Demonstration of Monolithically Integrated Graphene Interconnects for Low-Power CMOS Applications

by

Kyeong-Jae Lee

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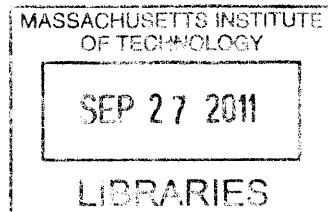in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Author ....
Department of Electrical Engineering and Computer Science
August 19, 2011

Certified by ....
Jing Kong
ITT Career Development Associate Professor
Thesis Supervisor

Certified by ....
Anantha P. Chandrakasan
Joseph F. and Nancy P. Keithley Professor
Thesis Supervisor

Accepted by ....
Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

Demonstration of Monolithically Integrated Graphene Interconnects for Low-Power CMOS Applications

by

Kyeong-Jae Lee

Abstract

In recent years, interconnects have become an increasingly difficult design challenge as their relative performance has not improved at the same pace with transistor scaling. The specifications for complex features, clock frequency, supply current, and number of I/O resources have added even greater demands for interconnect performance. Furthermore, the resistivity of copper begins to degrade at smaller line widths due to increased scattering effects. Graphene has gathered much interest as an interconnect material due to its high mobility, high current carrying capacity, and high thermal conductivity. DC characterization of sub-50 nm graphene interconnects has been reported but very few studies exist on evaluating their performance when integrated with CMOS. Integrating graphene with CMOS is a critical step in establishing a path for graphene electronics.

In this thesis, we characterize the performance of integrated graphene interconnects and demonstrate two prototype CMOS chips. A 0.35 μm CMOS chip implements an array of transmitter/receivers to analyze end-to-end data communication on graphene wires. Graphene sheets are synthesized by chemical vapor deposition, which are then subsequently transferred and patterned into narrow wires up to 1 mm in length. A low-swing signaling technique is applied, which results in a transmitter energy of 0.3-0.7 pJ/bit/mm, and a total energy of 2.4-5.2 pJ/bit/mm. We demonstrate a minimum voltage swing of 100 mV and bit error rates below 2x10^{-10}. Despite the high sheet resistivity of graphene, integrated graphene links run at speeds up to 50 Mbps. Finally, a subthreshold FPGA was implemented in 0.18 μm CMOS. We demonstrate reliable signal routing on 4-layer graphene wires which replaces parts of the interconnect fabric. The FPGA test chip includes a 5x5 logic array and a TDC-based tester to monitor the delay of graphene wires. The graphene wires have 2.8x lower capacitance than the reference metal wires, resulting in up to 2.11x faster speeds and 1.54x lower interconnect energy when driven by a low-swing voltage of 0.4 V. This work presents the first graphene-based system application and demonstrates the potential of using low capacitance graphene wires for ultra-low power electronics.

Thesis Supervisor: Jing Kong
Title: ITT Career Development Associate Professor

Thesis Supervisor: Anantha P. Chandrakasan
Title: Joseph F. and Nancy P. Keithley Professor

Acknowledgments

I will praise the LORD, who counsels me;

even at night my heart instructs me.

I keep my eyes always on the LORD.

With him at my right hand, I will not be shaken.

Psalm 16:7-8

Thank you Jesus, my rock and my fortress.

First, I sincerely thank my two advisors, Professor Jing Kong and Professor Anantha Chandrakasan. Your advice and support has made this work possible. Thank you for giving me this research opportunity and being patient with this project. I thank Jing for her kindness, openness, and continuous encouragement. I have learned how to maintain the right perspective on all matters, which extends far beyond research. I thank Anantha for his dedication, passion, and continuous support. Anantha is truly one of the best at identifying the problem and providing direction and motivation. I have learned a great deal from both of you.

I am also very grateful to my thesis committee member, Professor Tomas Palacios. Tomas has given this work a great deal of attention and support. Thank you for your enthusiasm and interest in this thesis. Your feedback and support has made this thesis better.

This work has also greatly benefited from the research of current and past students. I thank Mario Hofmann for helping create the RF probe mounts for the first test chip. I am also grateful for many of the discussions we have had over the years. I also thank Hyesung Park for assisting part of the graphene transfer work. I am grateful for Masood Qazi for discussing circuit design and testing issues. I also thank my co-workers, especially Allen Hsu, Kikang Kim, Yongcheol Shin, Soomin Kim, Sungmi Jung, Minseok Choi, Wenjing Fang, Roman Caudillo, Helen Zing, Alfonso Reina, Daniel Nezich, Hootan Farhat, Jikang Yuan, Marek Zebrowski, Ervin Miles, Sreekar Bhaviripudi, Stefan Thiele, Ya-Ping Hsieh, Xiaoting Jia, Federico Villalpando Paez, Liming Xie, Paulo Araujo, Daniela Mafra, Xialoing Shi, Ying Feng, Yumeng Shi,

Contents

List of Figures

12

13

14

15

List of Tables

Chapter 1

Introduction

1.1 Interconnect Challenges

The semiconductor industry has advanced at an exponential rate over the last few decades. In recent years, interconnects have become a major limiting factor on the performance of very large-scale integrated (VLSI) systems [1–3]. The relative performance of interconnects has not improved at the same pace with transistor scaling [2]. Latency, energy dissipation, and signal integrity have all become an increasingly difficult problem to cope with. With shrinking cross-sectional areas and hence increased electrical resistance, the interconnect delays have begun to exceed transistor delays and this trend worsens at advanced technology nodes [2, 4]. While the capacitance of a global wire remains fairly constant under technology scaling, the addition of more complex features has resulted in higher energy dissipation. Global wires often dominate the total power consumption in many VLSI systems [5, 6]. Furthermore, the specifications for clock frequency, supply current and voltage, number of I/O resources has added greater demands for higher levels of integration and interconnect performance [7].

Many solutions have emerged to address these challenges, ranging from new materials and processes to novel micro-architectures. At the system level, multi-core systems has emerged as a recent trend, where slower data transfers are managed across multiple dies and faster local communication is kept on-chip [8, 9]. Nonethe-

less, for high bandwidth systems, cross-chip communication can still limit total performance as it increases on-die cache delays and buffer resources. Three dimensional integration can also benefit certain applications as the length-reduction in wires leads to lower energy dissipation [10, 11]. Heat removal and I/O resource allocation remains a challenge for such integration schemes. Innovative circuit techniques have also contributed to more efficient data communication. On-chip transmission lines have shown near speed-of-light latency and high throughput, but this comes at the cost of significant wire resources [12–14]. Low-swing signaling methods reduce the voltage level primarily as a power-reduction technique, but often have higher latency and reduced noise margins [15–18]. Other solutions have combined CMOS repeaters with channel equalization techniques.

Furthermore, innovative device structures or new nano-materials have already found their way into prototype VLSI applications in an effort to reduce power dissipation [19, 20]. Improving fundamental material properties are expected to become more important in highly scaled technologies. At narrow line widths, surface scattering of conducting electrons are projected to be a major concern, drastically increasing the effective resistivity of copper interconnects [21, 22]. This results from a combination of smaller cross-sectional dimensions and increased liner thicknesses. Heat management will also be increasingly important, as higher energy dissipation of wires and poor thermal conductivity of low-K insulators contribute to substantial temperature increases.

Among many materials, carbon-based materials such as graphene or carbon nanotubes have received much attention in recent years as a replacement for copper interconnects. Graphene is a planar sheet composed of carbon atoms. Graphene exhibits ballistic transport [23, 24], high intrinsic mobility [24, 25], high thermal conductivity [26, 27], and high current carrying capacity [28–30], making it attractive not only for transistors [31, 32] but also for interconnects [33, 34] and even as a thermal interface material [35]. Theoretical projections show that at small line widths (< 8nm), graphene will outperform copper with a 1:1 aspect ratio [33].

Carbon nanotubes are formed when a planar sheet of carbon atoms are rolled up in

a cylindrical tube. Graphene and carbon nanotubes share many excellent properties but graphene is more attractive from a manufacturing standpoint. Carbon nanotubes are chemically stable but it is extremely difficult to control their size and placement. Graphene can be grown in large sheets [36, 37] and then be subsequently patterned and etched using standard lithography methods. This results in better control and higher reproducibility of graphene devices.

Despite these excellent properties, the fabrication process is not well controlled at the level required for integrated circuits. Befittingly, the majority of graphene research focuses on methods to improve the material quality or finding innovative device architectures. To explore the full potential of graphene-based electronics, the research focus must extend beyond materials and devices. We need to find new promising applications and understand their requirements throughout all phases of the design from material to system. One of the major advantages of graphene over other nanomaterials is that graphene can be lithographically processed. This allows an easy path toward integration with existing silicon technology. DC characterization of sub-50nm graphene interconnects has been reported [28, 34], but very few studies exist on evaluating their performance when integrated with CMOS. Integrating graphene with CMOS is a critical step in establishing a path for graphene electronics. Chen et al. have reported the first integrated graphene/CMOS system [38]. They use CMOS ring oscillators to indirectly measure the performance of short graphene wires. More importantly, studying the performance of graphene under real workloads is needed but demonstration of a full system using graphene has not been made. Although several reports exist on graphene applications [32, 39, 40], these are generally prototypes that have limited functionality and only use a few devices. Developing a complete graphene-based system not only helps establish graphene as a viable interconnect material but also provides a general roadmap for material, circuit, and system design.

1.2 Thesis Contributions

The objectives of this thesis are to characterize the performance of graphene interconnects and demonstrate a complete graphene/CMOS application. This thesis contributes in the following areas:

(1) Monolithic Integration with CMOS.

Providing a path toward integration is critical in establishing the use of graphene as interconnects. Here, we demonstrate monolithic integration of graphene with CMOS on two prototype test chips. The purpose of the first test chip is to characterize the performance of long graphene wires. Off-chip measurements have limited scope and often require expensive equipment. The first test chip provides a platform to directly measure the delay and energy associated with driving a signal on long graphene wires. The second test chip demonstrates a complete system application. Large sheets of graphene are synthesized and then transferred to the CMOS chip. We then use standard lithography steps to pattern narrow graphene wires and connect them with the underlying CMOS circuitry. Details of the process flow are outlined in Chapter 3.

(2) Characterization of Multilayer Graphene Interconnects.

In this work, we grow large-area graphene sheets by chemical vapor deposition [36, 37, 41, 42]. The underlying catalyst film differs among the various growth methods, but Cu foils are a popular choice since they yield highly uniform monolayer graphene sheets [41, 42]. However, the monolayer sheet needs to be transferred multiple times to achieve a lower sheet resistance. In contrast, the use of Ni catalyst films generally produces a thick multilayered stack of graphene sheets and does not require multiple transfers. Both methods are used throughout this thesis. Here, we apply the term 'multilayer' to indicate graphene sheets with more than 10 layers. While graphene interconnects using monolayer and few-layer sheets have been previously characterized [34, 43], no studies exist on using thick multilayer graphene sheets as interconnects. In

this thesis, we conduct both off-chip and on-chip measurement of multilayer graphene interconnects. We characterize the properties of multilayer graphene sheets as well as long graphene wires. We implemented a CMOS test chip onto which 1 mm length graphene wires are monolithically integrated. Unlike Chen's work [38], this test chip focus on end-to-end data communication on medium to long multilayer graphene wires. The performance of each graphene wire is measured in detail, using isolated transmitters and receivers.

(3) Demonstration of Graphene-based Subthreshold System.

The analysis and results from Chapter 2 and Chapter 5 point to the large wire resistance as a major limitation for using graphene for high-speed communication. Unless very thick stacks of high quality graphene layers can be fabricated, the sheet resistance of a multilayer stack cannot match that of a Cu wire. Instead, another way to leverage graphene wires is to fabricate ultra-thin wires which have low wire capacitance. The low capacitance of few-layer graphene devices offers great opportunities for ultra-low power applications, which often have moderate frequency requirements. In Chapter 2, we briefly discuss this trade-off between speed and energy and suggest that ultra-thin graphene wires can provide significant energy reduction in subthreshold applications. Furthermore, we develop a second CMOS chip that operates in subthreshold and takes of advantage of few-layer graphene interconnects. This test chip presents the first experimental demonstration of a system application using graphene devices. Graphene is monolithically integrated as part of the interconnect fabric in a field-programmable gate array (FPGA). An FPGA has a highly interconnect-centric architecture making it an ideal test vehicle for graphene integration. Interconnect delay is a significant portion of the delay due to multiple routing segments in an FPGA. Furthermore, global interconnects have been shown to dominate the total power consumption in FPGAs [5,6].

This thesis is organized as follows. Chapter 2 describes a physics-based circuit model for graphene and compares its performance with Cu interconnects. In Chap-

ter 3, we discuss various graphene synthesis methods and outline the process flow used in this work. The monolithic integration process is also explained. Next, we describe the characteristics of multilayer graphene sheets and wires in Chapter 4 and present the results for integrated graphene data links in Chapter 5. Chapter 6 then explains the FPGA architecture and measured chip results. Finally, Chapter 7 concludes the thesis.

Chapter 2

Benchmarking Graphene Interconnects

Graphene has large conductivity and large current capacity making them attractive for interconnect applications. Many reports highlight the potential of graphene but experimental results show that the resistivity of graphene is still quite larger than that of Cu. This chapter uses a physics-based circuit model to project and compare the performance of graphene and Cu interconnects.

2.1 Modeling Graphene Interconnects

2.1.1 Physics-Based Circuit Models

An accurate model is needed in order to benchmark the potential performance of graphene interconnects. Here, we use the well known physics-based model presented by Naeemi et al. [33,44,45]. Figure 2-1 shows the equivalent circuit model for quantum wires including graphene or carbon nanotubes [46]. The value of the circuit parameters depend on the electronic band structure of the material.

When a net current exists in a quantum wire, the kinetic energy of the electrons $(1/2LI^2)$ manifests itself in the kinetic inductance L_K. This can be observed at high frequencies and in high mobility conductors such as superconductors. For graphene
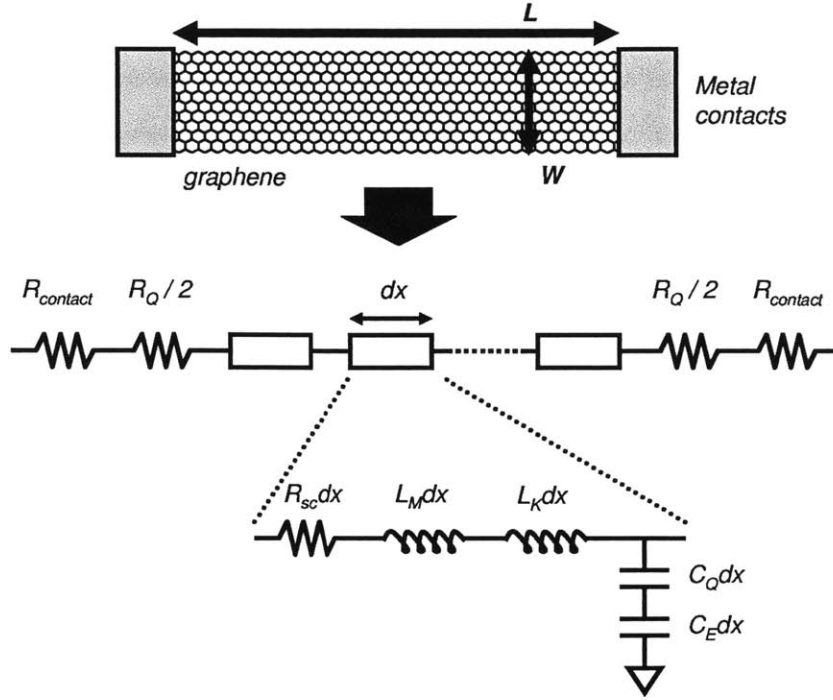
Figure 2-1: Equivalent circuit model for a graphene wire.

and carbon nanotube devices, the kinetic inductance is usually much larger than the magnetic inductance. For most practical dimensions, the frequency at which the inductive effects begin to be important is usually in the THz range or in some cases several hundreds of GHz. This is well beyond the practical range for most applications and hence will not be considered throughout this thesis. In addition, the wires are assumed to be long enough compared to the mean free path that the contact resistances can be ignored. Contact resistances as small as a few hundred ohms is reported [47, 48]. Although the contact resistance is expected to rise at smaller line widths, the exact values also depend on the fabrication process and are difficult to precisely model.

The quantum resistance, quantum capacitance, and kinetic inductance are determined by the total number of conduction channels in the device, which is in turn determined by the chirality and width of the graphene device. The chirality, or configuration, of the graphene ribbon depends on the pattern of the edge which can be in an armchair or zigzag configuration (Figure 2-2). While all zigzag edged graphene

26

devices are metallic, an armchair device may be metallic or semiconducting. An armchair device is metallic if the number of carbon atoms across its width is 3p+2, and semiconducting if the number is 3p or 3p+1, where p is an integer.
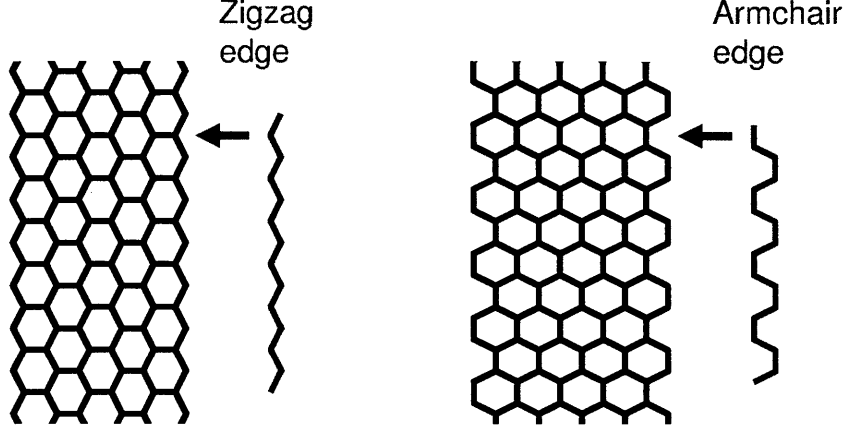
Figure 2-2: Diagram of zigzag and armchair configuration for graphene wires.

The electrostatic capacitance C_E is determined by the surrounding materials and geometry. In addition to C_E, in a quantum system one must add an electron at an available quantum state above the Fermi energy due to the Pauli exclusion principle. This additional extra energy cost can be equated with an effective quantum capacitance. This quantum capacitance C_Q can be expressed as:

$$C_Q = \frac{4e^2}{hv_f} N_{ch} \simeq (200aF/\mu m)\, N_{ch} \tag{2.1}$$

where e is electron charge, h is the Plank constant, v_f is the Fermi velocity in graphene ($\simeq 8 \times 10^5 m/s$), and N_{ch} is the number of conduction channels. Similarly, the quantum resistance R_Q is the resistance of an ideal quantum wire with no scattering and equals [49]:

$$R_Q = \frac{h/2e^2}{N_{ch}} \simeq \frac{12.9k\Omega}{N_{ch}} \tag{2.2}$$

In virtually all practical wires, electrons will get scattered by phonons, defects, and rough edges. The scattered resistance per unit length is $r_{sc} = R_Q/\lambda_{eff}$ where λ_{eff} is the effective mean free path and is modeled in 2.1.3. The total resistance and

capacitance of the wire then becomes:

$$R_{tot} = R_Q + r_{sc} \tag{2.3}$$

$$= R_Q \left(1 + \frac{L}{\lambda_{eff}} \right) \tag{2.4}$$

$$C_{tot} = C_Q // C_E \tag{2.5}$$

$$= \frac{C_Q C_E}{C_Q + C_E} \tag{2.6}$$

Both the conductance (or $1/R$) and quantum capacitance scale linearly with the number of conduction channels. The number of conducting channels or modes is a function of the chirality and width of the device and can be expressed using Fermi-Dirac statistics as:

$$N_{ch} = N_{ch,electron} + N_{ch,hole} \tag{2.7}$$

$$= \sum_n \frac{1}{exp\left(\frac{E_{n,electron} - E_F}{k_B T} \right) + 1} + \sum_n \frac{1}{exp\left(\frac{E_F - E_{n,hole}}{k_B T} \right) + 1} \tag{2.8}$$

where $E_{n,electron}$ ($E_{n,hole}$) is the minimum (maximum) energy of the nth conduction (valence) subband, E_F is Fermi energy, $k_B T$ is thermal energy, and n is an integer. Using a tight-binding approximation, the subband energy can be calculated as [44]:

$$E_n = \frac{h v_f}{2W} |n + \beta| \tag{2.9}$$

where β is 1/3 and 0 for semiconducting and metallic devices respectively. Several modifications need to be made to this equation especially when the line width becomes very small. When graphene is patterned into small ribbons, this geometric confinement causes the electronic band structure to change. In reality, the carbon atoms at the edge are spaced slightly closer than the atoms in the middle [50].This shifts the subbands, but most importantly, this opens up a gap in metallic 3p+2 armchair devices, which was experimentally verified in [51]. For zigzag graphene ribbons, a small gap also appears due to the staggered sublattice potential from magnetic or-

dering. The simple tight-binding models are accurate unless graphene ribbons that have a narrow width (< 5 nm) and low Fermi energies. The exact equations are found in [44].
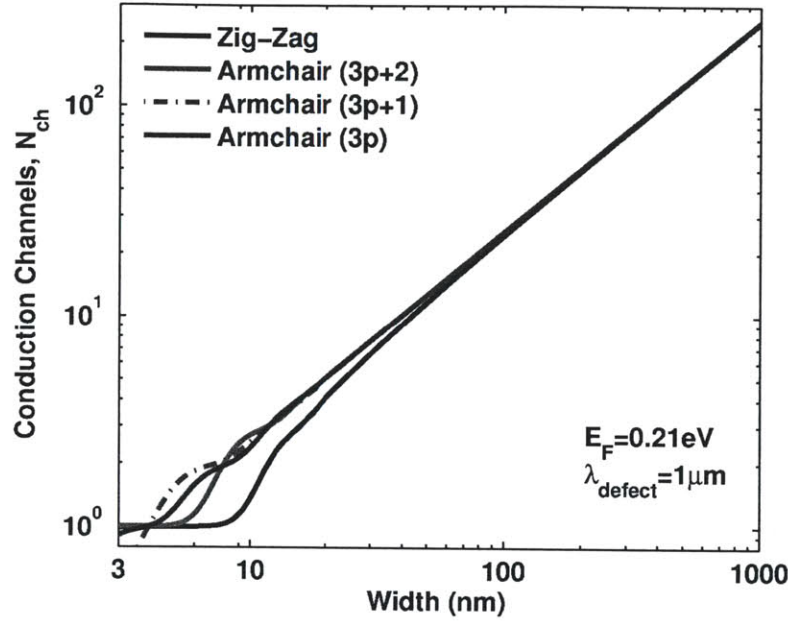
Figure 2-3: Number of conduction channels in a graphene nanoribbon.

Figure 2-3 shows the number of conduction channels as a function of graphene width based on the equations above. We chose an arbitrary Fermi energy value of 0.21 eV. Although the exact number of conduction channels depends heavily on the Fermi energy, the qualitative results do not change. The effect of varying the Fermi energy is discussed in the next section. In Figure 2-3, at large widths, the difference between semi-conducting and metallic devices disappears. As the width decreases, the band gap opening becomes more pronounced. For armchair graphene wires, semiconducting ribbons have larger quantum conductances compared to metallic wires, which appears counter intuitive. The reason behind this is that there are smaller gaps between subbands in semiconducting devices than those in metallic ones. Thus, depending on the Fermi energy, more subbands may be populated in semiconducting devices. Graphene nanoribbons with rough edges all become semiconducting [47,52,53], which

may not be problematic for interconnect applications since semiconducting wires conduct as well as metallic wires. [47] suggests that the detailed edge structure plays a more important role than crystallographic direction in determining the properties of GNR. Theory supports this and predicts the energy gap depends sensitively on the boundary conditions at the edges.

The difference between work functions of graphene and the substrate causes some charge to get trapped at the interface [52,54]. This causes the Fermi energy to shift from zero, where E_F of 0.13 eV [55], 0.21 eV [52], and 0.4 eV [54] have been previously observed. The shift in E_F is associated with the surface charges at the interface rather than the carrier concentration of the substrate. Wang et al. have observed that the conductance per layer saturates as the number of graphene layers increases which suggests that the conduction of graphene sheets is limited by the substrate [48]. A nominal value of E_F=0.2 eV is assumed in this chapter.

2.1.2 Multilayer Stacks

The model presented in the previous section assumes a monolayer graphene interconnect. Multilayer graphene wires can offer lower resistance, and ultimately, a thicker stack is more reliable for large-scale manufacturing. Depending on the stacking order, a multilayer stack of graphene turns in to graphite and the increased intersheet electron interactions lower the conductivity per layer [56]. Therefore, to take full advantage of multilayer graphene devices, the adjacent graphene layers must be non-interacting and electronically decoupled.

To date, the interaction between monolayer and multilayer flakes of graphene is not well understood. Some groups have provided theoretical and experimental evidence suggesting that the transition from graphene to graphite occurs around seven to eight layers of graphene [57,58]. Zhou et al. have suggested that epitaxial graphene behaves as bulk graphite beyond five layers [54]. In contrast, others have demonstrated electronically decoupled multilayer graphene films [37,59] which shows great promise.

Understanding the nature of multilayer stacking has a subtle yet very important

effect on the analysis of graphene wires. Xu et al. assumes that the multilayer graphene device is neutral (E_F=0eV) and extracts the mean free path from conductivity values of bulk graphite [60]. This results in overly pessimistic projections of multilayer graphene devices compared to Cu wires. In contrast, Tanachutiwat and Wang models the Fermi level shift resulting from multilayer stacks of graphene [61] and conclude more favorable results for multilayer graphene interconnects than Xu et al.

