, which take advantage of silicon’s high thermo-optic coefficient of $1.8 \times 10^{-4} \text{K}^{-1}$ \cite{25} to tune the resonance. However, such heaters typically suffer from low efficiency since the resistive metal heaters typically used must be placed far away from the resonator so as to not disturb the optical mode \cite{6}. Graphene offers a fortuitous solution, as the same sheet of graphene used to modulate absorption could also be used as a heater. On the other hand, the BEOL waveguide materials for which graphene technology is most appealing have much lower thermo-optic coefficients than silicon, with $dn/dT_{\text{SiN}} = 2.45 \times 10^{-5} \text{K}^{-1}$ \cite{3} and $dn/dT_{40\text{As}60\text{Se}} = (3.2 \pm 1.1) \times 10^{-5} \text{K}^{-1}$ \cite{17}. (Note that due to the wide nature of the chalcogenide glass material family, room for improvement here is possible.) The close contact between graphene heaters and BEOL waveguides may offer some compensation for these poor material properties. Beyond this, smaller resonators (such as photonic crystal cavities) can be tuned more efficiently simply due to their reduced thermal mass. Either way, tunability and the associated process design considerations are essential for practical, high-speed and high-performance graphene BEOL photonic circuits.

\section*{4.3 Conclusion}

As bandwidth requirements increase for telecom networks as well as for links within data centers and on/between ICs, photonic link technologies become increasingly appealing for smaller and smaller link lengths due to their potentially extremely high bandwidth and very low bandwidth-independent loss. There are a few different approaches and technologies for photonics-enabled network systems, with silicon photonics being one of the most promising and well-developed, especially for lower-bandwidth applications such as access networks or data centers (with respect to high-bandwidth applications, such as long haul/regional/metro telecom, for which III-V-based technology may be more suitable). Furthermore, for packaging cost reasons, it is desirable to put as much functionality as possible onto a single chip. However, combined Si photonics/CMOS may be an unappealing solution due to competition
between photonics and electronics for costly chip area. Thus, a better solution may be to integrate optoelectronics with BEOL waveguide materials such as SiN\textsubscript{x} or chalcogenide glasses. Here graphene is a promising active material due to its high speed, low cost and low processing temperature. We describe a simple process to fabricate graphene optoelectronic devices with two graphene layers and execute the process on a planarized silicon waveguide chip (although the process can be performed on planarized waveguides of any material stable to 350°C), fabricating two different types of combined detector/modulators as well as a simple modulator. The devices exhibit preliminary functionality indicating that further research and development may give rise to devices good enough for the aforementioned applications, and future avenues for investigation are discussed such as contact resistance and graphene quality improvement as well as h-BN encapsulation and integration with BEOL-compatible waveguides and resonators. All in all, although the technology for graphene optoelectronics is currently immature, collective progress in materials processing and device design indicates the graphene has great potential as a standard active material in future photonic microsystems.
Appendix A

Optimal graphene detector length

Consider a graphene photodetector of length $L$ with translational symmetry along the waveguide axis, which we shall label the $x$-axis. Imagine slicing the detector into identical slices of differential thickness $dx$. Each of these slices has an identical intrinsic responsivity $g$, with units of A/W, which measures the photocurrent contribution from that slice divided by the power that it absorbs (in graphene or otherwise). We assume that the distance between the gold contacts is much smaller than the optical absorption length of the loaded waveguide $1/\alpha$, which allows us to approximate the current in the graphene as flowing perpendicular to the waveguide (i.e. parallel current flow occurs only in the highly conductive gold contacts). This is very similar to the long-channel approximation used in MOSFET analysis. The intrinsic responsivity is not a function of slice thickness $dx$ for the following reason. In the PTE-based photodetector, an electromotive force voltage is generated by the Seebeck effect proportional to the electronic temperature increase, which is proportional to the power absorbed per length. The absorbed power per length is not a function of $dx$, so neither is the photovoltage. Since the resistance of the slice scales as $dx^{-1}$, the photocurrent scales as $dx$. However, the total power absorbed in the slice also scales as $dx$, so the intrinsic responsivity has no $dx$ dependence. Each slice can be modeled as an infinitesimal optical excitation-dependent current source in parallel with an infinitesimal resistor. The slices are connected to each other in parallel under the “long-device approximation,” yielding a total device resistance $R = \rho/L$ where $\rho$
is the resistance-length of the slices. The total photocurrent is given by the integral
\[ I = \int_0^L g \left( -\frac{dP}{dx} \right) dx, \]  
(A.1)
where \( P(x) \) is the power in the waveguide, equal to \( P(x) = P_0 e^{-\alpha x} \). Continuing,
\[ I = -gP_0 \int_1^{e^{-\alpha L}} dP = gP_0(1 - e^{-\alpha L}). \]  
(A.2)
Johnson noise is the predominant source of noise in unbiased graphene devices since there is no dark current. The Johnson noise current in the device is given by:
\[ i_{RMS} \equiv \sqrt{\langle i_{\text{noise}}^2 \rangle} = \sqrt{\frac{4k_B T \Delta f}{R}}, \]  
(A.3)
where \( i_{RMS} \) is the root mean square (RMS) noise current, \( i_{\text{noise}} \) is the instantaneous noise current, \( T \) is the temperature, \( k_B \) is Boltzmann’s constant, and \( \Delta f \) is the bandwidth in consideration. Then the noise equivalent power is given by the RMS noise current divided by the responsivity, divided by root bandwidth:
\[ \text{NEP} = \sqrt{\frac{4k_B T \Delta f}{\rho g L}} \frac{1}{g(1 - e^{-\alpha L})} \frac{1}{\Delta f} = \sqrt{\frac{4k_B T}{\rho g^2 \alpha} \frac{\sqrt{\alpha L}}{1 - e^{-\alpha L}}}. \]  
(A.4)
This function is minimized at \( \alpha L \approx 1.26 \), yielding an NEP of
\[ \text{NEP} \approx \frac{3.13}{g} \sqrt{\frac{k_B T}{\rho \alpha}} \approx \frac{2.50}{g} \sqrt{\frac{k_B T}{R}}. \]  
(A.5)
Note that one cannot simply shrink the NEP by adding series resistance to the device, since this will cause \( g \) to decrease faster than \( \sqrt{R} \) increases. In our devices, \( 1/\alpha \approx 50 \mu m \), so \( L_{\text{min}, \text{NEP}} \approx 63 \mu m \). We decided to round up to 70\mu m.
Bibliography