Throughout this chapter, we assume that each adjacent layer is decoupled [37,59] and assume each layer has the same parameters (i.e., E_F, P_{sc}, λ_{eff}, etc) that is equal to that of a high quality monolayer graphene device. Although this assumption can be readily validated for few-layer graphene wires (under ~10 layers) [37, 57–59], for high-performance applications, potentially hundreds of layers are needed to match the resistance of Cu wires.

2.1.3 Mean Free Path and Line-Edge Roughness

Rough edges can backscatter electrons and lower the effective mobility or mean free path. The detrimental effects of line-edge roughness have become more pronounced as the width of nano-scale devices continue to shrink. Controlling the edge of a graphene device is even more important since a rough edge occurs even when a single atom is displaced on the edge of a graphene wire. Recently, Ni nanoparticles have demonstrated the cutting and precise patterning of graphene devices [62]. Although this process achieves the atomic precision necessary to control the edge of a graphene device, this method lacks the control required for large-scale manufacturing. We must assume that some degree of backscattering will occur in the device, as smooth edges are extremely difficult to achieve if not impossible.

Experiments show that the intrinsic mean free path in graphene is in the μm range [25]. The mean free path of electron-phonon scatterings in graphene nanoribbons is expected to be extremely large and on the order of tens of μm [63] and hence has little effect on the overall mean free path. The effective mean free path can then be

modeled as [44]:

$$\frac{1}{\lambda_{eff}} = \frac{1}{\lambda_D} + \frac{1}{\lambda_{edge}} \qquad (2.10)$$

where the λ_D is the mean free path due to the substrate-induced disorders and defects and λ_{edge} is the mean free path associated with the edge roughness. Here, λ_D is assumed to be 1 μm, where a value between 400 nm and 1.2 μm have been demonstrated experimentally [25, 52, 64]. The mean free path associated with the edge roughness for the nth subband becomes [32, 33, 52]:

$$\lambda_{edge,n} = \frac{W}{P_{sc}} \sqrt{\left(\frac{E_F}{E_n}\right)^2 - 1} \qquad (2.11)$$

P_{sc} is the backscattering probability and has a value between 0 and 1. A value of $P_{sc}=0$ indicates that the device has a smooth edge and no backscattering occurs, and $P_{sc}=1$ indicates that transport along the edges is fully diffusive. The equation above also indicates that λ_{edge} is proportional to the width of the device. This width dependence was similarly modeled in [60].

Figure 2-4 shows the calculated mean free path. The graphs shows that the mean free path decreases at smaller line widths only when P_{sc} 0 (i.e., when backscattering occurs). The roll off can occur at much smaller line widths depending on the Fermi energy or backscattering probability. Therefore, the resistance of the graphene wire can be improved by increasing the Fermi energy or by fabricating graphene devices with smoother edges. The Fermi energy can be modulated by electrostatic gating or by means of chemical doping.

Yang and Murali have experimentally demonstrated mobility degradation in graphene nanoribbons as a function of the device width [43]. The mobility is limited by edge scattering at smaller line widths as expected. Using the equations in this section and in 2.1.1, we can extract the mobility of a graphene device as $\mu = 1/en_{gr}\rho_{gr}$ where n_{gr} and ρ_{gr} are the carrier density and effective resistivity, respectively.

Figure 2-5 plots the data from [43] and the calculated mobility assuming a monolayer armchair wire with $P_{sc}=0.5$ and $n_{gr}=5\text{x}10^{12}$ cm^{-2}. Typical carrier densities
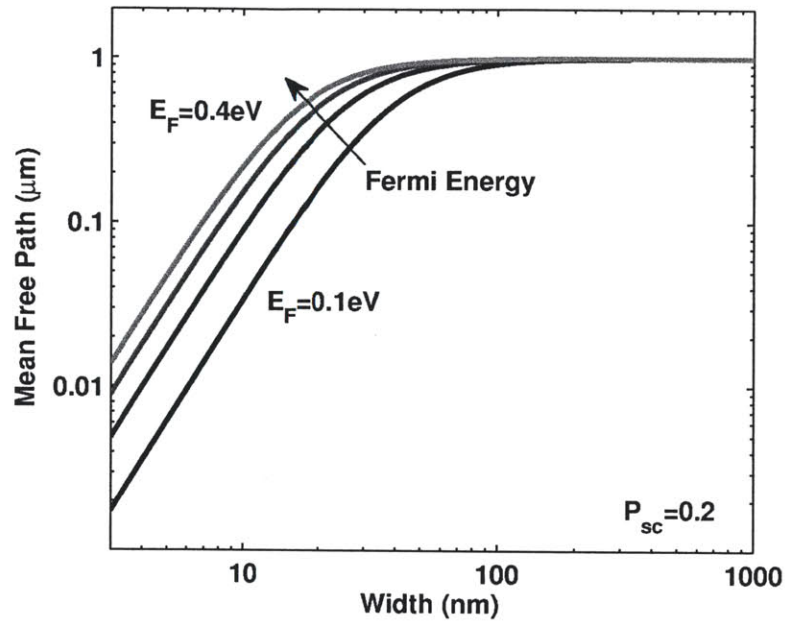
(a)
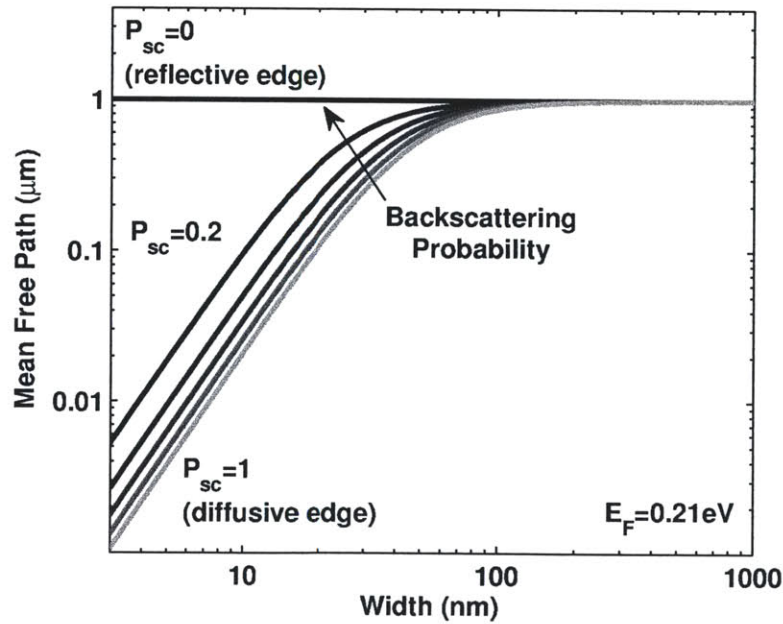
(b)

Figure 2-4: Calculated mean free path in a graphene nanoribbon. (a) A constant $\lambda_D = 1\mu$m is assumed while varying the Fermi level. (b) A constant $E_F = 0.21$ eV is assumed while varying scattering probability.
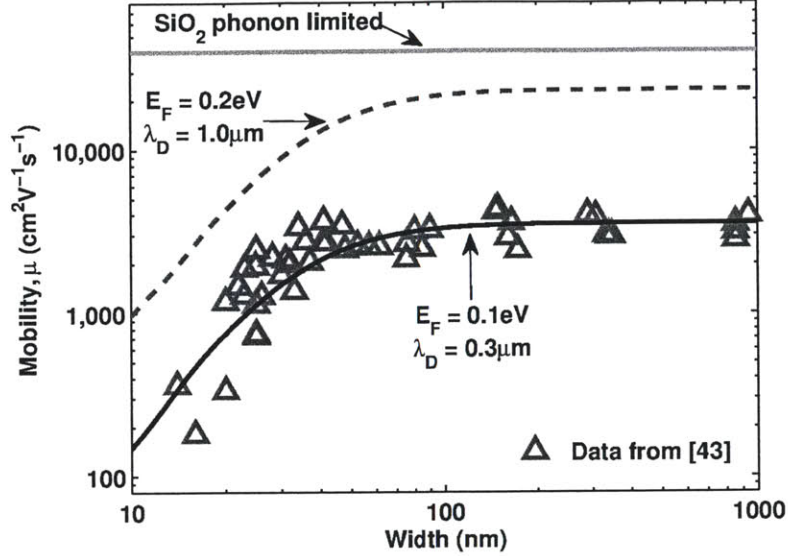
Figure 2-5: Mobility of graphene wires vs. wire width. Data points are from [43] and lines represent calculated mobility.

between 2×10^{11} and 9×10^{12} cm^{-2} have been reported [25, 36, 42, 47, 55, 64–66]. Due to phonon scattering of SiO_2, the room-temperature mobility limit of graphene on SiO_2 is 40,000 $cm^2V^{-1}s^{-1}$ [67] and is also plotted in Figure 2-5. The mobility degradation due to line-edge roughness is clearly visible when the width is below 50 nm. The calculated mobility fits the experimental data well when $P_{sc}=0.5$, $E_F=0.1eV$ and $\lambda_D=0.3\mu m$. Throughout Section 2.3 we assume that $P_{sc}=0.5$, $E_F=0.2eV$ and $\lambda_D=1.0\mu m$, which is a reasonable and yet optimistic projection. These values result in mobilities that are roughly 6x higher than the experimental data found in [43], but are comparable to those found in [25, 65, 66].

2.2 Modeling Copper Interconnects

Developing a closed form model for Cu is essential in projecting the resistivity values when the physical dimensions extend beyond the roadmap outlined by the International Technology Roadmap for Semiconductors (ITRS) [68]. The effective resistivity of a metal conductor is a strong function of the scattering processes at the surface

and grain boundaries. Such effects have been studied for a long time and a number of well-known models exist [22, 69–72]. Recently, Lopez et al. have added the effect of line-edge roughness [72], which has become increasingly more important.

When a Cu wire has an effective width and height of w_{Cu} and h_{Cu} respectively, and a bulk resistivity of ρ_0, the effective resistivity of Cu is given by [72]:

$$\rho_{Cu} = \frac{\rho_0}{\sqrt{1 - \left(\frac{LER}{w_{Cu}}\right)^2}} \left[G(\alpha) + 0.45\lambda_{Cu}(1 - p_{Cu}) \left(\frac{1}{h_{Cu}} + \frac{1}{w_{Cu}\left[1 - \left(\frac{LER}{w_{Cu}}\right)^2\right]} \right) \right]$$

(2.12)

where $G(\alpha)$ is the grain boundary component defined as [71]:

$$G(\alpha) = \frac{1}{3}\left[\frac{1}{3} - \frac{\alpha}{2} + \alpha^2 - \alpha^3 ln\left(1 + \frac{1}{\alpha}\right)\right]^{-1}$$

(2.13)

and α is given by:

$$\alpha = \frac{\lambda_{Cu}}{d_{Cu}} \frac{R_{Cu}}{1 - R_{Cu}}$$

(2.14)

where LER is the line-edge roughness amplitude, λ_{Cu}=40 nm is the bulk mean free path in copper [69], and d_{Cu} is the average separation of the grain boundaries and can be approximated as $\sim w_{Cu}$. The two primary parameters used to model and fit experimental data to is R_{Cu} and p_{Cu}. R_{Cu} is the fraction of electrons scattered at the grain boundary and p_{Cu} is the fraction of electrons elastically scattered. R_{Cu} is the grain reflectivity, where R_{Cu}=1 indicates that an electron will experience complete reflection within a grain. p_{Cu} is specularity, where a value of 0 indicates diffuse (inelastic) scattering and electrons completely lose their drift velocity.

2.3 Comparison of Copper and Graphene Interconnects

2.3.1 Sheet Resistivity

Recent demonstration of sub-50 nm graphene interconnects show that the best devices are comparable to copper in terms of their resistivity [34]. Although such reports show great promise of graphene as an interconnect material, comparing the resistivity often overlooks one of the most important challenges of graphene. Ultimately, thick stacks of multilayer graphene are needed to compete with Cu and yet no experimental demonstration has come close. In commercial CMOS technologies, it is often more useful to report the two-dimensional sheet resistivity (R_{sh}) since the height of each metal layer is fixed. The sheet resistivity is a function of the material properties and its thickness ($R_{sh}=\rho/thickness$).

Figure 2-6 plots the sheet resistance of various graphene samples found in literature. Sheet resistance as low as 30 Ω/sq was produced from HNO_3-doped 4-layer graphene sheets fabricated from a 30-inch graphene film [55]. Most commercial CMOS technologies have sheet resistances less than 0.1 Ω/sq for all metal layers [73], although this number is expected to increase at future technology nodes. The sheet resistance of graphene devices is generally 3-4 orders of magnitude higher than this limit primarily because the graphene films reported in literature are typically very thin and composed of 1-10 layers. As discussed in Section 2.1.2, fabricating thick multilayer graphene stacks that do not turn in to graphite is extremely difficult and has not yet been demonstrated.

As a result, one of the most promising applications of graphene and certainly the closest to reaching the market has been transparent electrodes. Transparent electrodes are widely used in displays, touch panels, and solar cells. Due to the limited supply and high cost of indium tin oxide (ITO), the standard material for transparent electrodes, graphene has been actively pursued as a low-cost alternative. In addition, recent demonstration of graphene-based touch-screen panels shows that graphene is

36
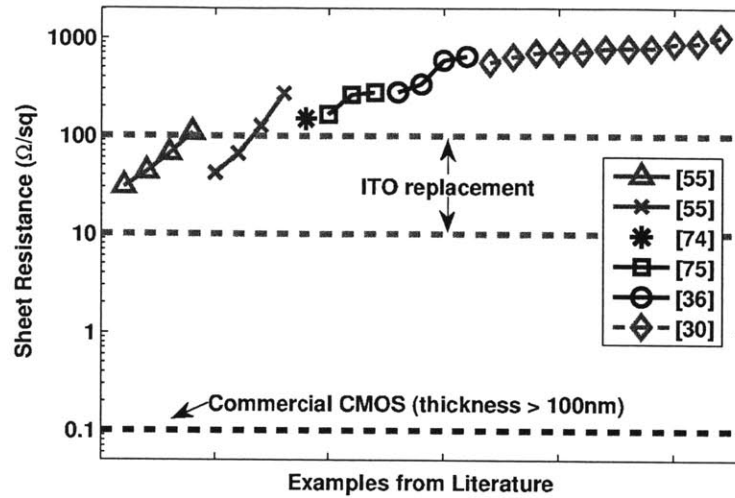
Figure 2-6: Sheet resistance of various graphene samples.

more tolerant to strain than ITO [55]. The required sheet resistance to replace ITO is typically between 10 and 100 Ω/sq [74]. The inherent requirement for having a thin and transparent metal conductor results in a sheet resistance for transparent electrodes that is much higher than what is required for CMOS interconnects. Co-incidentally, because of this requirement, few-layer graphene devices are a perfect candidate to replace ITO.

Figure 2-7 shows the sheet resistance of graphene and Cu at the extreme limits. In the optimistic case (E_F=0.2eV, λ_D=1μm), the effective R_{sh} for graphene is slightly lower than that of Cu. When the thickness is less than 1 nm, the data from [55] outperforms Cu at those thicknesses. A single graphene layer is one atom thick and represents the ultimate limit of a two-dimensional material. While traditional inter-connects are fabricated by evaporating a bulk metal source, single crystalline graphene films can be synthesized resulting in superior performance. Although existing fabrica-tion methods have produced highly uniform few-layer graphene films, thick multilayer graphene films have not been demonstrated. At the 11 nm node, the effective sheet resistance is expected to rise to 1 Ω/sq. In order to reach 1 Ω/sq, roughly 50 layers of graphene is needed in the optimistic case or 30 layers when the performance is limited by SiO$_2$. Reliably fabricating such thick layers remains a challenge.
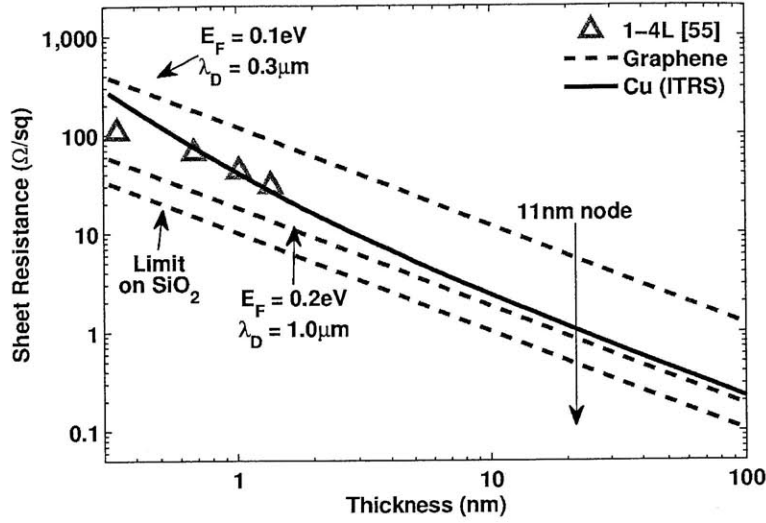
Figure 2-7: Sheet resistance of graphene and Cu. Data points are measured results from [55].

2.3.2 Wire Resistance and Capacitance

In this section, we compare the performance of graphene and Cu wires using a typical wire structure with adjacent ground signals and ground planes (Figure 2-8). The interlayer dielectric constant is $\kappa=2.2$ and the aspect ratio is fixed at 2 for Cu wires.

Figure 2-9 shows the resistance of graphene and Cu wires. When graphene is limited by the SiO_2 substrate, graphene begins to outperform Cu below 14 nm wire width for 10-layer graphene wires and below 4 nm wire width for monolayer graphene interconnects. As more layers of graphene are used, the width at which graphene begins to outperform Cu increases. However, these projections are based on the upper limit of graphene. ITRS projections also assume high quality Cu wires where scattering induced by line-edge roughness is limited. In Figure 2-9b, we assume more realistic values equivalent to those used in Section 2.1.3 and Section 2.2. Under these assumptions, line-edge roughness severely limits the Cu wires below 14 nm, resulting in the rapid rise in resistance. Both monolayer and 10-layer graphene wires show better performance than Cu wires below 14 nm.

The exact point at which graphene begins to outperform Cu is subject to change
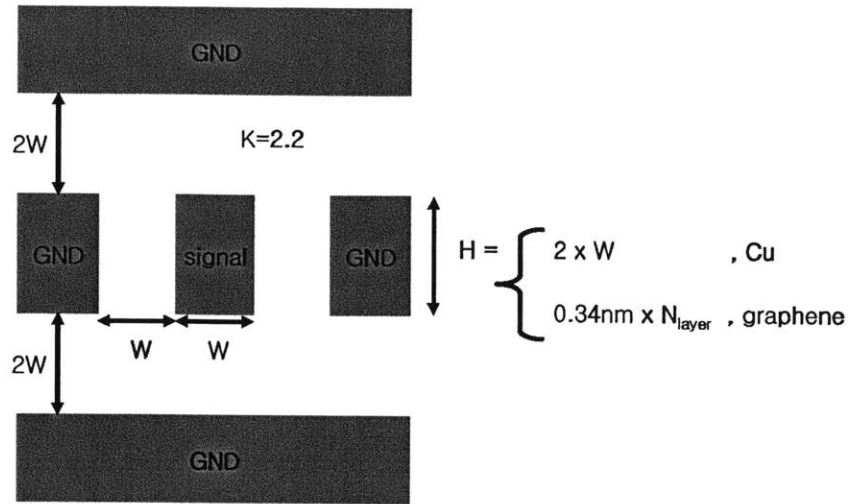
Figure 2-8: Cross section of interconnect and surrounding ground planes. The dielectric constant is $\kappa=2.2$ and the aspect ratio is 2 for Cu wires. The estimated height H of the graphene wire is the product of the interlayer spacing 0.34 nm and the number of layers.

as the modeled parameters change. However, in general, this crossover is not expected to occur until around 10 nm which is consistent with [33]. In contrast to resistance values, the capacitance of the graphene wire is known to be significantly less than that of a Cu wire. Figure 2-10 shows the capacitance per unit length as a function wire width assuming the geometry outlined in Figure 2-8. The capacitance of the Cu wire is nearly constant across all wire widths because the geometry scales accordingly. Recall that the total capacitance of a graphene wire can be expressed as the series combination of the electrostatic capacitance and quantum capacitance. As the wire width decreases, the capacitance of the graphene wire slightly increases since we assume that the graphene wire has constant thickness in contrast to the Cu wire, which has a constant aspect ratio. The difference between a monolayer and 20-layer graphene wire is relatively small when $W=100$ nm. When $W=100$ nm, the capacitance of a 20-layer graphene wire is only 7.4 % higher than that of a monolayer graphene wire. This difference becomes 79.6 % when $W=10$ nm. Overall, if we use less than 10 layers, the capacitance of a graphene wire is 2x lower than that of a Cu wire.

Although more layers are needed to reduce the overall resistance, adding more
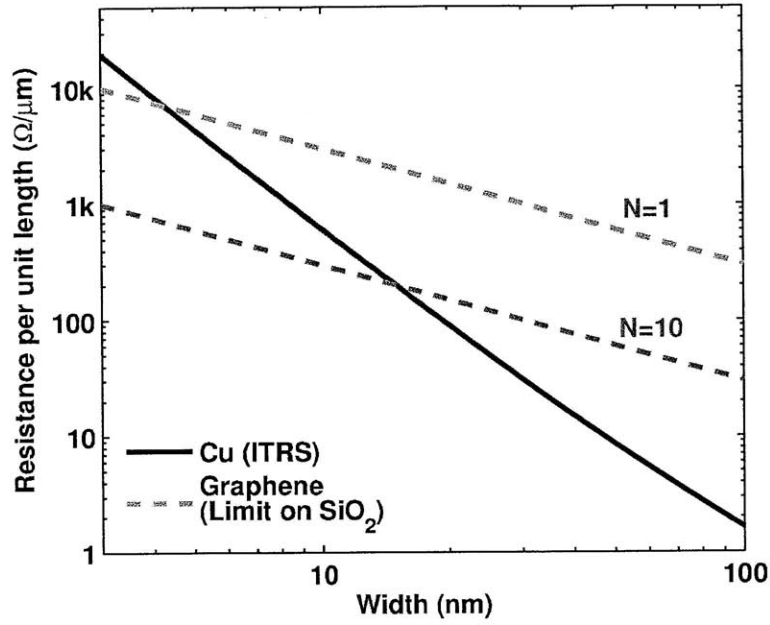
(a)
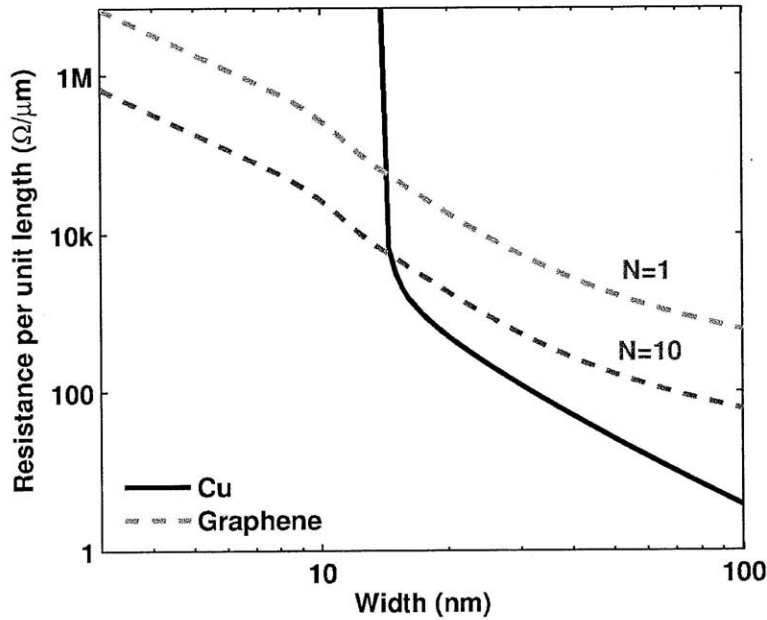
(b)

Figure 2-9: Resistance per unit length. (a) Graphene is calculated as the limit on SiO_2. Cu model assumes ITRS projection where (p_{cu}=0.95,R_{cu}=0.45,LER=0, H=2xW). (b) zigzag graphene assumes E_F=0.2eV, λ_D=1 μum, P_{sc}=0.5. Cu model assumes experimental values from [72] where p_{cu}=0,R_{cu}=0.79, and LER=14nm.

40
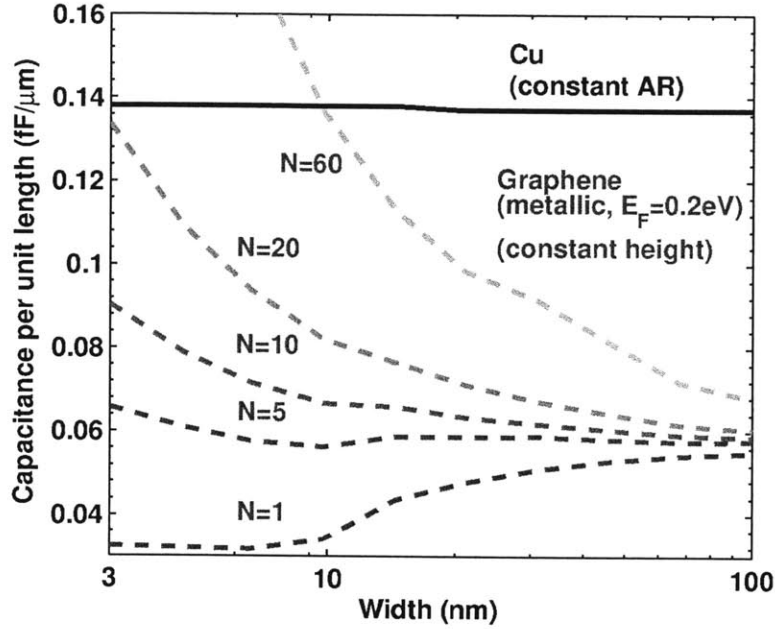
Figure 2-10: Capacitance of graphene and Cu wires. Cu wires assume a constant aspect ratio=2. Graphene wires are assumed to be zigzag with E_F=0.2eV.

layers also increases the capacitance and energy dissipation. Figure 2-11 plots the wire capacitance at a fixed width of 10 nm. Generally, the quantum capacitance is much larger than the electrostatic capacitance and thus has very little effect on the overall capacitance especially for thicker graphene films. As the number of layers increase, the capacitance also increases. In contrast, the resistance shows a more pronounced decrease as the number of layers increase (Figure 2-12). If we assume a smooth edge with no scattering, roughly ~16 layers is needed to match the resistance of a Cu wire. However, more than 60 layers is required when the wire is completely diffusive. Thus, fabricating multiple graphene layers with very little interlayer and line-edge scattering is necessary to have resistance values comparable to that of Cu wires.

Similarly, the combined effect of the wire resistance and capacitance is shown in Figure 2-13. Because the wire capacitance shows a rather weak dependence on the number of layers, the RC time constant shows a similar form as the wire resistance. Nonetheless, the small capacitance of graphene does lower the overall RC time con-

41
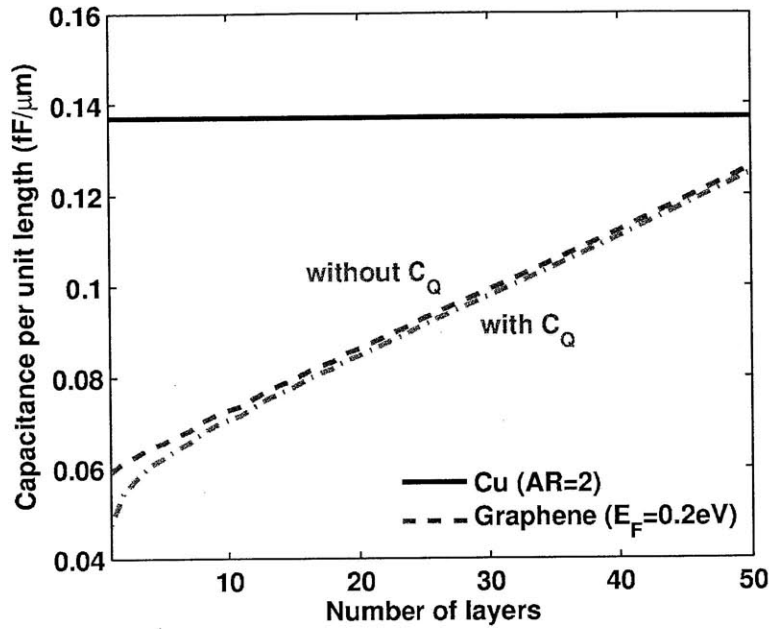
Figure 2-11: Wire capacitance as a function of number of graphene layers. Wire width is fixed at 10 nm. Cu wire assumes an aspect ratio of 2.
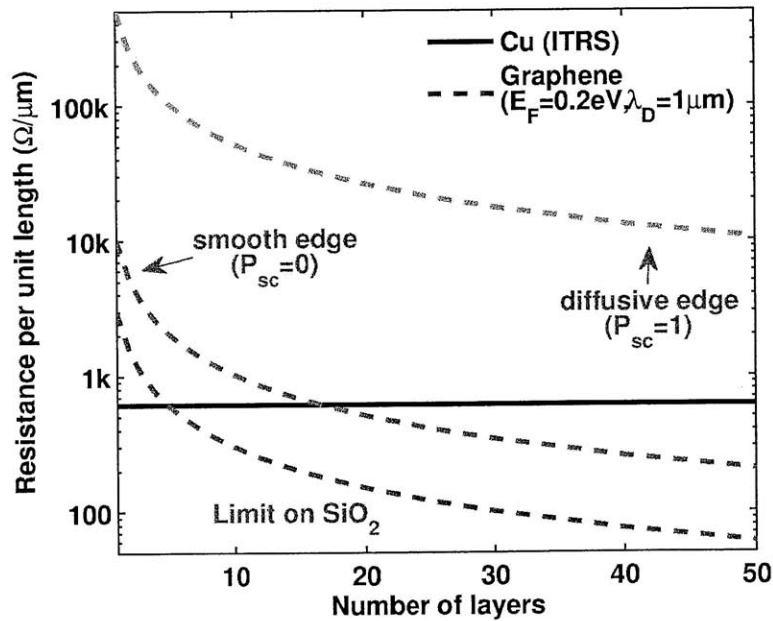
Figure 2-12: Wire resistance as a function of number of graphene layers. Wire width is fixed at 10 nm. Cu wire assumes an aspect ratio of 2.
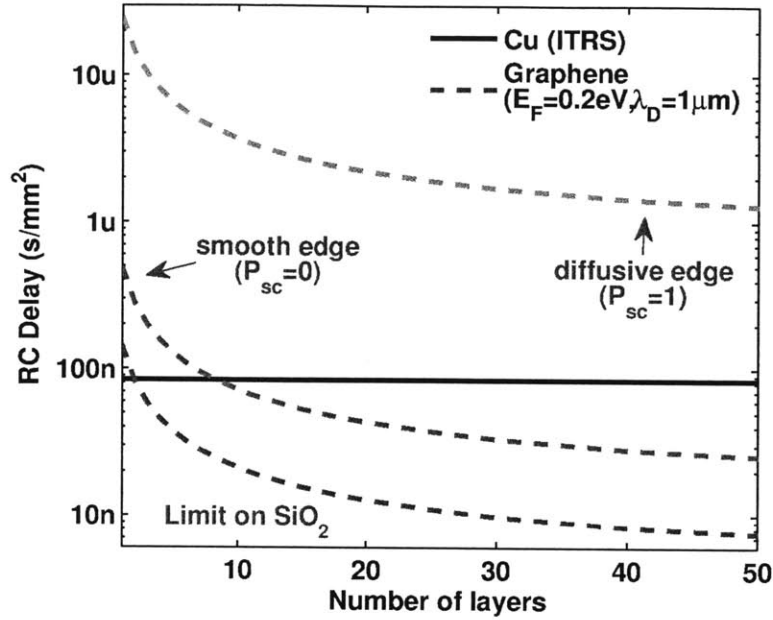
Figure 2-13: RC delay as a function of number of graphene layers. Wire width is fixed at 10 nm. Cu wire assumes an aspect ratio of 2.

stant and only ~ 8 layers is needed to match the RC time constant of a Cu wire when a smooth edge is assumed. When the wire is completely diffusive, more than 60 layers is still required to match the performance of Cu wires. When graphene approaches its limit on SiO_2, only two layers of graphene is needed to match the RC time constant of Cu. This suggests that the graphene films needs to be of extremely high quality.

2.3.3   Interconnect Performance for Subthreshold Circuits

Enhancing the quality of the graphene wire and fabricating multilayer stacks is critical for lowering the delay of graphene wires. Overcoming these challenges may prove to be too difficult as no known solutions exist yet. One attractive alternative is to take advantage of the small capacitance of graphene wires. Chip makers will often produce two different silicon technologies depending on the system needs. For example, interconnect stacks are often optimized either for RC performance or wire density [75]. In high-performance applications, thick wires are used to optimize the

43

resistance and capacitance of the wire. However, the energy of a wire is predominantly determined by its capacitance. Lowering the capacitance is more advantageous in low-power applications although this also results in a higher wire resistance.

From earlier discussions of Figure 2-10, roughly 10 layers or less is needed at small dimensions for graphene wires to have a competitive edge over Cu in terms of wire capacitance. When 10 or fewer layers are used, the graphene film must have extremely high quality to have a resistance value comparable to that of Cu. However, these design margins may be relaxed for subthreshold circuits, where the supply voltage is less than the threshold voltage of a transistor. Subthreshold applications often have modest speed requirements. Subthreshold circuits are inherently slow due to the decreased current levels and can tolerate more resistive wires. The use of low-capacitance wires is fitting because the primary motivation is to lower the energy dissipation in such systems.

Naeemi has previously suggested changing the aspect ratio or using carbon-based devices for subthreshold circuits [76]. The analysis presented in the previous section assumed a fixed aspect ratio of 2 for Cu wires. At a fixed wire width, the aspect ratio can be decreased by decreasing the height of the bulk Cu wire or by stacking less layers of graphene.

Figure 2-14 shows the capacitance as a function of the wire aspect ratio. Wire capacitance decreases as the aspect ratio decreases because the coupling and fringe capacitance is reduced. Graphene has a slightly smaller wire capacitance than that of Cu because of the quantum capactiance of graphene. However, except at very small aspect ratios, the difference in wire capacitance between Cu and graphene is very small because the wire capacitance is mostly determined by the electrostatic capcitance. Therefore, graphene wires do not have a significant advantage over Cu wires in terms of capacitance since capacitance is predominantly determined by the wire geometry.

However, graphene is an ideal candidate for low-capacitance wires because it is intrinsically very thin. Graphene sheets are one atom thick and the ability to fabricate such wires with atomic precision presents a significant manufacturing advantage over
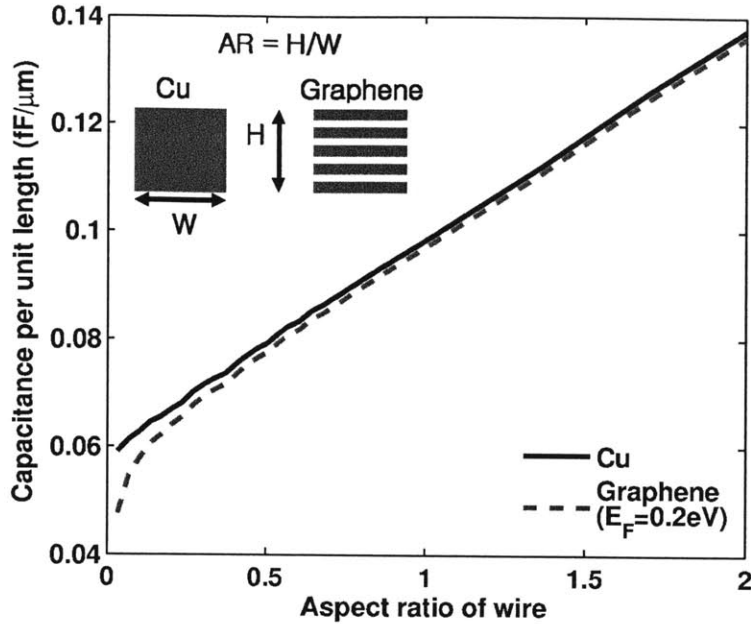
Figure 2-14: Wire capacitance as a function of aspect ratio. As the aspect ratio increases, the height of the Cu wire increases or more layers of graphene is used.

bulk materials. The back-end process flow in a CMOS technology typically includes physical deposition of metallic sheets, which result in a polycrystalline film. Therefore, bulk materials such as Cu suffer at such small thicknesses (Figure 2-7). Figure 2-15 shows the wire resistance as a function of the wire aspect ratio. Wire resistance increases as the aspect ratio decreases. When we assume smooth line edges for both graphene and Cu wires, graphene wires are projected to have lower resistance than Cu wires. In subthreshold circuits, although the wire resistance is often not the dominant resistance term, this may become problematic for thin Cu wires. The actual resistance of extremely thin Cu wires is likely significantly higher than what we project here if edge scattering effects are considered. When fabricating extremely thin wires for subthreshold circuits, graphene is more reliable and robust and thus a better candidate for low-capacitance wires.

When a CMOS inverter is driving a distributed RC wire (Figure 2-16), the RC time constant is given by [77]:

$$\tau = R_{on}C_w + \frac{R_w C_w}{2} \tag{2.15}$$

45
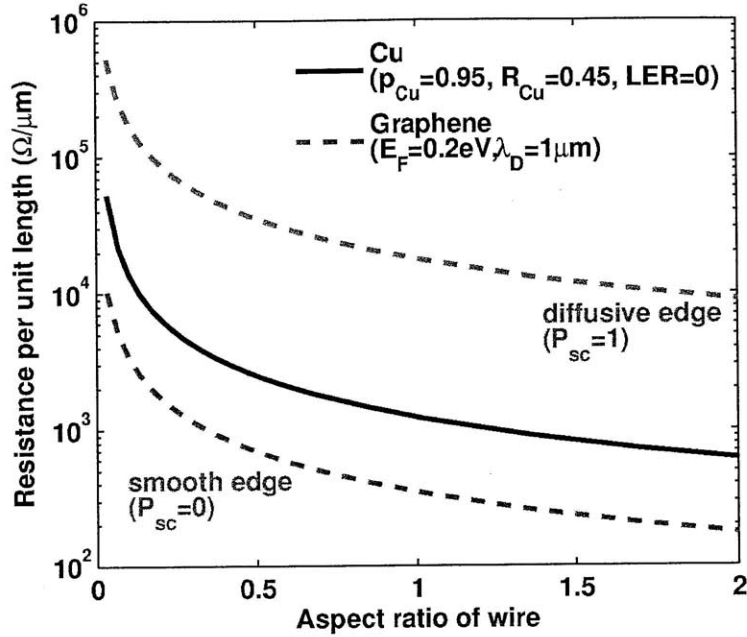
Figure 2-15: Wire resistance as a function of aspect ratio.

where R_{on} is the equivalent on resistance of the transistor and R_w and C_w are the distributed resistance and capacitance respectively. At the nominal supply voltage, R_{on} is typically a few kΩ and the wire resistance dominates. However, as the voltage is scaled below threshold voltage of the transistors, R_{on} increases exponentially and becomes the dominant resistance term as long as the wire resistance does not degrade significantly at very small thicknesses.
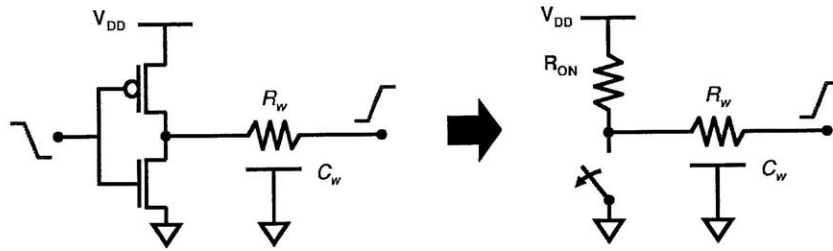
Figure 2-16: Diagram of CMOS inverter driving a distributed RC wire.

Figure 2-16 shows this effect in detail. We assume a simple buffer driving a distributed RC line at the 16 nm node [68, 78]. To compare the effect of the metal thickness, we compare 5-layer graphene wires with Cu wires with a nominal aspect

46

ratio of 2. The wire is modeled using a $\pi 3$ distributed RC model (Figure 2-17). The distributed resistance and capacitance values for the Cu wire are 197.3 kΩ and 138 fF respectively. The values for the 5-layer graphene wires are 7.3 MΩ and 61.8 fF respectively. The graphene wire has a resistance that is roughly 37x higher and capacitance that is 2.2x smaller than the Cu wire.
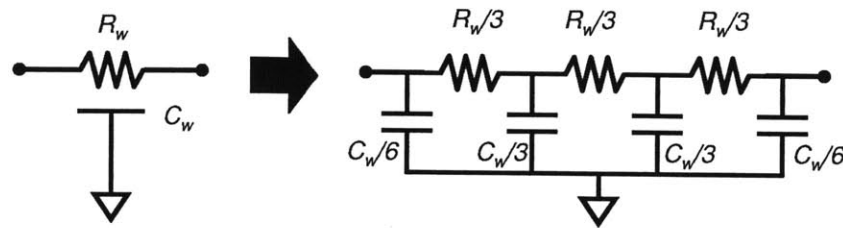
Figure 2-17: Equivalent $\pi 3$ RC wire model used for simulating a distributed RC wire with total resistance R_w and total capacitance C_w.
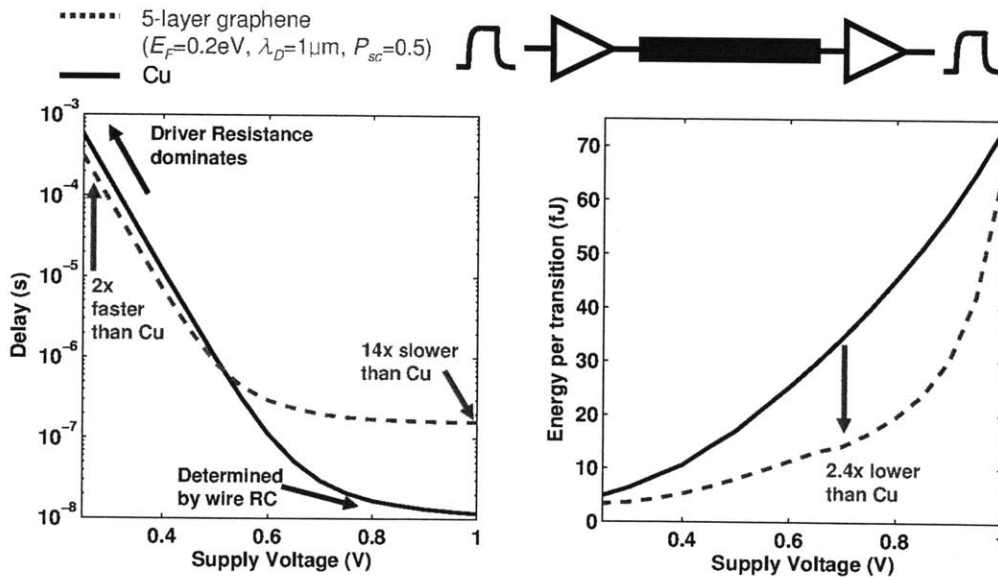
Figure 2-18: Performance of graphene and Cu wire in subthreshold.

For long interconnects, the energy is largely determined by the capacitance of the wire and favors graphene. At high supply voltages, the high resistance of the graphene wire negatively affects the RC delay. However, at low supply voltages, the large resistance of the transistors begins to dominate and the wire capacitance becomes important, resulting in a lower delay for the graphene wire. Although we demonstrate

47

this at the 16 nm node, the qualitative results are true regardless of the underlying CMOS technology because the transistor resistance begins to increase exponentially when the supply voltage begins to drop below the threshold voltage. In Section 6.3.3, we experimentally demonstrate that graphene wires outperform similarly sized metal wires in subthreshold operation.

2.4 Summary

In this chapter, we use a physics-based circuit model to compare the performance of graphene and Cu wires. Rough edges can backscatter electrons and lower the effective mobility or mean free path. Controlling the edge of a graphene device is even more important since a rough edge occurs even when a single atom is displaced on the edge of a graphene wire. When line-edge roughness is accounted for, we expect that graphene will begin to outperform Cu in terms of resistance around 10 nm line widths. However, this requires very thick graphene films or graphene wires with very smooth edges. Alternatively, we can take advantage of the small capacitance of graphene wires. Roughly 10 or fewer number of graphene layers are needed to have 2x smaller capacitance than Cu wires. Although this results in large wire resistances, this is advantageous in subthreshold circuits where the large transistor resistance dominates. Ultra-thin graphene wires have lower delay and energy dissipation under such conditions than similarly sized Cu wires.

Chapter 3

Fabricating Graphene Devices

Graphene has shown promise for a large range of applications. Since the introduction of graphene, many research groups have explored various ways to fabricate graphene devices. This chapter briefly surveys the existing synthesis methods and then outlines the process flow for producing graphene wires. The integration process with CMOS is also discussed here. Being able to leverage the established infrastructure of CMOS foundries is extremely important as graphene-based electronics gain traction.

3.1 Survey of Existing Methods

The various arrangements of pristine carbon atoms lend itself to many well known compounds, such as diamond, graphite, or even soot. Graphite has a layered, planar structure. In each layer, the carbon atoms are arranged in a hexagonal lattice. The in-plane bonding forces are much stronger than that in the perpendicular direction. For example, graphite is commonly used as the marking material in pencils. This weak interlayer bonding force allows the graphite source to be mechanically cleaved, thus allowing a pencil to easily leave a mark on paper.

The term graphene strictly refers to a single layer sheet of graphite [79, 80]. However, throughout literature the term graphene is often loosely used to indicate monolayer, few-layer, and even thick multilayer stacks of graphene sheets.

The widespread interest in the many possible applications of graphene and the

49

search for large-scale and low-cost manufacturing methods has led to the discovery of many different fabrication techniques. Traditionally, copper or aluminum interconnects are fabricated by evaporating a bulk metal source and physically depositing the material onto a substrate. Because graphene is one atom thick and represents the ultimate limit of a two-dimensional material, many research groups have attempted to produce large sheets by chemically synthesizing graphene. This bottom-up approach presents a stark departure from the conventional approach used in modern CMOS processes. The ability to manufacture high quality metallic sheets is perhaps one of the most important advantages of graphene over Cu at such thicknesses.

Monolayer graphene has existed for a long time. However, the first isolation of high quality monolayer graphene did not occur until 2004, which was achieved by micromechanical cleaving of a highly-oriented pyroltic graphite (HOPG) source [64]. To date, HOPG graphene still remains as one of the most effective ways to obtain high-quality devices [28, 34]. This method is assisted by an adhesive tape, which repeatedly peals off graphite layers from HOPG until monolayer or few-layer regions are obtained [64]. The graphene flakes are then transferred to a target substrate, typically a 300 nm SiO_2/Si substrate, by rubbing the adhesive tape.

Mechanical exfoliation of HOPG remains one of the most popular methods due its relative ease and high sample quality. However, this method is not suitable for large-scale integration. This technique has poor area coverage and produces small flakes on the order of tens of μm at best.

Chemical vapor deposition (CVD) was one of the first methods that reliably grew large areas of graphene sheets [36, 37]. CVD is attractive for large-scale integration as it enables an arbitrarily large area of graphene to be grown [36, 37, 41], which can then be subsequently patterned using standard lithography methods to produce narrow graphene wires. The growth process starts by placing a catalyst film, typically Ni or Cu, in a thermal furnace. At high temperatures, a carbon-rich gas source fills the chamber and the carbon atoms begin to diffuse into the catalyst film. As the temperature decreases, the excess carbon begins to precipitate out forming layers of graphene. The CVD process is not limited to small substrates as Bae et al. have

recently demonstrated roll-to-roll production of 30-inch graphene films [55].

Graphene sheets produced by CVD are usually transferred to another substrate after the growth process. Both Si and flexible plastic substrates are commonly used as target substrates. The main advantage of this aspect is that the high-temperature CVD process is isolated and independent from the target substrate. However, the additional transfer step also adds to the manufacturing complexity and introduces defects (Section 3.3).

Another method of growing epitaxial graphene has been facilitated by thermal decomposition of single-crystal silicon carbide wafers [81]. The SiC substrate is heated in vacuum to temperatures above 1250°C [81]. The thickness of the graphene film is determined by the temperature and the graphene layers are electronically decoupled [59, 81]. However, SiC wafers are more expensive than Si wafers and most graphene applications are processed on Si or plastic substrates.

Rather than synthesizing graphene by carbon precipitation, others research groups have taken a different route to chemically derive graphene [51, 82]. Li et al. have produced ultra smooth graphene ribbons by suspending graphite in solvents with polymer functionalization [51]. Although their approach yields narrow ribbons with well defined edge structures, controlling the device yield and placement is extremely difficult. In contrast, Yamaguchi et al. demonstrated uniform graphene films on 300 mm SiO_2/Si wafers by spin coating graphene oxide in a methanol/water solution [51]. The film thickness is uniform and controllable by adjusting the volume of the solution and spin rate. Nonetheless, because the film is made up of overlapping flakes, the mobility is an order of magnitude lower than films produced by other methods.

Table 3.1 summarizes the performance of graphene sheets grown by various methods. The mobility of HOPG graphene is unmatched. Murali et al. have also demonstrated HOPG graphene devices that have achieved the best experimental resistivity values and are comparable to that of Cu [34]. Because HOPG graphene devices are typically isolated, the sheet resistivity or number of layers is not usually reported. The quality of CVD graphene films vary quite a bit, but the lowest reported sheet resistance of graphene films is obtained by roll-to-roll CVD [55].

Table 3.1: Summary of graphene performance from various fabrication methods.

Method	Mobilities ($cm^2V^{-1}s^{-1}$)	Sheet Resistance (Ω/sq)
Ni CVD [37]	100-2000	770-1000
Ni CVD [36]	3750	280-600
Ni/Cu CVD [41]	500-3000	280-510
Ni CVD ($AuCl_3$ doped) [83]	-	150
Cu CVD [42]	4050	-
Cu CVD (HNO_3 doped) [55]	4000 - 12,000	30-125
Chemically derived [82]	10 - 15	600 - 2400
SiC [84]	-	170-280
SiC (Temperature=4K) [52]	2700	-
HOPG (suspended) [25]	200,000	-
HOPG [65]	2000 - 20,000	-
HOPG [66]	10,000 - 25,000	-

While many synthesis methods exist, CVD is exclusively used throughout this thesis. A key objective of this thesis is to examine the integration of graphene with CMOS technology and CVD enables large-scale integration on arbitrarily large areas. Although the CVD and transfer methods has certain limitations for large-volume manufacturing, these methods present the best known solution and are quite suitable for developing a prototype graphene/CMOS solution. The limitations of CVD are further discussed in Section 3.3.3.

3.2   Fabrication of Graphene Wires

3.2.1   Synthesis of Graphene Sheets

In this work, we grow large graphene sheets by CVD [36,37,41,42]. Two methods are commonly used in literature. Both methods have limitations that must be overcome. The first method uses Ni films and a fast cooling rate to produce thick, multilayered graphene sheets [85]. Although the electrical properties need improvement, multilayer

graphene films produced from Ni-based CVD are more suitable for interconnect applications. Unlike transistor applications, a thicker graphene layer is generally desired as it increases performance [86] and manufacturing reliability.

The second method uses Cu foils instead of Ni. Recently, Cu foils have been used in CVD to produce mono- or bi-layer graphene films [41, 42]. The Cu-grown graphene films have shown superior performance than Ni-grown films due to their high uniformity (Table 3.1). Because Cu CVD produces monolayer graphene films, repeating the transfer process is often required to obtain thicker stacks and better performance [55].

We use both methods in this work. Chapter 4 and Chapter 5 focus on Ni-grown graphene wires. Devices made from Cu CVD have been widely studied. But interconnect properties of Ni-grown multilayer thick graphene have not yet been reported. In Chapter 6, we use Cu-grown graphene to demonstrate complete CMOS system.

Next, we describe the process flow for Ni CVD (Figure 3-1). The starting material is a 500 nm Ni film on SiO_2/Si substrate. First, the substrate is placed in a quartz tube. The quartz tube is placed in a thermal furnace. The substrate is then annealed at 900°C for 20 minutes. Then the growth is carried out at 1000°C for 5 minutes, using 5-30 sccm of methane and 1300 sccm of hydrogen [37]. For Ni CVD, the cooling rate has an important impact on how carbon precipitates out of the Ni film [37]. This results in a large-area multi-layered graphene sheet directly on top of the Ni film.

The CVD process is fundamentally identical for both Ni and Cu catalysts except that Cu replaces the Ni film. For the Cu CVD process, a free standing Cu foil is generally used instead of depositing Cu on a Si substrate. The other difference is that the Ni process is done at atmospheric pressure and Cu process is done in the tens of millitorr range.

3.2.2 Substrate Transfer and Graphene Patterning

Once the growth process is complete, the graphene sheet can be transferred to any substrate. Figure 3-2 shows this process. A layer of poly(methyl methacrylate) (PMMA) is spun on top of the graphene/Ni or graphene/Cu substrate. A wet etchant
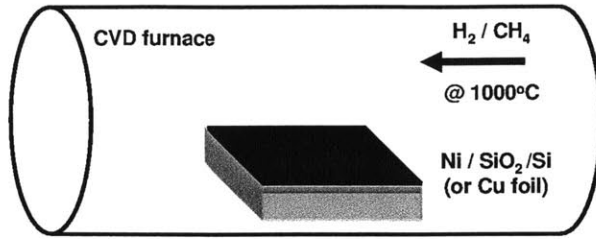
Figure 3-1: Diagram of CVD process. Substrate is placed in a thermal furnace where a mixture of hydrogen and methane flow at 1000°C.

is used to detach the PMMA/graphene layer from the underlying substrate. Here, we use a 10% HCl aqueous solution for the Ni process and a Cu etchant and HCl for the Cu process. This PMMA/graphene film can then be placed on an arbitrary target substrate, where after the PMMA layer is rinsed off with acetone [37].
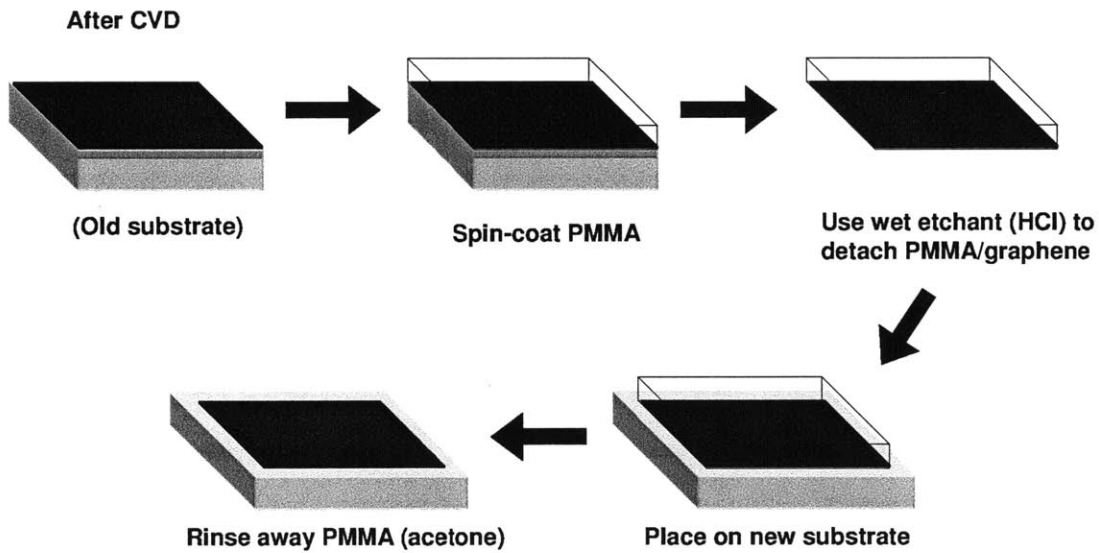
Figure 3-2: Process flow for transferring graphene film on to an arbitrary target substrate.

For off-chip measurements in Chapter 4, the graphene sheet is transferred to a blank 300 nm SiO_2/Si substrate [37]. When integrating with CMOS, the graphene sheet is directly transferred to the CMOS test chip (Section 3.3).

Once the graphene film is on the target substrate, the graphene sheet is patterned into graphene wires. Figure 3-3 shows the process flow for fabricating graphene wires on a blank substrate. First, we use a negative resist in an electron-beam lithography

process to define the wire patterns. Then oxygen or argon plasma is used to etch away the unwanted graphene leaving the wires patterns intact. The etch time needs to be monitored since the oxygen plasma also etches the electron-beam resist. The graphene wires are generally less than 20 nm. Thus, under nominal conditions, a short etch time less than 30 seconds is usually sufficient to fully etch the unwanted graphene regions. However, very thick regions often form on Ni CVD graphene films and these spots do not get removed even after prolonged plasma etch steps. Other limitations are discussed in Section 3.3.3. Finally, a second lithography step is used to define the metal contact region. The process is complete after depositing a Ti/Au or Cr/Au films and lifting off the remaining metal films.
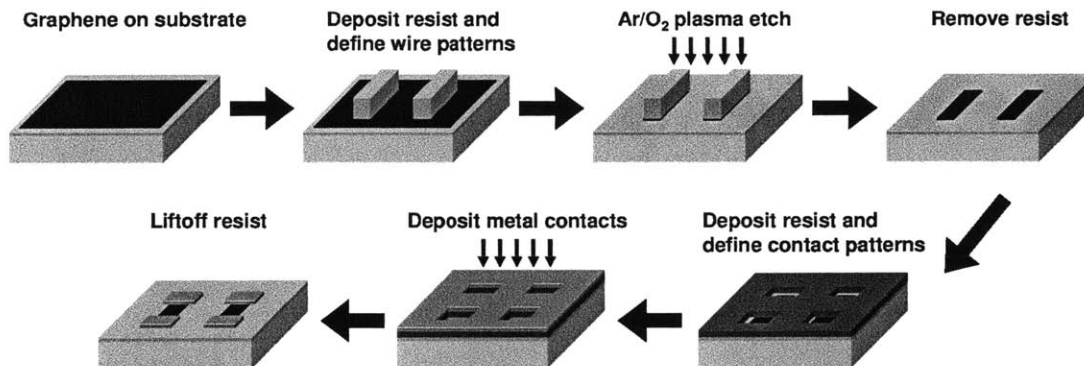
Figure 3-3: Process flow for fabricating graphene wires.

3.3 Monolithic Integration with CMOS

The previous section explained the fabrication process of graphene wires on a blank substrate. Here, we describe the process flow for integrating graphene on CMOS chips. Although both test chips follow a similar process flow, each test chip was optimized for different functions. The implementation details of the two CMOS prototype chips are described in Chapter 5 and Chapter 6.

Before graphene was discovered and gained widespread interest, carbon nanotubes were considered a promising material for many electronic applications [87–89]. Since its initial discovery, a number of successful commercial applications have emerged [90–

93]. However, most of these applications only leverage the extreme lightweight and strength of carbon nanotubes or its chemical sensing properties. It is extremely difficult if not impossible to precisely control the placement and type of carbon nanotubes at the level that is required for integrated circuits. A number of research projects have successfully integrated carbon nanotubes with CMOS [94] but the device yield is typically very low.

Graphene on the other hand retains many of the properties that made carbon nanotubes widely popular and attractive. In addition, the planar nature of graphene allows it to be processed using conventional lithography steps. This feature makes graphene more easy to integrate with CMOS and opens new opportunities for many electronic applications [95,96].

3.3.1   A 0.35 μm CMOS Prototype

The first test chip was implemented in a standard 0.35μm CMOS process. As the integrated graphene/CMOS chip serves primarily as a technology demonstrator, a more advanced CMOS technology was not needed and not used here. The test chip implements a low-swing driver and receiver to characterize the performance of long graphene wires. We use Ni CVD to produce thick multilayer graphene sheets. All remaining integration steps are CMOS compatible and follow a similar flow as in [38].

First, we expose the CMOS chip in a CF_4-based plasma etch process to partially thin down the silicon nitride passivation layer. Electron-beam resists typically have very poor etch selectivity. Therefore, we perform this blanket exposure to thin down the passivation layer to make it easier for later process steps. Next, we use the same process outlined in Section 3.2.2 to transfer the Ni-grown graphene sheets to the CMOS chip. Fig. 3-4 shows both a drawing and an optical image of the CMOS chip at one end of the graphene wire.

After the graphene sheet is transferred to the CMOS chip, a lithography step then defines the wire patterns. The unwanted graphene regions are exposed and etched away using oxygen plasma. A second lithography step defines the top metal contacts. After Ti/Pt contacts are deposited, a third lithography step defines the via openings.
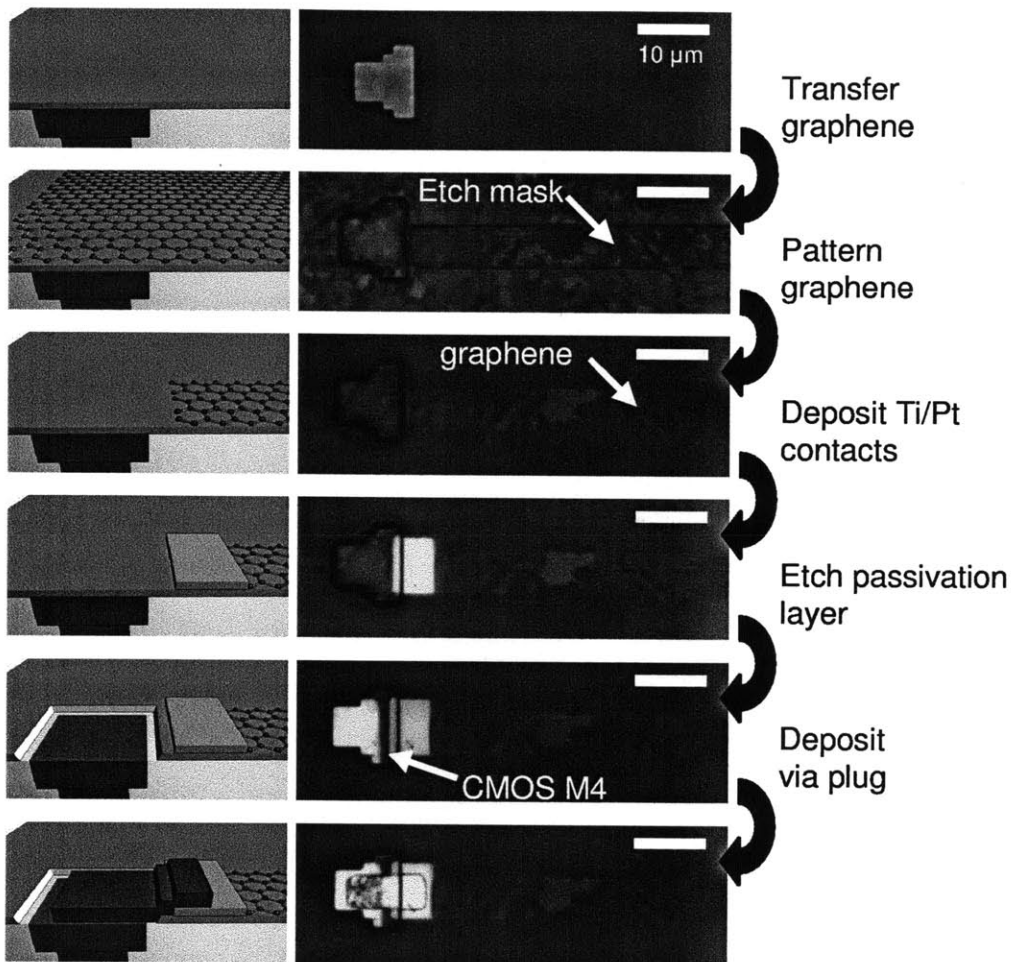
Figure 3-4: Process flow for graphene and CMOS integration. A drawing and optical image at one end of the graphene wire is shown.

A CF_4/Ar etch process etches through the passivation layer and exposes the top metal layer of the CMOS chip. A fourth lithography step defines the final via plug patterns. The chip is placed in a sputterer and a brief Ar plasma is used to clean the native oxide formed on the exposed Al wires. Lastly, without breaking vacuum, Ti is sputtered to make the via plugs between the graphene contacts and the CMOS metal layer.

In this test chip, the graphene wires are 4 μm wide and 0.5 or 1 mm long. The total yield of working devices was about 60 % for 0.5 mm wires, and about 3 % for 1 mm wires. Unlike test devices fabricated on flat SiO_2/Si wafers, the device yield for the longer wires has been affected by the topology of the CMOS passivation layer and the increased number of processing steps. Graphene wires that did not function were mainly a result of lithography alignment errors or tears/voids along the graphene channel. After each step, we visually inspected the chip and observed that graphene tore apart mostly during lithography/etch steps. Handling small CMOS chips also contributed to lithography challenges.

3.3.2   A 0.18 μm CMOS Prototype

The second test chip implements a field-programmable gate array (FPGA) in a 0.18 μm CMOS process. The underlying circuit was designed with missing wires, onto which graphene wires were subsequently integrated. The second test chip was designed for ultra-low power operation and thus we decided to use thin few-layer graphene sheets to obtain lower wire capacitance. We follow a similar flow as in the first test chip, but the order of the steps has changed.

Fig. 3-5 shows the process flow for integration. We first perform a blanket exposure on the CMOS test chip to thin down the silicon nitride passivation layer. The passivation layer did not etch uniformly across the chip. The relevant via patterns are placed at the center of the chip but the passivation layer in these regions is etched the least. This non-uniform etching is likely due to the dense patterns and different nitride composition in this test chip. Therefore, unlike the first test chip, we then lithographically define the via holes. A CF_4/Ar plasma step etches through the pas-

sivation layer and exposes the top CMOS metal layer. This allows us to over-etch the via holes while protecting the perimeter of the chip.
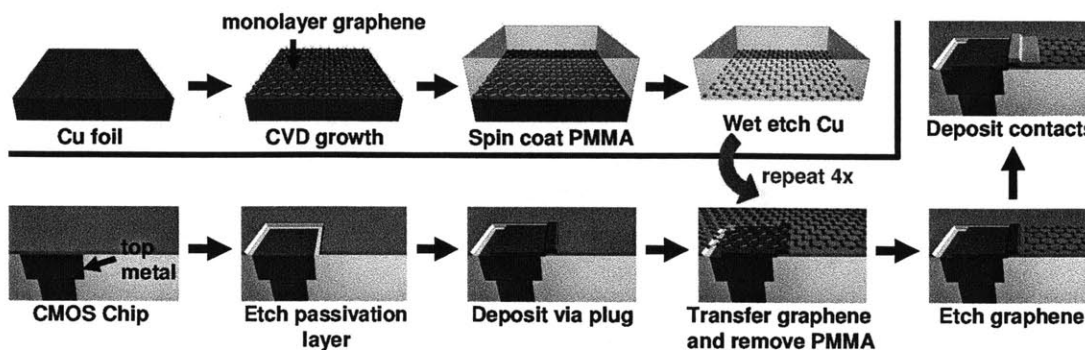
Figure 3-5: Process flow for synthesizing/transferring graphene sheets and integrating graphene wires with the CMOS test chip.

After forming the via holes, a second lithography step defines the via plug patterns. Ti/Au is then deposited to fill the via holes. Next, we use the Cu CVD process to grow the monolayer graphene sheets. The growth is carried out at 1000°C and 10 mT, with 20 sccm and 10 sccm flow of CH_4 and H_2, respectively. Using 4 separate graphene sheets, we repeat the transfer process 4 times to achieve a thicker graphene stack and lower sheet resistivity. For clarity, the transfer process is also depicted in Figure 3-5. The resulting sheet resistance from the 4-layer stack is roughly ~270 Ω/sq.

After all 4 layers have been transferred, the graphene sheet is patterned into wires using electron-beam lithography and an Ar/O_2 plasma etch step. Finally, Ti/Au contacts are deposited on top of the wires/vias. Although the order is different, the same four lithography steps are performed on this chip as on the first test chip. Figure 3-6 shows images of the the integrated graphene wires. Roughly 40 % of the integrated graphene wires worked. Visual inspection of the chip shows that the device yield is largely impacted by the transfer process. Parts of the graphene sheet get wrinkled or tear apart, resulting in either holes or non-uniform multilayer regions. We discuss this further in the following section.
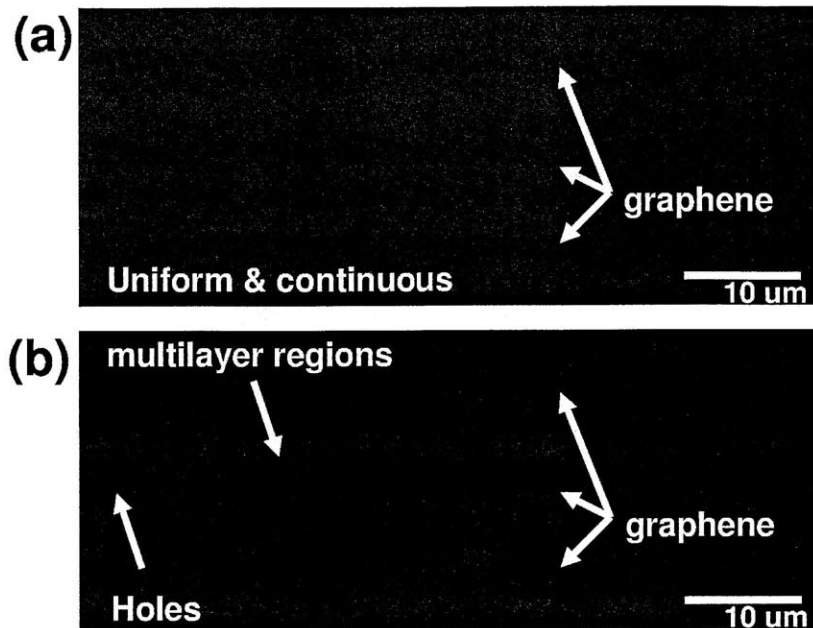
Figure 3-6: Image of graphene wires on top of the CMOS chip. We observe areas of (a) uniform 4-layer wires, as well as (b) non-uniform and non-continuous regions due to wrinkling and tearing that occurs during the transfer process.

3.3.3 Optimizing Process Flow

In the previous sections, we have described the process flow for integrating graphene with CMOS. The planar nature of graphene allows conventional lithography methods to be used. This integration process is fully CMOS compatible but also accentuates several limitations that must be overcome.

Several groups have demonstrated wafer-scale fabrication of graphene devices [41, 82, 97, 98]. In their work, graphene devices are fabricated on flat substrates and only require one or two lithography steps. The actual device yield is seldom reported. In our work, the graphene wires are narrow and up to 1 mm in length and involve more processing steps. The device yield is significantly impacted by the topology of the CMOS passivation layer and the increased number of processing steps. For example, the longest graphene wires (1 mm) had the lowest device yield of only 3 %. Visual inspection of the chip after each step revealed that graphene tore apart during various lithography/etch steps. This affects the longer wires more than the shorter wires. The shorter wires (0.3 - 0.5 mm) generally had between 40 to 75 %

of the wires functioning properly. This may be attributed to the weak bonding force between graphene and the substrate. Although most wires appear robust and adhere well to the surface, the graphene wires are extremely delicate and are often knocked out of their place during resist liftoff or sonication.

One aspect that contributed to these challenges was the small chip size. In this work, both test chips were only 5 mm x 5 mm. Although this is generally considered very large for a typical integrated circuit, this is rather difficult to handle in a clean-room environment. When spin coating the electron-beam resist, the small chip size necessitates a higher spin speed to reduce the amount of electron-beam resist that builds up on the perimeter of the chip. The uneven thickness of the resist affects not only the lithography and development stage but also the subsequent etch and deposition steps. Nonetheless, these problems can be optimized by tweaking the process conditions or by processing an entire wafer rather than a single chip.

Perhaps a more difficult and important issue is transferring the graphene sheet to the CMOS chip. As shown in Figure 3-6, parts of the graphene sheet tear apart or get folded during the transfer process. This appears to be more problematic for Cu-grown graphene sheets which are only one atom thick. Ni-grown graphene sheets are generally between 10 and 20 nm and more rigid than Cu-grown graphene sheets. Achieving nearly 100 % coverage is absolutely critical for interconnect applications. Recently, Bae et al. have demonstrated production of 30-inch monolayer graphene sheets [55]. They use graphene for a touch-screen panel which has less stringent requirements. For example, a 10 μm long tear in the graphene sheet will likely not prevent the touch-screen panel from functioning but creates major issues for integrated circuits. The transfer process must be significantly improved to meet the manufacturing reliability needed for microelectronics. In this regard, Ni CVD appears more promising but must overcome the non-uniform growth of graphene layers.

3.4 Summary

In this chapter, we have described the process flow for fabricating graphene wires. We grow large graphene sheets by CVD using Ni and Cu catalysts. Ni CVD is performed in atmospheric pressure and produces thick multilayered graphene sheets. Unlike transistor applications, a thicker graphene layer is desired for interconnects as it decreases wire resistance. In contrast, Cu CVD produces monolayer graphene sheets. Because of the extremely high uniformity, Cu-grown graphene sheets generally have superior performance than Ni-grown graphene sheets. However, Cu CVD only produces a single sheet and thus multiple transfers are often required. The Cu-grown sheets are extremely delicate and seem to break more easily than the thick Ni-grown graphene sheets. Both methods are important for interconnect applications and we designed two CMOS test chips to characterize the performance of either Ni or Cu-grown graphene. Because CVD produces an arbitrarily large area of graphene, conventional lithography methods were used to monolithically integrate graphene wires with CMOS. Ultimately, since the graphene sheets are very delicate, the transfer process needs to be optimized to increase device yield and improve performance.

Chapter 4

Characterizing Multilayer Graphene Devices

CVD enables growth of large-area graphene sheets. Using Cu catalyst produces high quality monolayer graphene sheets but requires multiple transfers. Ni catalyst yields thick multilayer graphene sheets and is more suitable for interconnection applications. While the property of monolayer or few-layer graphene interconnects have been widely studied, few reports exist on characterization of multilayer graphene interconnects. This chapter explores the fabrication of thick multilayer graphene sheets and its properties.

4.1 Sheet Properties

The graphene sheets grown here are produced from a Ni catalyst using CVD in ambient pressure. Thick multilayer graphene sheets are grown by utilizing the fast-cooling method [85].One precursory way of increasing the amount of carbon precipitation is to increase the concentration of methane. Figure 4-1 shows optical images of the graphene sheet as the methane concentration is increased. Generally, the number of thick layers increases with the methane concentration.

The number of graphene layers can be readily identified from the optical images [85, 99–102]. Even a single layer of graphene creates an amplitude shift of the
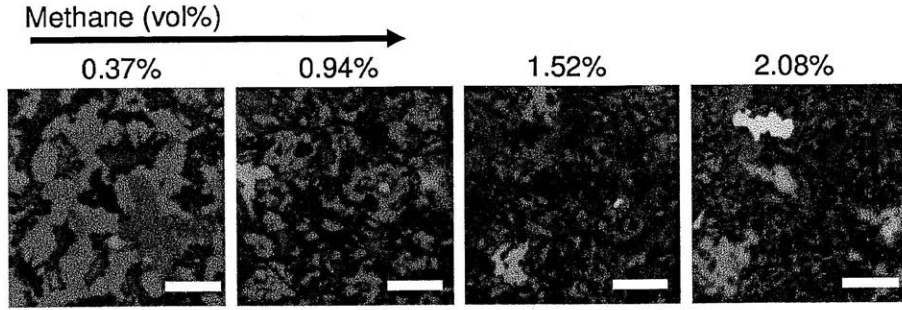
Figure 4-1: Optical image of graphene sheets on Si substrate with 300 nm SiO_2 (SiO_2/Si). Dark colored regions represent thicker graphene layers. Scalebars are 10μm.

light that is reflected from the surface. This results in a color contrast that depends on the number of graphene layers. A digital image has three components, red, green, and blue. Graphene on 300 nm SiO_2/Si creates an enhanced absorption at wavelengths around 500 nm [103, 104], which corresponds to the green component in the RGB color space. Using a similar method found in [85], we can estimate how many graphene layers exist at each pixel.

From Figure 4-1, we observe that the percentage of thicker layers increases with methane concentration.In Figure 4-5, we quantify the estimated area coverage of 1-3 layer graphene regions. The area coverage of 1-3 graphene layers decreased as methane concentration increased up to roughly CH_4=0.6 %(vol). For larger methane concentrations the area coverage stayed relatively constant around 10 %. This saturation point could indicate that the solubility limit of carbon diffusion into the Ni film during CVD growth.

In other words, at low methane concentrations, the graphene film has predominantly 1-3 layers with small regions of very thick graphene layers. At higher methane concentrations, the film is generally much thicker with a more even distribution of thin and thick regions. Figure 4-3 also supports this and shows that the average surface roughness R_a decreases as methane concentration increases.

Figure 4-4 shows the measured sheet resistance as a function of R_a obtained from AFM measurements. Sheet resistance is extracted from both the large sheet, using a 4-point probe method, and from each batch of graphene wires. Measured results
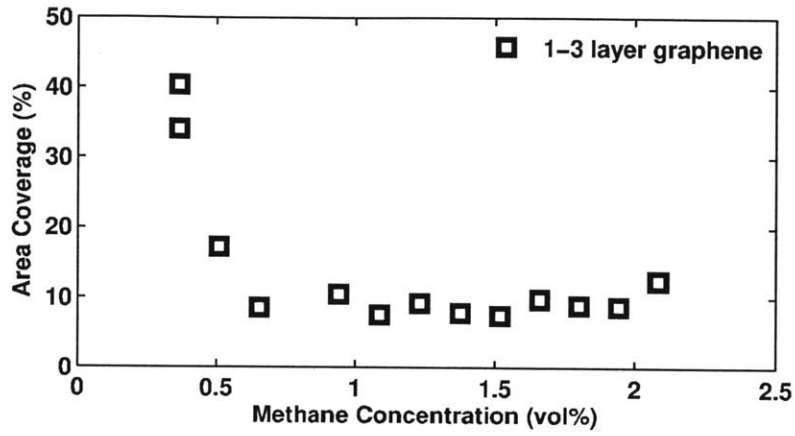
Figure 4-2: Estimated area coverage of 1-3 graphene layer regions.
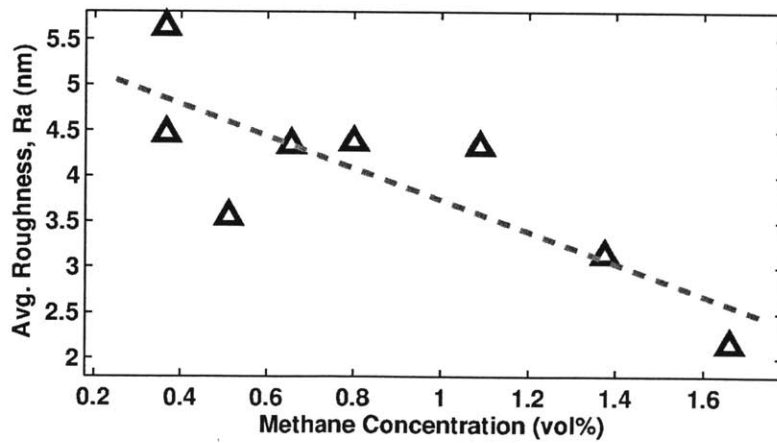
Figure 4-3: Average roughness of graphene sheets.

of the graphene wires are explained in the following section. First, 4-point probe measurements are conducted on the large graphene sheets. A second method plots the resistance of each fabricated wire against the length/width ratio. The slope of this curve also yields the effective graphene sheet resistance.
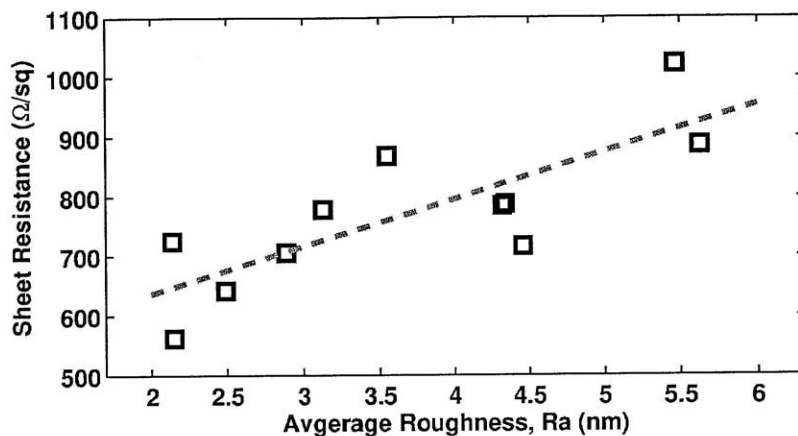
Figure 4-4: Measured sheet resistance.

The sheet resistance shows a strong correlation (r=0.79) with R_a, which is partly related to the methane concentration and cooling rate during the CVD process. This plot indicates that achieving good uniformity is critical in obtaining lower sheet resistance. Generally, high-quality devices made from HOPG have uniform graphene flakes and the best devices have resistivities that are comparable to Cu [34]. This result is also consistent with recent work, where sheet resistance as low as 30 Ω/sq was obtained from a highly uniform p-doped 4-layer graphene sheet grown by Cu CVD [55]. However, such methods require performing multiple transfers of monolayer graphene sheets. The devices here have higher sheet resistance (500-1000 Ω/sq) than recent reports of Cu CVD (30-280 Ω/sq) [55] and high temperature decomposition of SiC (170-280 Ω/sq) [84]. Nonetheless, directly growing multilayer sheets significantly increases manufacturing throughput.

4.2 Wire Properties

For off-chip electrical characterization, the graphene film is placed on a blank SiO_2/Si substrate. Figure 4-5 shows the height distribution measured from a graphene sample. The average height of the Ni-grown graphene sheets ranges from 10 to 20 nm. Once the graphene sheet is transferred to a SiO/Si substrate, the wires are defined lithographically using the procedure outlined in Section 3.2.
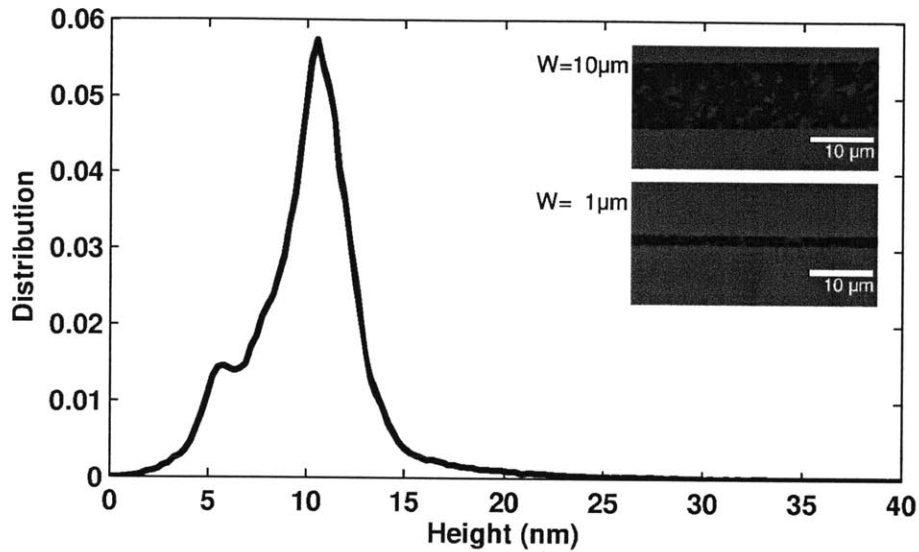
Figure 4-5: Height distribution of graphene (CH_4=1.7%) from AFM measurements. Inset shows optical images of fabricated graphene wires.

The 4-point probe measurements yield an approximate sheet resistance between 500-1000Ω/sq. We have fabricated graphene wires with widths of 1 and 10 μm and lengths from 2-1000 μm. Either Ti/Au, Cr/Au, or Ti/Pt metal pads have been deposited to make contacts for probing. The choice of metal does not affect the measurements. A set of I-V measurements confirm that graphene wires have near-ohmic contact.

Figure 4-6 shows the resistance of graphene wires as a function of length/width ratio. The wire resistance scale linearly with length, indicating diffusive transport [38]. For this particular batch of devices, the extracted sheet resistance from the slope was roughly 790 Ω/sq, which is close to the value of 740 Ω/sq obtained from the 4-point probe measurements.
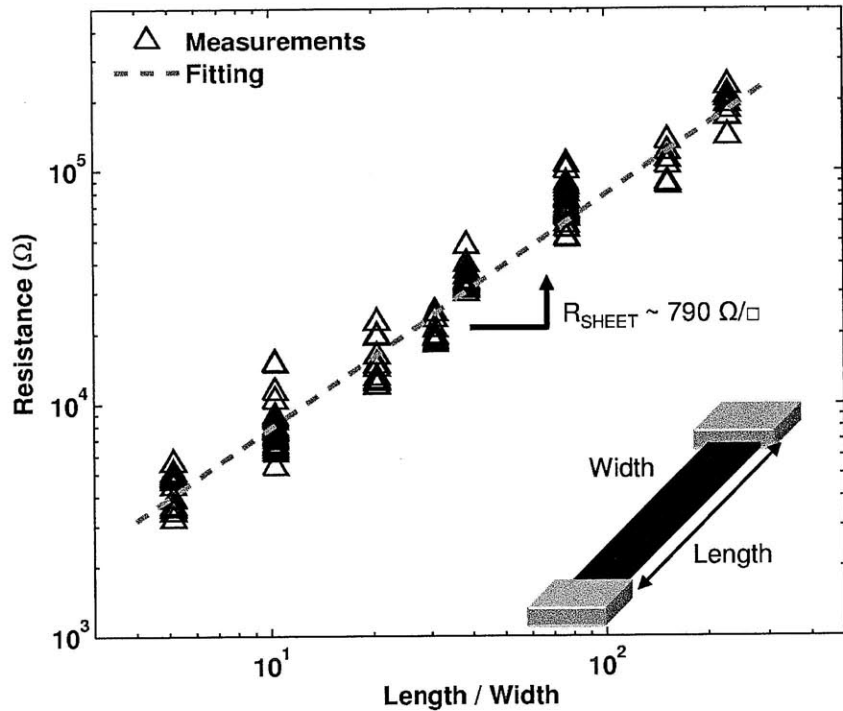
Figure 4-6: Resistance of graphene wires fabricated on SiO_2/Si test wafers.

4.3 Breakdown Current Density

The high mobility, ballistic transport, high current carrying capacity, and high thermal conductivity are all important factors that show promise for use of graphene as the interconnect channel material [23–28]. In particular, a high current carrying capacity is important for reliability and meeting the power density requirements. Recently, HOPG graphene has been reported to have current densities on the order of 10^8 A/cm^2 [28]. We have also characterized the current capacity of Ni-grown multilayer graphene.

Figure 4-7 shows the I-V curve of a graphene wire that is undergoing electrical breakdown. All measurements were taken in ambient air. The breakage occurs in the middle of the wire, indicating resistive heating. Unlike [28], the current does not saturate prior to breakdown. The graphene wires fabricated in this work are much larger and induce less self-heating. The maximum current density (J_{MAX}) and resistivity (ρ) are calculated using the width and average height of the wire.

68
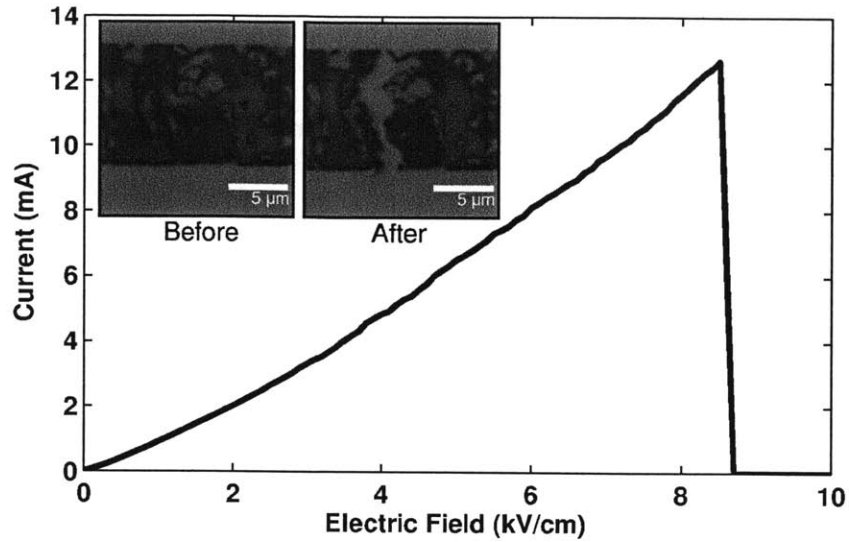
Figure 4-7: Measured I-V curve of a 50 μm long graphene wire undergoing electrical breakdown. Insets show an optical image of the graphene wire before and after the measurement.

Similarly, Figure 4-8 shows the I-V curve of a shorter (L=10 μm) graphene wire undergoing electrical breakdown. Unlike the longer (L=50 μm) wire shown in Figure 4-7, the slope of I-V curve increases possibly suggesting thermal annealing or increased electrical contact to the underlying graphene layers. In addition, the breakage of the shorter wire occurs over a much larger area and suggests that both the thin and thick graphene flakes were conducting current.

Figure 4-9 plots J_{MAX} and includes devices made from different CVD conditions and wire widths, but these do not have a noticeable effect. The plot shows that, at the same current density, high resistivity wires have a lower J_{MAX} and are more likely to breakdown. Defects and impurities increase the effective resistivity and also contribute to wire failure. Most of the data points are in the range from 2×10^6 to 4×10^7 A/cm^2. A power law relation in the form of $J_{MAX} \propto 1/\sqrt{\rho}$ has been previously proposed [28, 105]. In this work, a fit to $J_{MAX} = A\rho^{-n}$ yields an exponent of $n=0.86$ (L=2 μm), 0.94 (L=10 μm), and 0.79 (L=50-200 μm).

These results suggest instead that $J_{MAX} \propto 1/\rho$, which is consistent with having a constant breakdown electric field. Figure 4-10 shows the measured electrical field, which is roughly constant here for each length of devices (5-34 kV/cm). A

69
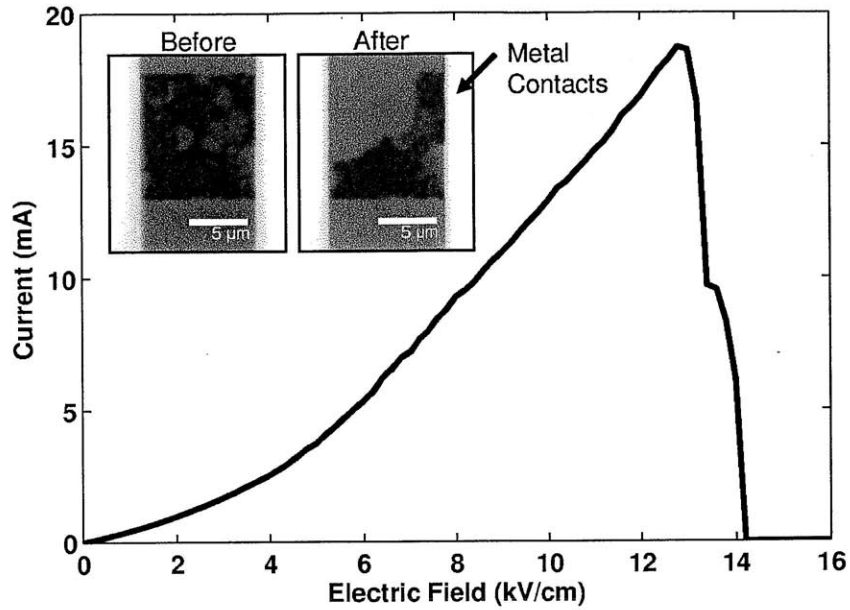
Figure 4-8: Measured I-V curve of a 10 μm long graphene wire undergoing electrical breakdown. Insets show an optical image of the graphene wire before and after the measurement.
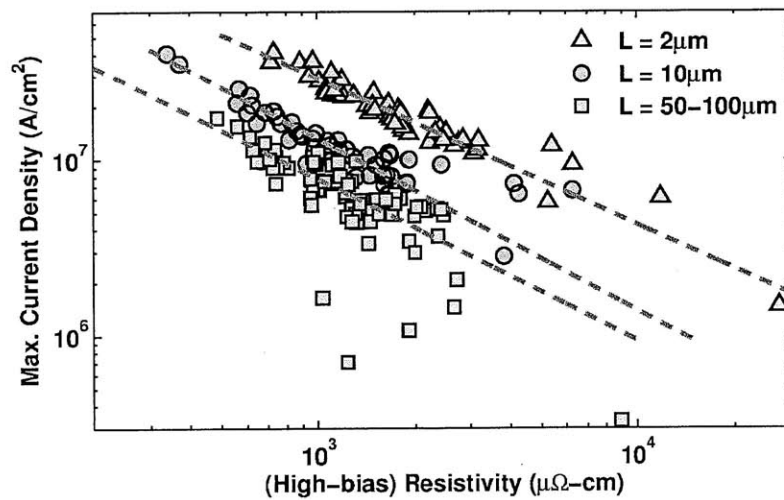
Figure 4-9: Measured maximum breakdown current density as a function of resistivity. The resistivity is defined at the point just prior to breakdown, as in [28, 105].

constant breakdown field has been previously observed for wall-by-wall breakdown of multiwalled carbon nanotubes [106] and graphene nanoribbons (33 kV/cm) [28].
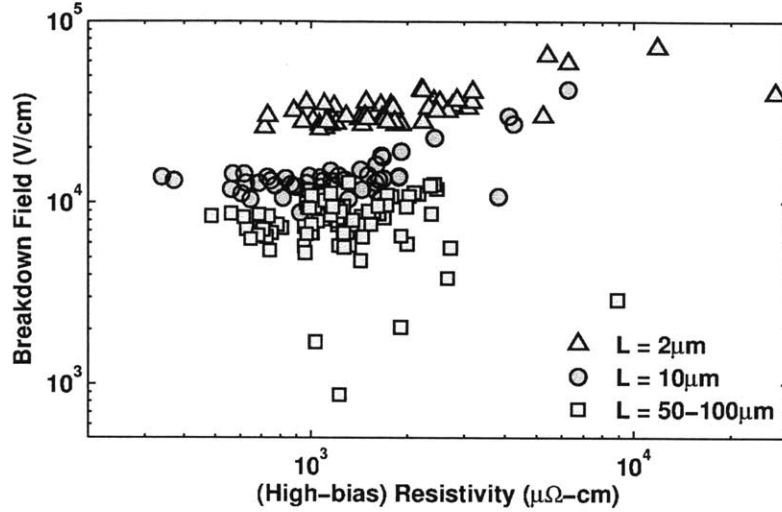
Figure 4-10: Measured electrical field at the point of current breakdown.

Furthermore, from Figure 4-9 and Figure 4-11, we see that J_{MAX} is generally higher for shorter wires. Although a similar length dependence has been observed in carbon nanofibers [107], the size of graphene flakes with uniform thickness appears to be a dominant factor in this work. The flake size with uniform thickness of CVD-grown graphene presented in this work is roughly ~5 μm. The graphene channel appears more uniform for shorter wires while longer wires show a large variation in the number of layers. The inset of Figure 4-7 also shows that the breakage tends to be centered on thinner layers, which may not necessarily be in the middle of the wire. Thus, increasing the uniformity of the multilayer graphene sheets is critical for both reducing sheet resistance and further increasing J_{MAX}.

Figure 4-12 shows the J_{MAX} of our data alongside the HOPG data found in [28]. The best HOPG devices approach the resistivity of Cu, and exceed the current capacity of Cu ($\sim 10^6$ A/cm^2) [108] by 100x [28, 29]. Due to layer non-uniformity, our CVD samples have higher resistivities than HOPG but nonetheless achieve J_{MAX} up to at least an order of magnitude higher than that of Cu. Although this work presents results for large wire widths, at the same resistivity, our samples exhibit similar J_{MAX}

71
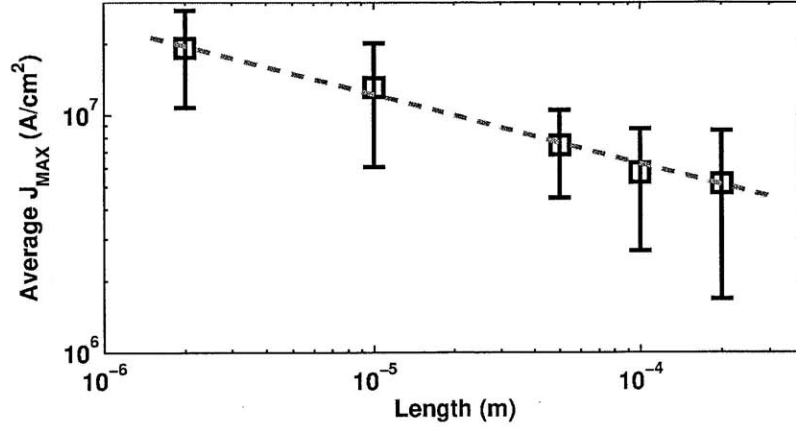
Figure 4-11: Measured average breakdown current density wire length. The error bars plot the standard deviation of J_{MAX} at each length.

values with sub-52nm HOPG samples.
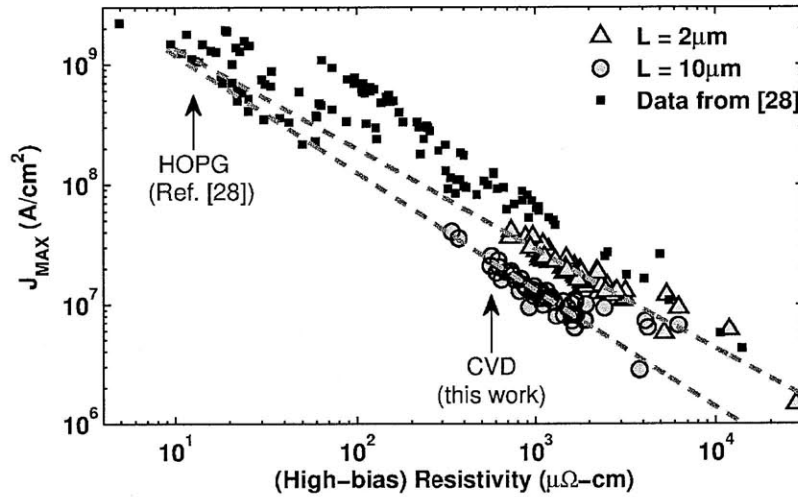
Figure 4-12: Measured maximum breakdown current density as a function of resistivity.

As the CVD process improves, we expect the quality and performance of the graphene to improve, which will be especially important at smaller dimensions where edge scattering effects are expected to be more pronounced. Reliably growing uniform multiple graphene layers remains a critical challenge. Although chemical doping can decrease the effective resistivity, the fundamental growth conditions need to be further

optimized. Recently, the use of Cu-Ni alloy [97], interface engineering [109], and optimizing the cooling rate [85,98] have proven to be effective methods in increasing the film quality and reliability.

4.4 Summary

In this chapter, we have characterized the properties of thick multilayer graphene wires, fabricated from large-area sheets grown by CVD. The CVD process results in an average thickness of 10-20 nm with a sheet resistance between 500-1000 Ω/sq. Maximum current densities up to $4x10^7$ A/cm^2 have been measured in ambient air. This exceeds the current capacity of Cu by an order of magnitude but falls short of the measured current capacity of HOPG graphene samples. The variation in the number of graphene layers directly affects the sheet resistance and current carrying capacity. Further process optimization is required to produce cleaner and more uniform graphene sheets, which can lead to lower resistance and higher current densities.

Chapter 5

Graphene Data Links

DC characterization of sub-50nm graphene interconnects has been previously reported [28, 34], but very few studies exist on evaluating their performance when integrated with CMOS. Off-chip measurements have limited scope and often require expensive equipment. Here, we characterize the performance of medium to long graphene wires integrated on a CMOS chip. The test chip implements an array of transmitters and receivers. Graphene sheets are grown by Ni CVD and then transferred to the test chip.

5.1   Overview

We present a CMOS test chip that characterizes the performance of integrated graphene wires. Chen et al. have recently reported the first integrated graphene/CMOS system [38], where CMOS ring oscillators are used to indirectly measure the performance of short graphene wires. In contrast, this work focuses on end-to-end data communication on medium to long graphene wires as shown in Figure 5-1. The performance of each graphene wire is measured in detail, using isolated transmitters and receivers. Wire lengths range from 0.5 mm to 1 mm for ease of delay measurements and for demonstrating reliable fabrication and signal transmission over longer length scales.
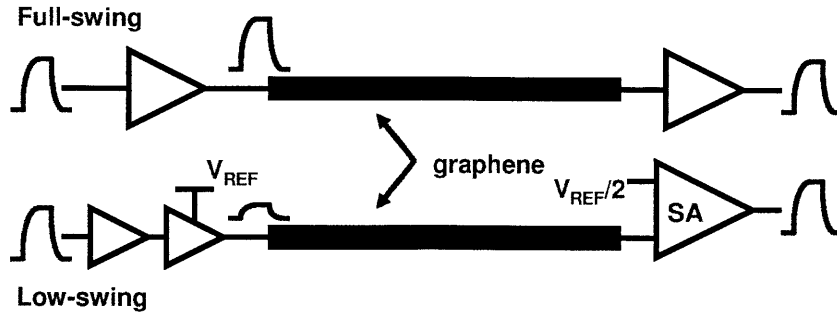
Figure 5-1: Overview of chip. Graphene wires are monolithically integrated and electrically connected to a transmitter and receiver. The low-swing topology uses a sense-amplifier (SA) at the receiver.

5.1.1 Implementation of Low-Swing Data Links

The CMOS chip includes an array of drivers/receivers to test signal transmission on integrated graphene wires. A full-swing and low-swing design is included. The full-swing topology uses a simple inverter chain to drive the signal and is unrepeated. The energy for signal transmission on a wire is roughly $C_w V_{sw}^2$, where C_w is the capacitance of the wire and V_{sw} is the voltage swing on the wire. Thus by reducing the voltage swing, the energy used to transmit a signal can be significantly reduced.

Figure 5-2a shows the diagram of the low-swing design. While many low-swing topologies exist [15, 17, 18, 110], this chip implements an NMOS push-pull driver and a secondary reference voltage source (V_{REF}) to generate a low-swing signal on the graphene wire. The receiver is a single-ended pseudo-differential latch-based sense amplifier [18].

The graphene wire was simulated using a $\pi 3$ distributed RC wire model. The distributed resistance (R_{GR}) was estimated from experimental values of the graphene sheet resistance. The capacitance (C_{GR}) of the wire was estimated from a field solver using the dimensions of the CMOS metal stack and the passivation layer. The parasitic capacitance due to integration (C_{INTEG}) was also estimated as 2-5fF.
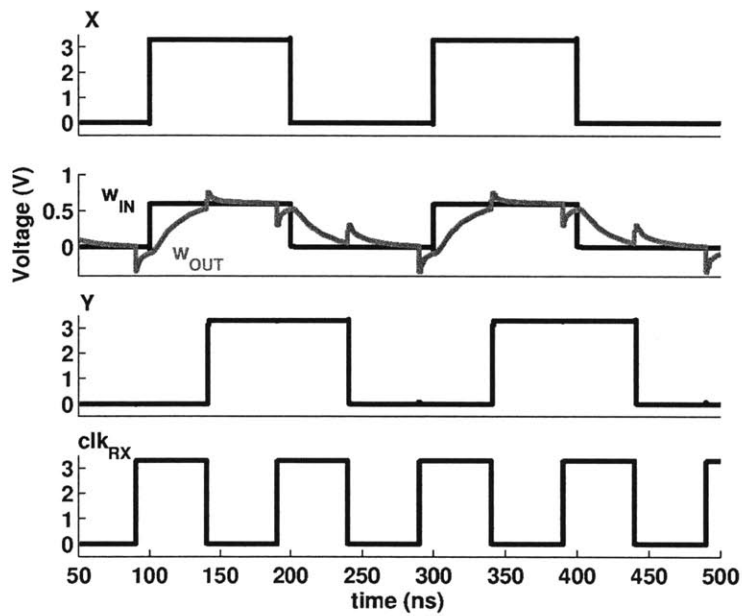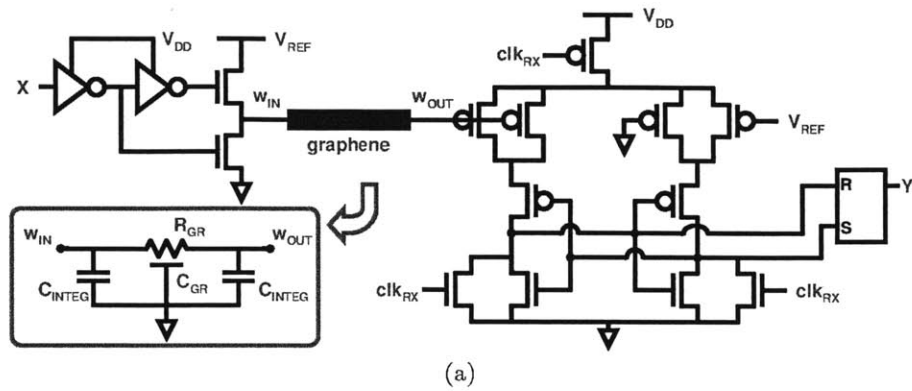
Figure 5-2: (a) Diagram of low-swing design. The graphene wire is simulated as a $\pi 3$ distributed RC wire. (b) Waveform of simulated low-swing design using V_{DD}=3.3V, V_{REF}=0.6V, R_{GR}=180kΩ, C_{GR}=200fF, and C_{INTEG}=3fF.

77

5.1.2 Experimental Setup

The measurement setup is shown in Figure 5-3 and Figure 5-4. The transmitter and receiver are independently clocked with a programmable delay. The balanced clock tree for the transmitter and receiver are matched to provide minimal skew. Furthermore, the worst-case delay of the transmitter and receivers are 0.59 ns, which is at least an order of magnitude smaller than the delay of the wire. The data input for each channel can either be an alternating 0101 pattern or a 2^{31}-1 pseudo-random binary sequence (PRBS).
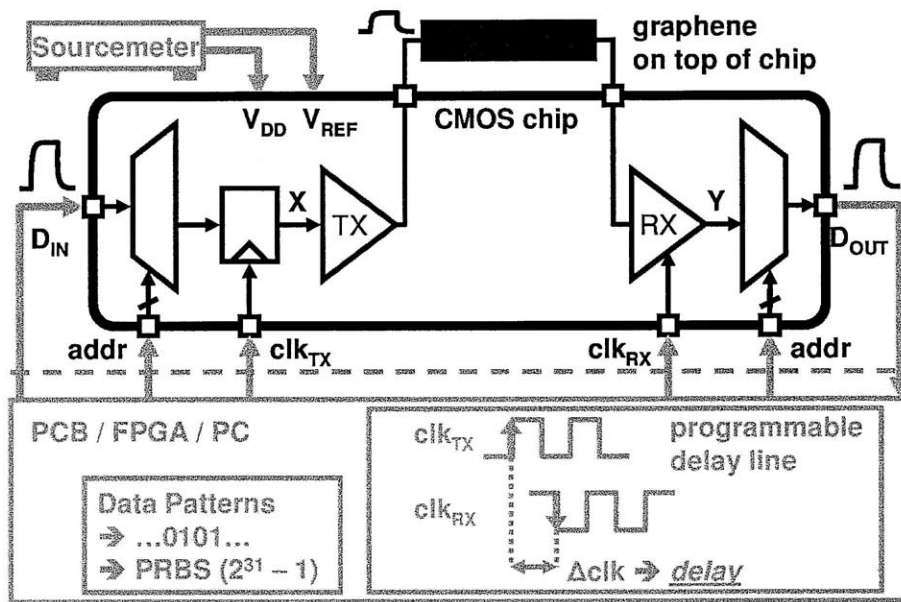
Figure 5-3: Diagram of experimental setup.
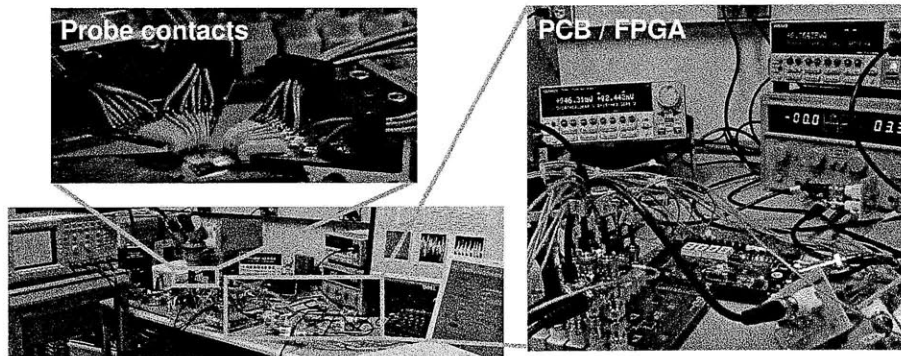
Figure 5-4: Photo of experimental setup.

5.2 Measured Results

The test chip was implemented in a standard 0.35 μm CMOS process. Since the test chip serves primarily as a technology demonstrator, a more advanced CMOS technology was not needed and not used here. The fabrication details were previously outlined in Section 3.3.1. Fig. 5-5 shows the CMOS die and the graphene wires. The graphene wires are integrated and electrically connected to the underlying CMOS drivers and receivers.
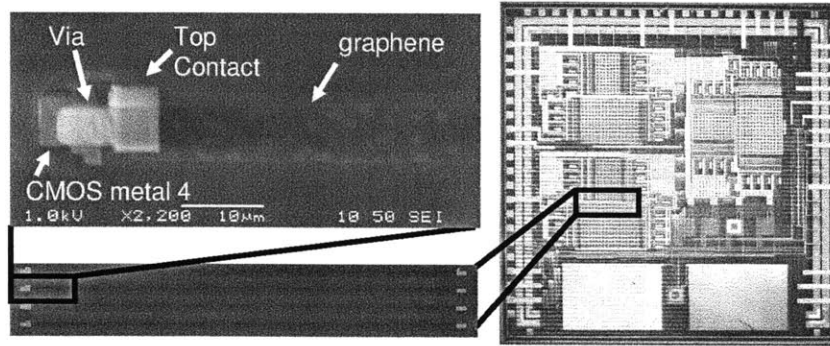
Figure 5-5: SEM image of CMOS die and integrated graphene wire.

The graphene wire widths are relatively large (4 μm) for ease of fabrication, but the results are expected to scale as line widths are reduced. The lengths of the wires are 0.5 and 1 mm. All measurements have used data rates between 1-50 Mbps. Figure 5-6 shows the transient waveforms for the low-swing design. A similar waveform is obtained for the full-swing topology. The data output (D_{OUT}) signal shows that it follows the input (D_{IN}) pattern and confirms connectivity of the graphene data channel.

The total yield of working devices was about 60% for 0.5 mm wires, and about 3% for 1 mm wires. Unlike test devices fabricated on flat SiO_2/Si wafers, the device yield for the longer wires has been affected by the topology of the CMOS passivation layer and the increased number of processing steps. However, similar to [38], the integration process did not appear to alter the graphene quality. The extracted R_{GR} values from measurements are close to the expected values. Graphene wires that did not function were mainly a result of lithography alignment errors or tears/voids along
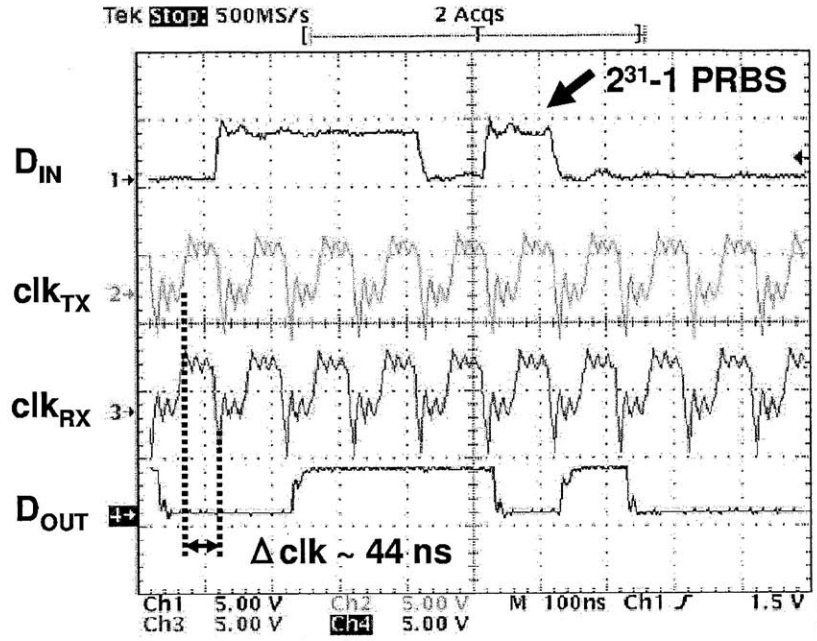
Figure 5-6: Transient waveform of digital signals for the low-swing design with a 2^{31}-1 PRBS input pattern (V_{DD}=3.3V, V_{REF}=0.5V).

the graphene channel.

Overall, the measurements were repeatable and did not show any signs of degradation over the period of a month. Figure 5-7 shows the measured bit error rates (BER) using a 2^{31}-1 PRBS. The transmitter/receiver clocks are fixed at a relative skew of roughly 12.8 ns. As V_{REF} decreases, the delay of the channel increases. A sufficient timing margin is guaranteed when the channel delay is much smaller than the preset clock skew, resulting in a very small BER. This timing margin becomes more difficult to meet as the channel delay approaches and exceeds the preset skew, which is reflected in the increase in BER. For nominal measurements, the graphene wires show BERs less than 2×10^{-10} at data rates between 10-25 Mbps. This shows reliable operation and that the properties of graphene wires do not drastically change under operating conditions.

For the low-swing topology, V_{REF} sets the voltage swing on the wire. Figure 5-8 plots the minimum voltage swing. Simulation results are in good agreement with measured results. The minimum V_{REF} is limited by the noise margin and sensitivity
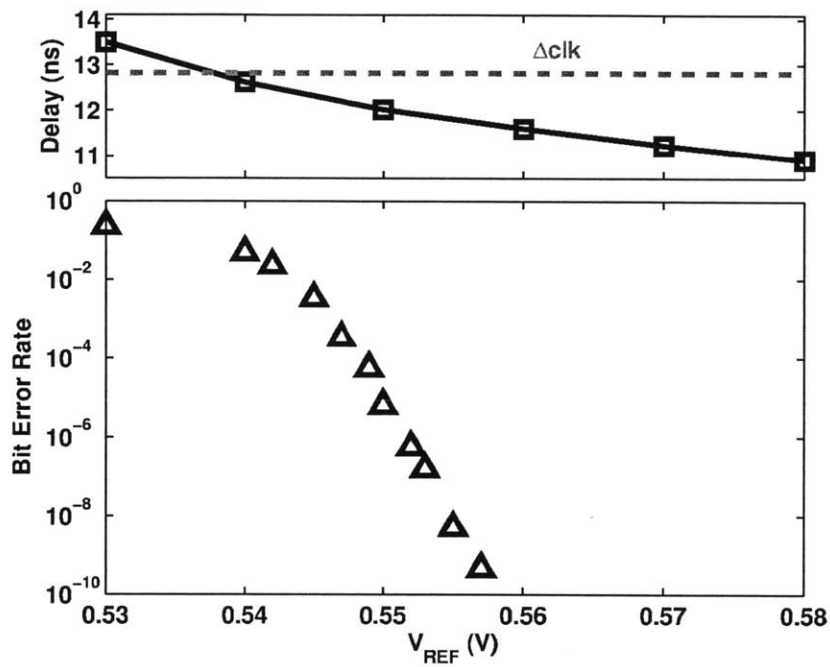
Figure 5-7: Measured delay (top) and bit error rates (bottom) of a low-swing graphene channel. The clock delay (Δclk) is fixed at 12.8 ns. The BER is measured using a 2^{31}-1 PRBS at V_{DD}=3.3V.

of the sense amplifier at the receiver.
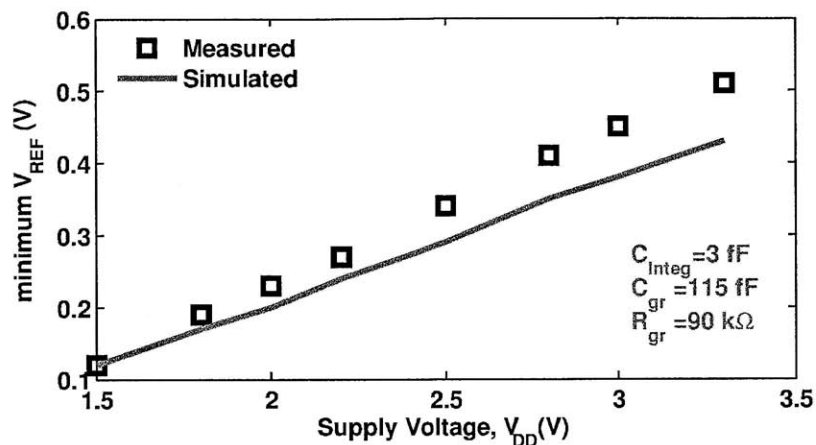
Figure 5-8: Minimum operable V_{REF}.

5.2.1 Delay Performance

The distribution of the measured channel delay is shown in Figure 5-9. In Figure 5-10, the delay of the full-swing design slightly increases at lower supply voltages. For the low-swing topology, V_{REF} sets the voltage swing on the wire. As V_{REF} is reduced, the delay of the channel also increases (Figure 5-11). This appears to be mainly limited by the noise margin and sensitivity of the sense amplifier at the receiver. For lower supply voltages (V_{DD}), the kickback voltage at the receiver input node is reduced and thus allows the low-swing design to operate at lower voltage levels. The data in Figure 5-11 could be further improved with better amplifier design but nonetheless illustrates an advantage of the low-swing design. At V_{DD}=1.5 V, the delay of a full-swing design increases, whereas the low-swing design maintains a smaller delay by tuning V_{REF}.

A minimum width M4 aluminum wire (L=1 mm) is included as a reference wire. For the low-swing topology, the measured delay on the M4 wire is 0.497 ns, and the delay on the graphene wire with L=1 mm is 13.915 ns. The actual M4 wire delay is likely smaller than 0.497 ns as the measurement is dominated by the transmitter and receiver delay. The graphene wire width is 6x that of the M4 wire but 58x smaller
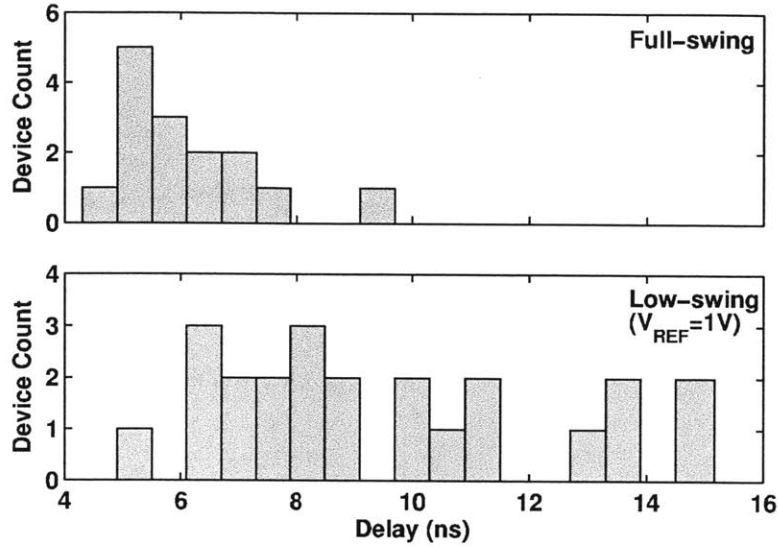
Figure 5-9: Histogram of delay measurements (L=0.5 mm) at V_{DD}=3.3V. For the low-swing design, the data plotted is at V_{REF}=1 V, where the channel delay is minimized.
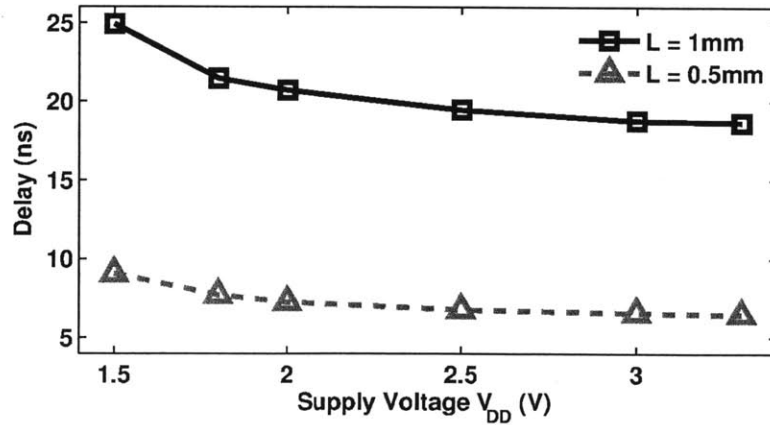
Figure 5-10: Measured channel delay of graphene wire in a full-swing topology.
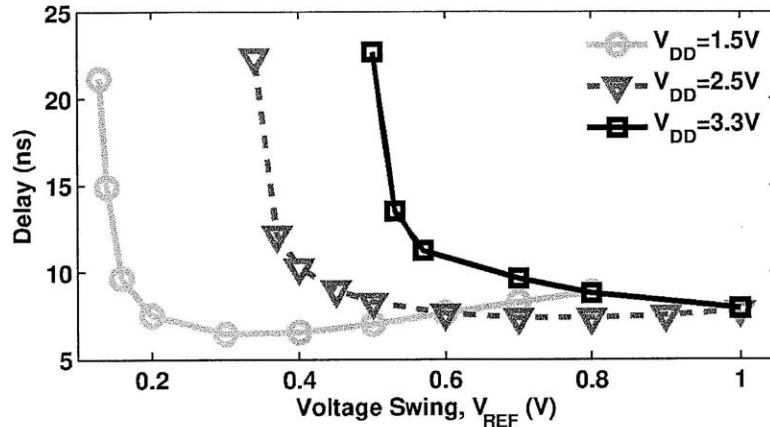
Figure 5-11: Measured channel delay of graphene wire in a low-swing topology (L=0.5 mm).

in thickness. The resulting cross-sectional area of the M4 wire is roughly 9x larger than that of the graphene wire. While a fair comparison may be difficult, the (per width) delay performance of the graphene wire clearly underperforms the M4 Al wire. The performance of the graphene wire is limited by its high sheet resistance (600-900Ω/sq). Ultimately, better control over the growth process is needed to grow highly uniform multi-layer graphene films.

5.2.2 Energy Performance

While the total energy of the low-swing design shows modest improvement (1.4-2.1x) over the full-swing design, this is limited by the receivers in this work. In comparison, the transmitter energy of the low-swing design is 4-4.7x lower than that of the full-swing design. Figure 5-12 shows the energy profile of two graphene wires. For both design topologies, only 10-40 % of the total energy is dissipated through the transmitters. While the transmitter energy scales with wire length, the receivers are clocked and require relatively constant conversion energy. For longer wire lengths and higher density interconnect fabrics, the transmitter energy, and hence the overall energy improvements, is projected to increase.

Although the main advantage of the low-swing design comes from reducing V_{REF}
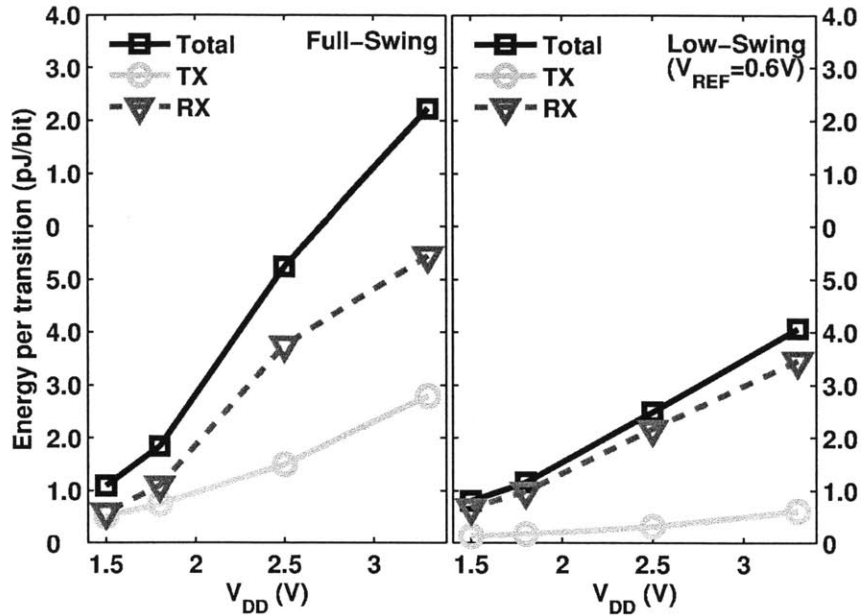
Figure 5-12: Measured energy profile of graphene wire.

and the energy dissipation on the wire, this energy only accounted for a small portion of the total energy. The energy dissipated from the (low-swing) NMOS drivers is less than 22 % of the transmitter energy and less than 4 % of the total energy as shown in Figure 5-13). Nonetheless, the solid lines in Figure 5-13 are from simulation results using values of R_{GR}=90 kΩ, C_{GR}=115 fF, and C_{INTEG}=3 fF, which shows excellent agreement with measured results.

To compare the capacitive load of the wires, another reference wire was included on top of the passivation layer. A 40 nm Ti/Pt wire with the same width as the graphene wire was fabricated. At V_{REF}=1 V, the measured energy dissipation of the NMOS transistors to drive the wire load is 103.23 fJ and 212.73 fJ for the graphene and Ti/Pt wires, respectively. The Ti/Pt wire is roughly 2.5x thicker than the graphene layer, and this is reflected in the larger energy dissipation from the NMOS drivers.

85
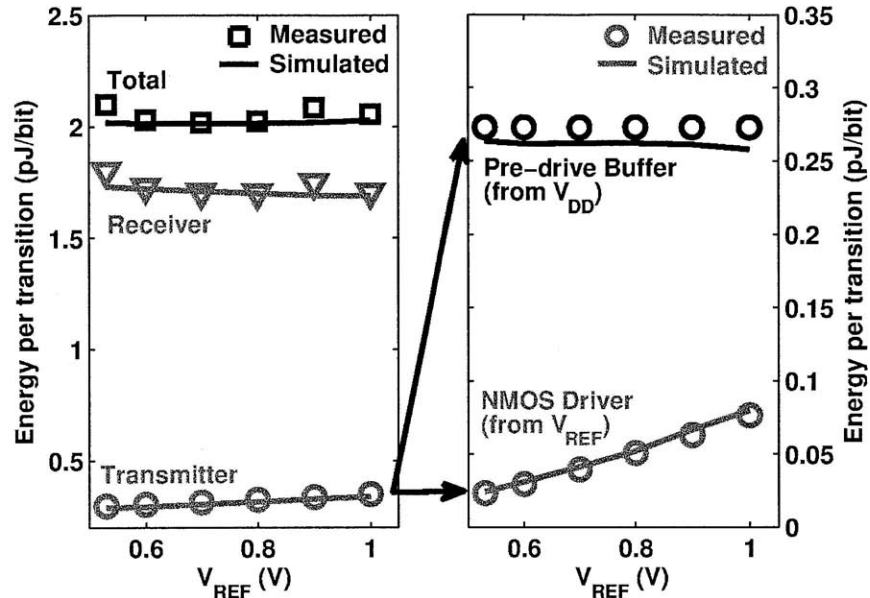
Figure 5-13: Comparison of measured and simulated energy profile of a low-swing data channel at V_{DD}=3.3V. The data points are measured values, whereas the solid lines are from simulations using values of R_{GR}=90 kΩ, C_{GR}=115 fF, and C_{INTEG}=3 fF.

5.2.3 Limitations

A summary of the on-chip measurements is shown in Table 5.1. The delay and energy performance show good agreement with expected values of R_{GR} and C_{GR}.

Most importantly, the graphene quality needs to be improved to be competitive with copper interconnects. The graphene sheet resistance is orders of magnitude higher than the CMOS aluminum wires. Despite the smaller wire capacitance, the resulting graphene data link delay is at least an order of magnitude higher than that of the CMOS wires. Recently, sheet resistances as low as 30 Ω/sq from 4-layers have been reported from p-doped graphene films [55]. Chemical doping of $AuCl_3$ has also been proven to be very effective on 1-2 layer graphene films, reducing their sheet resistance up to 77 % [83]. Ultimately, multilayer stacking and doping is necessary to achieve comparable numbers to copper. Despite this large performance gap, this chip demonstrates reliable signal transmission over long graphene wires which has not been demonstrated before. Except for the fabrication problems outlined in Section 3.3.3,

Table 5.1: Summary of measurements

Technology	0.35 μm, 3.3V CMOS	
Graphene Wires	W=4 μm, L=0.5, 1 mm $\sim$ 800 Ω/sq	
	Full-Swing	Low-Swing
Delay (L=0.5mm)	4.7 - 9.6 ns (avg. 6.2 ns)	5.2 - 15.2 ns (avg. 9.5 ns)
Transmitter Energy Total Energy	2.6-2.8 pJ/bit/mm 6.2-8.2 pJ/bit/mm	0.3-0.7 pJ/bit/mm 2.4-5.2 pJ/bit/mm
Energy-Delay-Product	(low-swing vs. full-swing) up to 3.3x improvement	
Bit error rates	less than 2x10^{-10}	
Minimum voltage swing	100 mV (@ V_{DD}=1.5 V) 500 mV (@ V_{DD}=3.3 V)	
Data rates	1 - 50 Mbps	

the graphene sheet itself does not deteriorate throughout the integration and testing process. The measured results also show excellent agreement with expected values. This gives us a baseline for projecting future enhancements in both graphene quality and circuit designs. We expect the performance to improve as the film becomes more uniform, although finding a reliable method that grows multiple layers while suppressing the interlayer and edge scattering effects remains a future step.

5.3   Summary

This chapter characterizes the performance of low-swing signaling on monolithically integrated graphene wires. CVD-grown graphene was fabricated into wires up to 1 mm in length on a 0.35 μm CMOS chip. This work primarily focuses on technology demonstration, and thus a more advanced CMOS process was not needed and not used here. Reliable operation of end-to-end data communication on these graphene wires was demonstrated, achieving bit error rates below 2x10^{-10}. A low-swing sig-

naling technique was used to achieve a transmitter energy of 0.3-0.7 pJ/bit/mm and a total energy efficiency of 2.4-5.2 pJ/bit/mm. Despite the high sheet resistivity of as-grown graphene ($> 600\Omega/\text{sq}$), integrated graphene links running at 50 Mbps was demonstrated. On-chip measurements were in good agreement with the simulated results, and recent work shows promise as the reported sheet resistivity is an order of magnitude smaller than the graphene devices used in this work. Ultimately, cleaner processing steps and optimized CVD growth conditions are necessary to produce higher quality graphene films. Other unique electrical and thermal properties of graphene may also play an important role in establishing graphene as a viable replacement for copper interconnects.

Chapter 6

Field-Programmable Gate Array using Graphene Wires

In Chapter 2, we suggested that using few-layer graphene films can be advantageous for subthreshold circuits. Graphene wires have small wire capacitance and thus can have a large impact on lowering the energy dissipation. In this chapter, we demonstrate a subthreshold field-programmable gate array (FPGA) using graphene wires as parts of the interconnect fabric.

6.1   Designing a FPGA

FPGAs are a common form of reconfigurable logic. FPGAs are often used for rapid prototyping or in systems that require flexibility. Building such flexibility comes at the cost of slower speeds and higher energy dissipation. The trade-off between flexibility and performance in different hardware paradigms is well known and has been previously documented [111–113]. Dedicated hardware units achieve the best performance but are limited to executing a single task albeit very efficiently. On the other hand, general-purpose processors exhibit the most flexibility but have poor efficiency due to instruction fetching and decoding. Field programmable gate arrays (FPGA) bridge this gap between flexibility and performance. By allowing the hardware to be reconfigured, FPGAs have greater flexibility than dedicated hardware units. Since the

hardware configuration is fixed at runtime, FPGAs also achieve better performance than general-purpose processors [5, 111, 112].

With the rising cost of designing custom circuits, using a flexible and energy-efficient FPGA has become a more attractive solution in many IC systems. In this chapter, we describe the implementation of a low-power FPGA that uses graphene interconnects. Demonstrating a FPGA is a perfect candidate to implement a graphene-based CMOS platform. FPGAs inherently have an interconnect-centric architecture that is highly regular which also makes it an attractive and easy platform to integrate graphene wires. Furthermore, because the overhead of an FPGA lies in its reconfigurable interconnect structure, the importance of replacing metal wires with graphene has increased. Furthermore, global interconnects have been reported to dominate the total delay and energy of FPGAs [5, 6, 114].

6.1.1 Chip Architecture

A conventional FPGA includes an array of configurable logic blocks (CLB) that are connected by programmable switch matrices. The CLB is designed to implement an arbitrary logic function and typically uses lookup tables (LUT) to accomplish this. The content of the LUTs can be programmed to map the inputs to the desired output values. As several interconnect buses run throughout the logic array, the switch matrix at each junction directs the traffic by enabling and disabling the appropriate switches in the interconnect fabric. The FPGA is reconfigured by streaming in a set of bits that programs the LUTs, switch matrices, and other related logic elements. Because of the extra routing structure, FPGAs are typically larger, slower, and consumes more power than a dedicated hardware unit [113].

Various approaches have been taken to reduce this overhead. In [19], a nano-electromechanical relay replaced the NMOS pass transistors in a switch matrix. More conventional approaches have tried to optimize the interconnect architecture since their performance dominates the delay and energy of the FPGA [5, 6, 114]. Our approach is to replace the metal wires with low-capacitance graphene wires and demonstrate better energy efficiencies. Figure 6-1 shows a diagram of the test chip which

implements a typical LUT-based FPGA. The CLBs are configured in a 5x5 array.
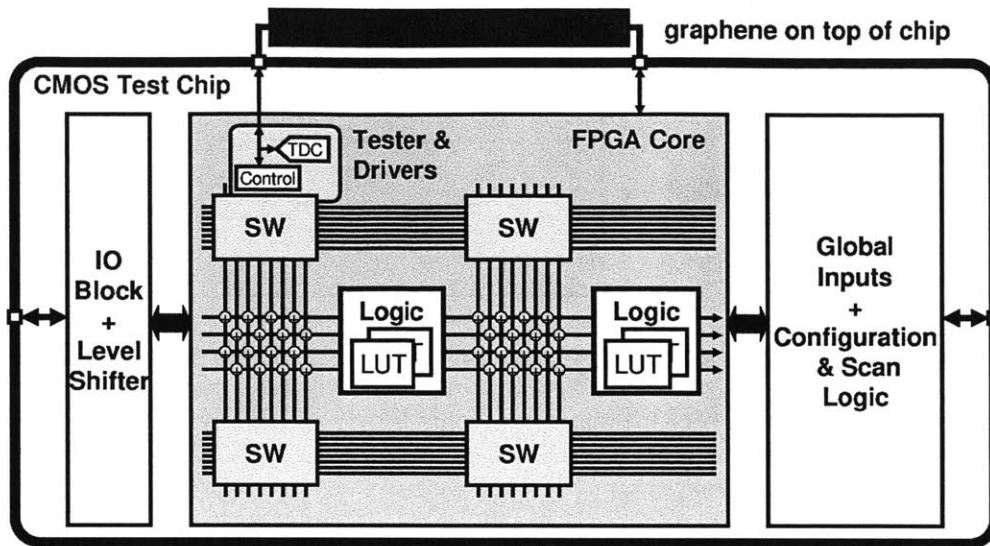
Figure 6-1: Overview of FPGA test chip. Graphene wires are integrated on top of the CMOS chip and interface to the switch matrices (SW).

The most simple logic block structure is a LUT connected to a flip-flop. Since the LUT can be programmed and the flip-flop can be bypassed by a multiplexer, any arbitrary combinational and sequential logic function can be realized. More advanced FPGAs often include dedicated hardware such as memory blocks or arithmetic units to improve the performance [115, 116]. Heterogeneous FPGAs structures are not considered in this work. Figure 6-2 shows the CLB structure where each logic block includes a cluster of two 4-input LUTs and two flip-flops, totaling 50 LUTs in the chip. The number of the inputs to the CLB that results in the most efficient resource utilization is given as $inputs = LUT/2 \times (cluster + 1) = 4/2 \times (2 + 1) = 6$ [117–119].

In recent years, many circuit designs have achieved very high energy efficiencies by operating in the subthreshold regime [120–124], where the supply voltage is smaller than the transistor threshold voltage. Despite the extreme robustness and scalability of CMOS logic, in subthreshold operation, the on current of a transistor and the on/off current ratio are significantly reduced. This results in slower operating speeds and in some cases may lead to functional failure as the supply voltage is decreased. One of the primary disadvantages of subthreshold circuits is the increased sensitivity
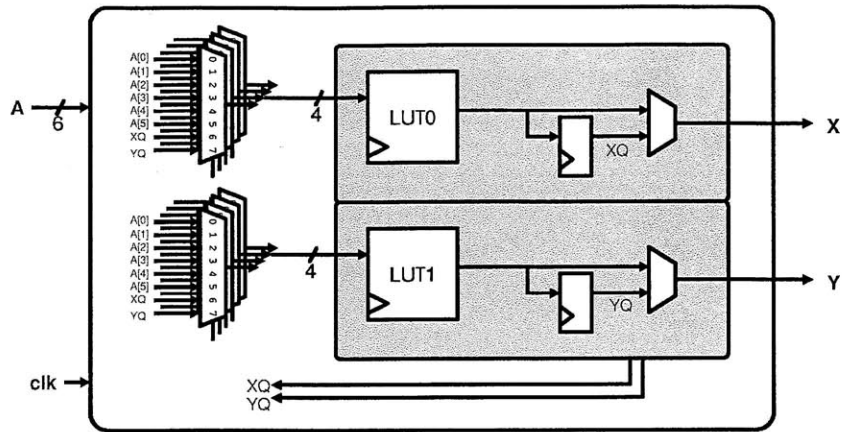
Figure 6-2: Implementation of Configurable Logic Blocks (CLB). Each CLB contains two clusters of LUT/FFs.

to variations. Therefore, strict design rules and statistical timing analysis are often used to ensure reliable operation of a subthreshold circuit [122, 125].

Ryan and Calhoun have recently demonstrated an energy-efficient subthreshold FPGA with low-swing dual-supply interconnects [114]. Their result highlights the potential of subthreshold FPGAs which can provide both design flexibility and low energy dissipation. In this work, the majority of the FPGA core blocks were synthesized using a standard cell library. We applied similar design rules that were found in [5, 120–123, 125]. The synthesis was optimized for subthreshold operation by: (1) limiting the fan-in to 3; (2) only using static CMOS logic; (3) using logic gates with short stacks of less than 4 stacked transistors; (4) re-mapping large multiplexers in to 2-1 multiplexers; and (5) upsizing critical clock and data buffers to ensure rail-to-rail outputs.

6.1.2   Interconnect Architecture

Figure 6-3 shows the most simple and conventional implementation of a switch matrix, where each connection point uses six NMOS transistors to connect the adjacent wires. This structure allows the wire to be driven in any direction and requires very little area. Despite this simplicity, NMOS pass transistors degrade the signals and a level-restoring buffer is required for every two or three switch matrices. A variant of this

92

structure uses back-to-back tri-state buffers instead of a single NMOS pass transistor. This bi-directional structure provides the same connectivity as before and provides stable output levels. However, since only one direction is enabled on each wire at any given time, the system always utilizes less than 50 % of the tri-state buffers.
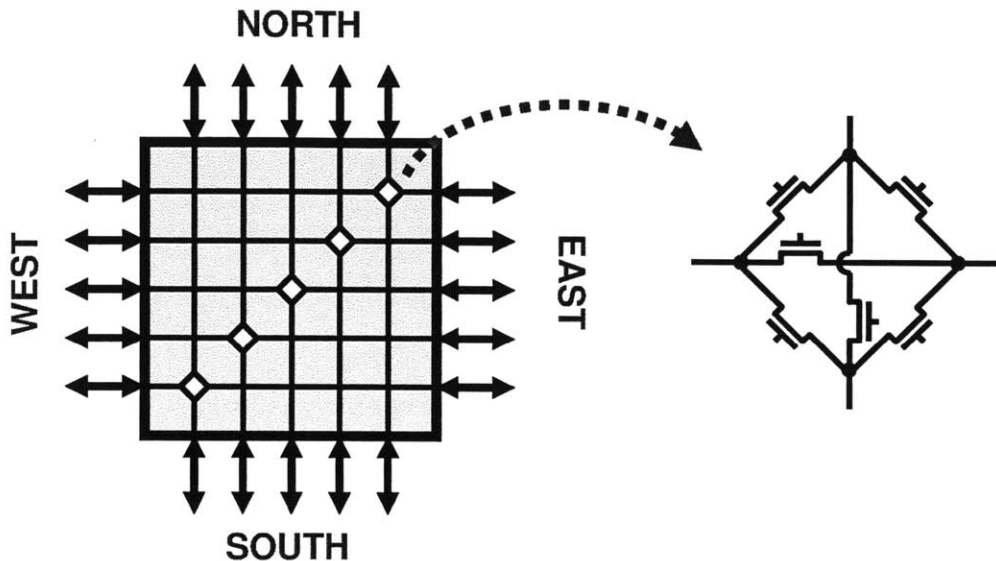
Figure 6-3: Diagram of a conventional switch matrix. At each intersection, six NMOS pass transistors are used.

In contrast, a uni-directional scheme allows each wire to be driven in only one direction and thus the tri-state buffers can be replaced with buffers which reduce area and improves the drive strength. One potential drawback is that the number of wires must be doubled to match the routing flexibility of the bi-directional scheme. However, the actual number of wires needed for each direction is less than expected because not all wires have the same directionality in a bi-directional scheme. Routing flexibility can be added by using multiplexers at the input of each driver [126]. In general, although more wires are needed in a uni-directional scheme, each wire has a lower critical path delay and thus the system is more tolerant of routing stress and allows smaller channel widths to be used to achieve a given performance level. From an architectural standpoint, recent examples show that uni-directional wires have better overall performance than the bi-directional approach since it provides better performance with less area [126, 127].

Many advanced commercial FPGAs have a combination of bi-directional and uni-directional wires to optimize the performance and flexibility in routing wires [115,116]. In addition, the interconnect hierarchy commonly includes double lines, hex lines, and long global lines [115,116]. The length of the line indicates the number of segments in between a switch matrix. A single line connects adjacent switch matrices and only travels one wire segment. Similarly, a double line will skip one switch matrix and connect every other switch matrix which is equivalent to traveling two wire segments. Global lines often span across the entire row or column and therefore have longer delays than a shorter line. However, routing a signal on a global line is often much more efficient than routing it through many single lines due to the overhead of each switch matrix. Commercial compilers maximize these resources and distribute the signal routing on all levels of wires to achieve the optimal performance.

Here, we implement a 10-bit uni-directional bus that runs throughout the logic array, and the programmable switch matrices provide connection points at each intersection. In addition, we have chosen a Wilton switch topology [128] which has better routability than the conventional switch matrix shown in Figure 6-3. For example, in the conventional switch matrix, the bottom wire on the west side can only connect to the bottom wire on the east side and the leftmost wire on the north and south side. To overcome this limitation, a Wilton switch box alternates the order in which each wire connects to its neighboring direction. Figure 6-4a shows the connections between the north and west side. Figure 6-4b shows the connections between all wires.

All single-length wires are implemented using the CMOS metal layers. The small size of the array precluded us from implementing long global lines and instead we designed double-length (L=2) wires. A total of 16 wire segments exist for double-length wires, which connect every other switch matrix. Figure 6-5 shows a diagram of the interface between the graphene wires and the switch matrices. At each L=2 segment, in addition to a reference (M5) wire, 3 redundant graphene wires are added for increased reliability. In addition, a local tester is added at each wire segment to monitor the quality of each graphene wire. The tester uses a low-overhead 3-bit time-to-digital converter (TDC) to measure the delay of the graphene data link (between
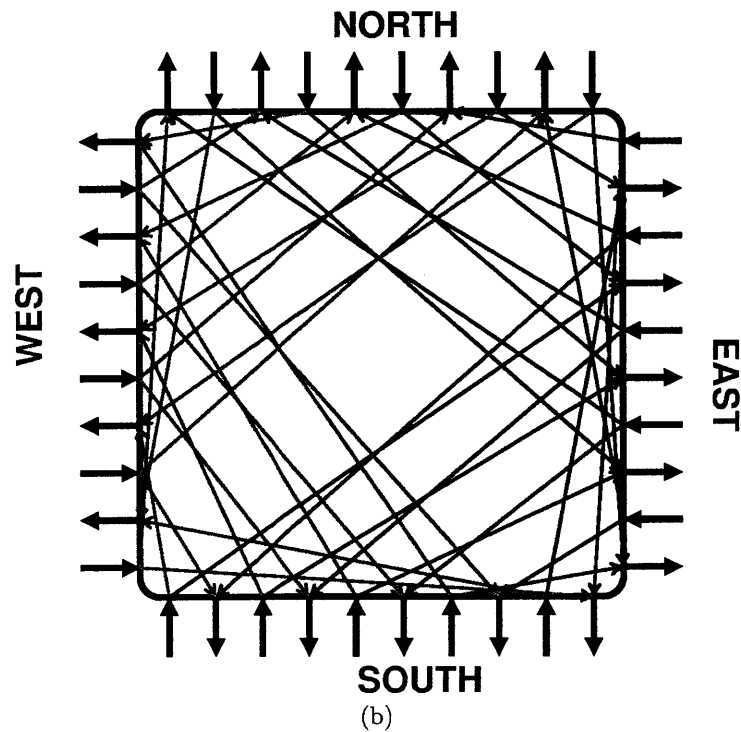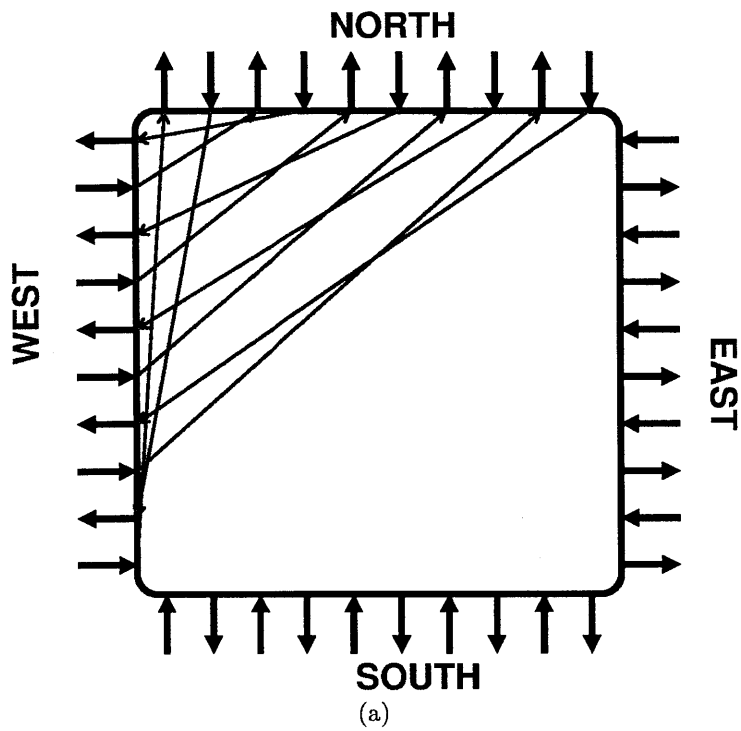
Figure 6-4: Implementation of programmable switch matrix (SW). (a) Only the North-West connections are shown. (b) All connections are shown.

A-B in Figure 6-5). Details of the TDC are described in Section 6.1.3.
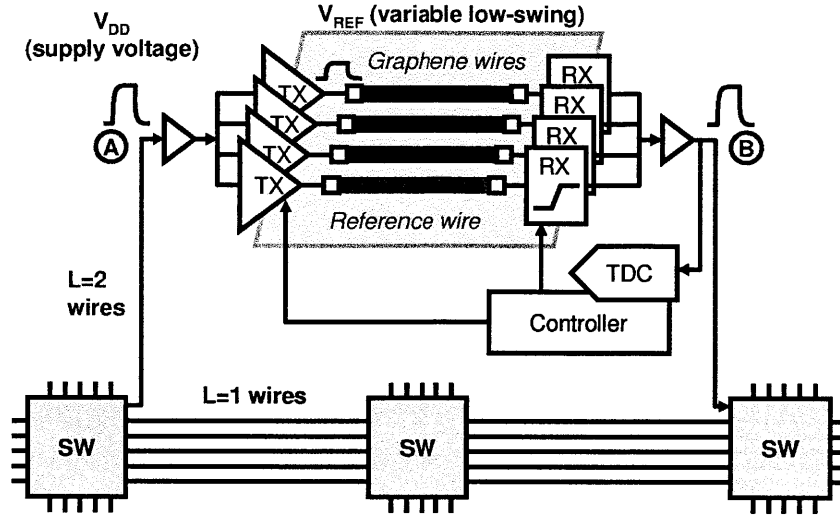
Figure 6-5: Diagram of graphene interface. Graphene is used to replace the horizontal double-length (L=2) wires. Each L=2 segment includes 3 redundant graphene wires and one reference (M5) wire. A TDC-based tester measures the delay (between A-B) for each wire.

We use a low-swing topology for energy reduction. The voltage swing on the graphene wire is set by V_{REF}, which also powers parts of the driver/receiver. The rest of the chip is powered by a single supply voltage (V_{DD}). Figure 6-6 shows the schematic of the low-swing driver and receiver. We implement a dual-supply tri-state buffer as the driver. This topology is robust and scalable.

Unlike the CMOS chip presented in Chapter 5, the receiver should be asynchronous in this test chip. Synchronous receivers like the pseudo-differential sense-amplifier implemented in Chapter 5 are robust, fast, and highly scalable. However, these advantages do not outweigh the cost of extra energy and complexity of routing a separate clock signal in an FPGA. Previously, level-converters have been used as a receiver to improve the energy efficiency of FPGAs [129]. However, conventional level-converters become too slow and power-hungry in the subthreshold regime. Several techniques have been explored to design a suitable subthreshold level-converter with a large dynamic range and low delay [130–132]. Here, we implement a level-converter with diode-connected PMOS devices [132]. This topology has low delay and large operating range and only uses a single voltage conversion stage unlike the

96
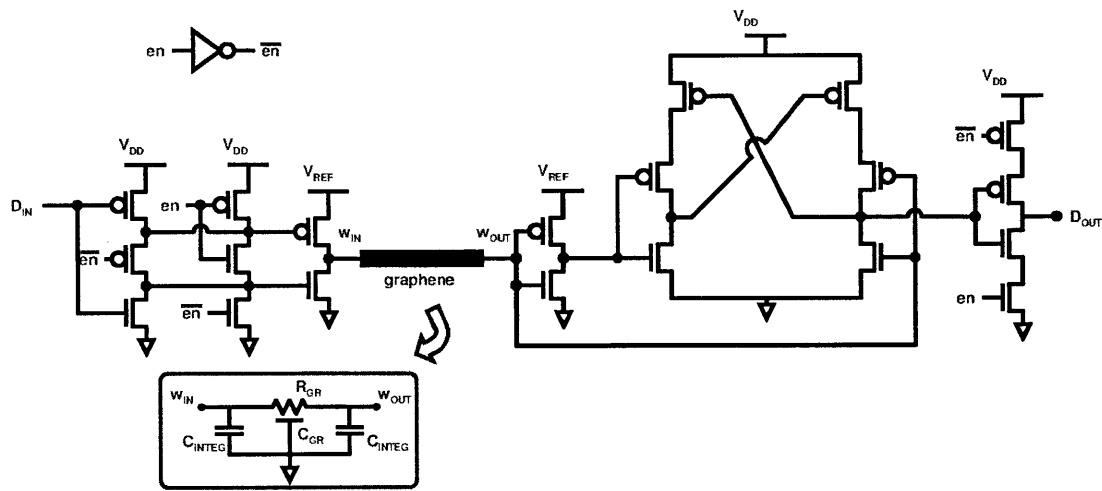
Figure 6-6: Implementation of low-swing driver and receiver.

design presented in [131].

One limitation of this low-swing implementation is that the the input node of the receiver inverter floats when the driver is disabled or when a graphene wire is not connected. An inverter can be added to the input of the receiver to prevent the voltage from floating in such cases, but this feature was mistakenly omitted in the final revision of the chip design. Although this results in static power dissipation in the receiver inverter, the measured static current through these inverters was less than half of the estimated worst case scenario. Overall, although the current levels were not negligible, the static power dissipation was not high enough to significantly alter the measurements.

6.1.3 Graphene Tester

The fabrication and integration process of graphene is not yet fully optimized and results in variations in the performance of graphene wires. Furthermore, some L=2 wires may be disconnected due to holes in the graphene wire. We have designed a local tester at each L=2 wire segment to monitor the quality of each graphene wire (Figure 6-5). The tester uses a TDC to convert the delay of the graphene data link into a 3-bit digital code. This measure of reliability is necessary for the test chip to function properly. The measured delay of each L=2 wire indicates not only how fast

the data link is but also whether graphene is properly connected or not.

In most integrated circuits, the analog information begins in the form of a voltage or current level. In contrast, a time variable or time difference can be chosen as the mode of information. Time-based signal processing has a wide range of applications, such as digital oscilloscopes, various CMOS sensors, and radio-frequency transmitters for digital phase-locked loops [133, 134]. TDCs are important building blocks used for time-based signal processing. A TDC converts a time interval between two signal edges into a digital code [133–136]. TDCs are frequently implemented with tunable delay lines where each buffer produces an equal delay. TDCs often require a temporal resolution in the picoseconds range and the main drawback of this topology is that the temporal resolution is limited by the delay of a single gate. Furthermore, the number of buffers increases exponentially with the resolution or the number of bits in the digital code. However, the objective of our test chip is to have distributed low-resolution TDCs rather than a single high-resolution TDC. Thus, each TDC only requires a modest number of bits and the delay of each buffer needs to be rather large to cover the entire dynamic range.

Figure 6-7 shows the circuit implementation of a single delay unit. Each buffer has a delay equal to τ. An NMOS footer device is added to the first inverter. This device is biased by a control voltage V_{CTL} which effectively modulates the delay of the buffer. Figure 6-8 shows that the delay of the buffer has a wide tunable range. When the NMOS device is biased in the subthreshold regime, the delay of the buffer increases exponentially.
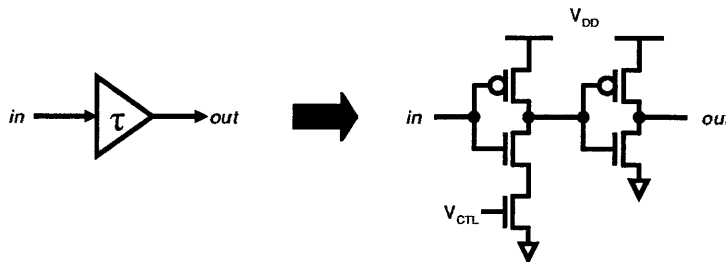
Figure 6-7: Schematic of delay unit in TDC.

Figure 6-9 and Figure 6-10 show histograms of the simulated buffer delay after
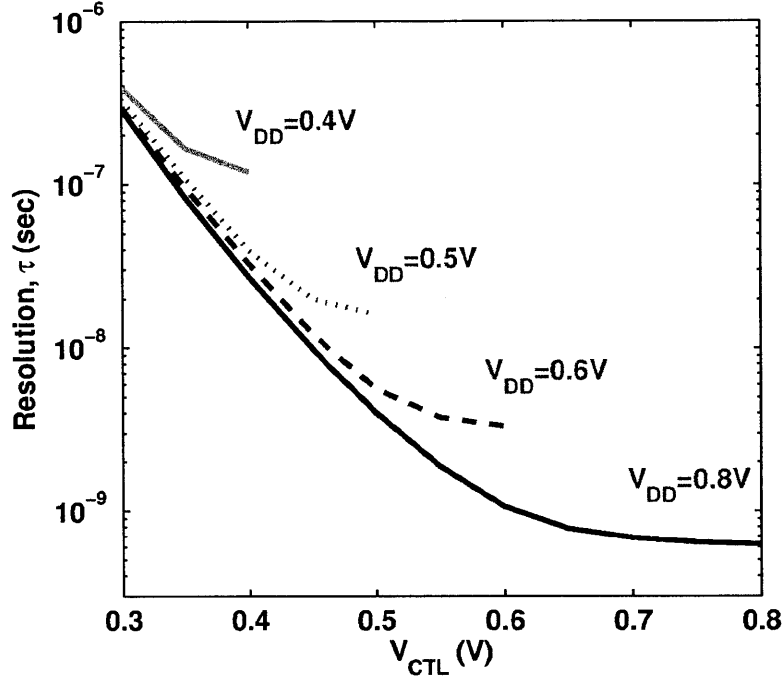
Figure 6-8: Simulated resolution of delay unit in TDC.

running Monte Carlo simulations. As expected, τ has much larger variations at low values of V_{DD} or V_{CTL}. However, τ is more sensitive to V_{CTL} than the power supply. The variation in τ increases rapidly as V_{CTL} decreases below the threshold voltage. Therefore, using a higher V_{CTL} and smaller τ guarantees better uniformity across all TDCs but has a limited delay range. If the graphene data links have excessive delays, one can increase the temporal resolution by lowering V_{CTL} at the cost of larger variations within a TDC and throughout the chip.

Figure 6-11 shows the diagram of the entire TDC. A reference signal is generated from the TDC controller (Figure 6-5) and travels through the driver, graphene (or M5) wire, receiver, and finally reaches the TDC. The reference signal is also directly fed into the TDC and this signal propagates through the delay line. Each buffer has an equal delay and provides the sampling clock edge to a register. Thus, each register will sequentially sample the data line at fixed time intervals.

Because of the time delay between the graphene data link signal and the reference signal, each register will sample a different value. Based on the encoded value of

99
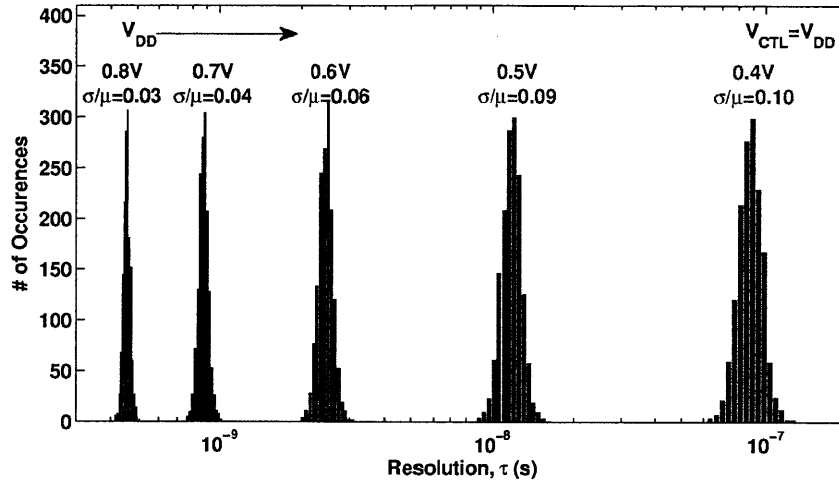
Figure 6-9: Simulated histogram of TDC vs. V_{DD}
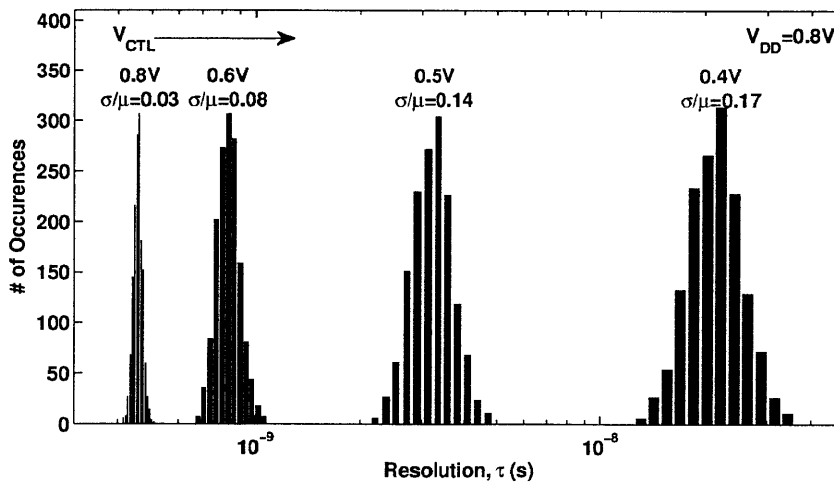
Figure 6-10: Simulated histogram of TDC vs. V_{CTL}
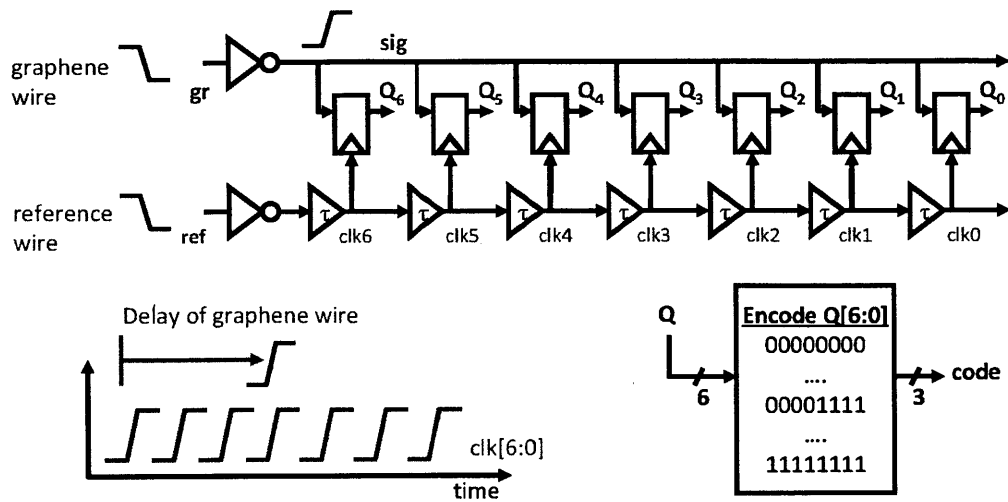
Figure 6-11: Schematic diagram of TDC. Each buffer has a delay equal to τ.

the registers, we can infer when the graphene data link signal arrived in comparison to the reference signal. Figure 6-12 shows this sequence in detail. The reference signal triggers the internal clock signals. The inverted graphene signal transitions to a binary one right after the third clock edge resulting in a digital code of 2. The TDC essentially converts the delay of the data link into a 3-bit digital code. The actual delay is between code×LSB and (code+1)×LSB.

This relationship is shown in Figure 6-13. Because the registers are sequentially triggered, only 8 different possibilities exist for $Q[6:0]$. These values are then encoded in to a 3-bit digital code. Similarly, each digital code, or unique value of $Q[6:0]$, indicates a possible range for the actual delay. For example, a digital code of 2 means that the delay is between 2τ and 3τ, where τ is determined by V_{DD} and V_{CTL} as shown in Figure 6-8. In this example, we can estimate the delay to be 2.5τ. Therefore, the absolute delay of each wire segment can be estimated by obtaining the 3-bit code at a given operating condition for V_{DD} and V_{CTL}.

6.2   Testing Methodology

Section 6.1 described the implementation details of the FPGA test chip. Since the FPGA chip serves as a platform rather than an isolated system, the test chip by itself

101
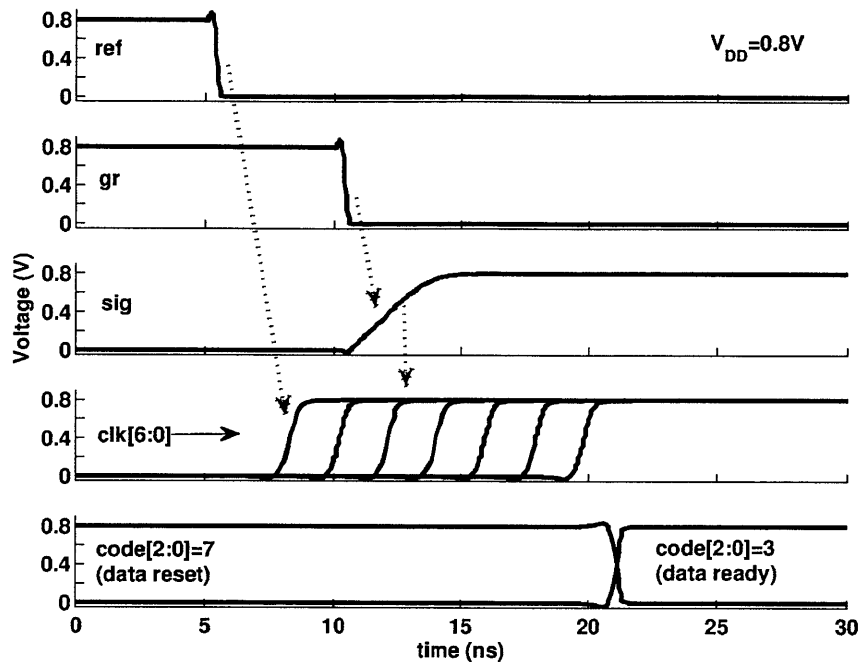
Figure 6-12: Simulated waveform of graphene tester.
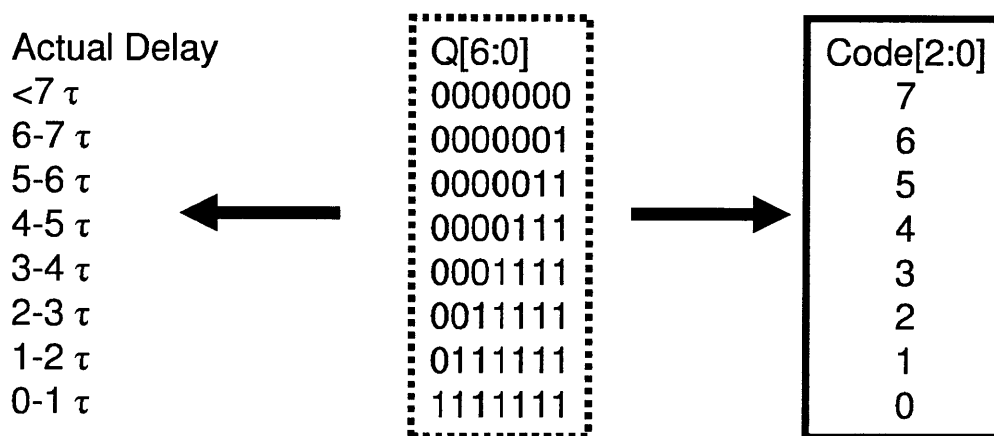
Figure 6-13: Relationship between TDC code and absolute delay.

does not perform anything interesting. The FPGA test chip needs to be configured to perform a specific task, for example, a 8-bit multiplier. This section describes the methodology to configure and test the FPGA chip.

6.2.1 Hardware Setup

The user must first define the function or application for which the FPGA will be used. This can range from simple arithmetic functions, such as binary addition or multiplication, to very complex digital control systems. The FPGA must then be configured to realize this function. FPGAs are a type of reconfigurable logic and thus the ability to program the internal core of the design is built into the chip. Each CLB, switch matrix, or other connection boxes have configuration data bits that can be programmed but are static when the test chip is running. The configuration bits typically provide the LUT contents or the select bits for multiplexers. The bitstream or bitfile is made up of all the configuration data bits.

In our design, the core of the FPGA design has roughly 300,000 transistors and 3,300 configuration data bits. A register scan chain is implemented so that the bitstream can be serially shifted in one bit at a time. Generating the bitstream is an important task and this process is described in Section 6.2.2. After the bitstream is assembled and uploaded to the FPGA, the test chip now behaves as a hardwired instantiation of the pre-defined function.

We have added an additional feature in our FPGA test chip. Before the FPGA is configured to run a benchmark, the test chip is programmed to assess the quality (delay) of each L=2 wire. Thus, the FPGA test chip has three primary functions: (1) run the graphene testers and retrieve the results; (2) upload the bitstream; and (3) run a benchmark application. Figure 6-14 and Figure 6-15 show the experimental setup.

A commercial Xilinx FPGA was connected to the test chip and serves as a *master* device to: (1) scan the tester results; (2) load the bitstream; and (3) apply the appropriate benchmark inputs to the test chip. The test chip is packaged and placed in a pin-grid array socket. Although the core of the test chip runs at voltage levels

103
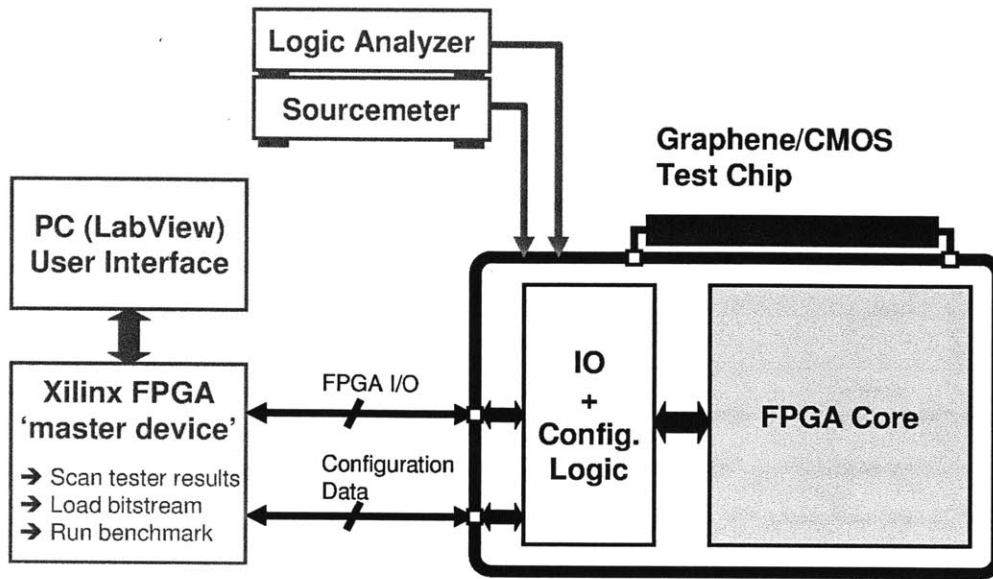
Figure 6-14: Diagram of experimental setup. A commercial Xilinx FPGA is used as a master device to configure the test chip and run benchmarks.
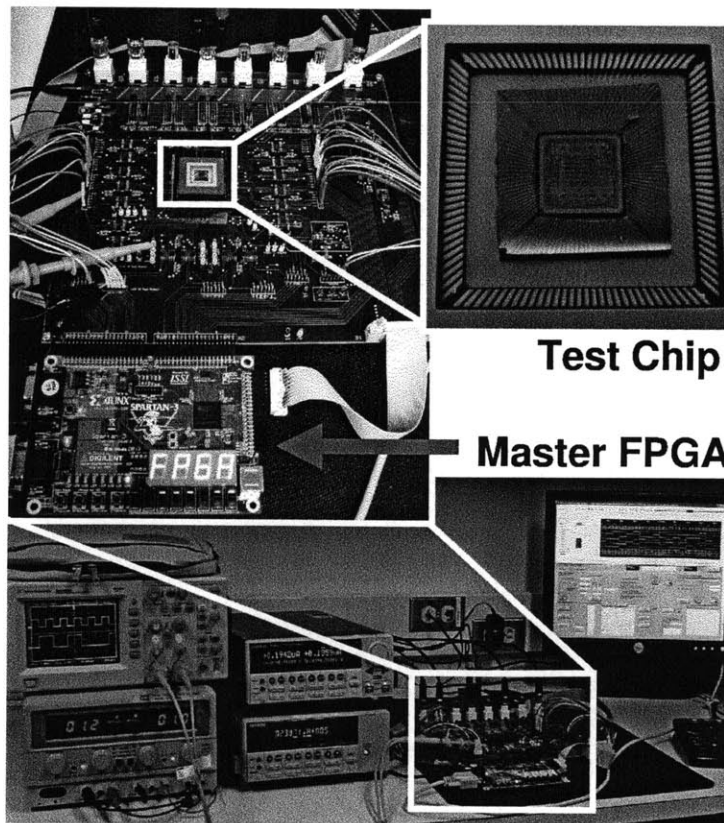
Figure 6-15: Photo of experimental setup. The test chip is wire-bonded to a pin-grid array package.

down to 300 mV, the voltage supply of the I/O ring was set between 0.8 V and 1.2 V. Multiple level-shifters were placed on the printed circuit board to interface to the Xilinx FPGA which operates on 3.3 V I/O. A digital-analog converter and programmable delay unit were added for calibrating the test structures on the FPGA chip.

The Xilinx FPGA board also provides serial communication to a computer. The user interface was designed using LabView. The results of the graphene testers are scanned from the test chip to the Xilinx FPGA which then get transmitted and are displayed on the LabView interface. Similarly, the bitstream is first loaded on the LabView interface and then gets transmitted to th Xilinx FPGA. The Xilinx FPGA then initiates the programming sequence and serially shifts in the bitstream to the FPGA test chip. The user may also select the clock speed and type of benchmark application through the LabView interface. The maximum clock speed that is allowed for the FPGA test chip is 100 MHz. The typical clock speed is in the kHz range under most operation conditions of interest.

6.2.2 Software Flow

The description of each benchmark application is written in Verilog, which is a type of hardware description language (HDL) that is widely used in the electronic design industry. Among many other uses, source codes written in Verilog are used in the development, verification, synthesis, and testing of digital designs [137]. In a typical computer-aided design (CAD) flow, the compiler takes a HDL source and generates a bitstream by synthesizing the design and mapping the configuration bits for a specific FPGA architecture. However, we adopt a different CAD flow since our FPGA is small in size and has various parts that must be customized. Figure 6-16 shows the CAD flow we use.

We use a set of tools that are frequently used in conjunction with *Versatile Place and Route (VPR)* [138–140]. *VPR* is used for FPGA placement and routing and is commonly used for FPGA research. First, *Odin* converts the Verilog source to a flat netlist of logic gates [141]. Then, *ABC* optimizes the logic structures and

Verilog source

```
module counter(clk,cntr);
input        clk;
output [15:0] cntr;

reg [15:0] cntr;
always @(posedge clk)
    cntr <= cntr + 1;

endmodule
```

Architecture description

```
Array size: 5x5

Cluster size: 2
Inputs: 6

Switch topology: Wilton
Interconnect:
    uni-directional
    single-driver
...
```

Odin

Front-end
synthesis

ABC

Optimization
& mapping

T-VPack

Clustering

VPR*

Placement &
routing

** Modified
original program*

**Existing
tool flow**

custom bitstream generator

...11001101010... **Bitstream
(~3300 bits)**

Figure 6-16: CAD flow for generating a bitstream from a Verilog source.

maps them into 4-input LUTs and flip-flops [142, 143]. *T-VPack* further packs these LUTs and flip-flops into clusters [140, 144, 145]. Finally, *VPR* places each cluster and finds the most efficient signal routing between all logic blocks [138–140, 146, 147]. In addition to the Verilog source file, a description of the FPGA architecture must be supplied. The source code of *VPR* was modified so that the final structure of the *VPR*-generated switch matrices would more closely resemble our FPGA design. The output files from *Odin*, *ABC*, *T-VPack*, and *VPR* describe the names of all signals, logic function of each LUT, connection between each wire and switch matrix, and placement of each logic cell. A number of *perl* scripts were written to parse these files and extract the configuration of each block in our FPGA design. The final bitstream is a long sequence of binary digits which can then be uploaded to the FPGA through the LabView interface.

Commercial FPGAs have hundreds of thousands of logic cells and can be configured to perform very complex digital systems. Although the test chip implemented in this work has only 50 LUTs, we have created a suite of 11 benchmark applications that include arithmetic, memory, and logic functions. Both combinational and sequential logic structures are considered. Table 6.1 provides a brief description for each benchmark.

Figure 6-17 shows the amount of resources each benchmark uses. The test chip has a total of 50 LUTs and 25 CLBs. The percentage of logic resources used, or resource utilization, for LUTs and CLBs do not necessarily match since a CLB may have only one active LUT. The 3-stage pipelined multiplier used the most number of resources; 48 out of 50 LUTs were occupied. The non-clocked multiplier implements the same function (4-bit multiplication) but uses only 32 LUTs. In general, applications that are clocked use more FPGA resources. Figure 6-18 shows a screen capture of running two different benchmark applications.

Table 6.1: Summary of benchmark applications.

Identifier	clocked?	Description
mult (3-st)	Yes	4-bit multiplication, 3-stage pipelined
mult	-	4-bit multiplication
mac	Yes	3-bit multiply-accumulate, 2-stage pipelined
adder	-	8-bit addition
comparator	-	8-bit comparison (<,>,==, =0)
decoder	-	4-to-16 decoding
counter	Yes	up/down 16-bit counter
filter	Yes	4-bit 4-tap moving average
freq divider	Yes	1-256 frequency divider
fifo	Yes	synchronous 3-bit first-in first-out buffer
mux	-	1-to-16 multiplexer
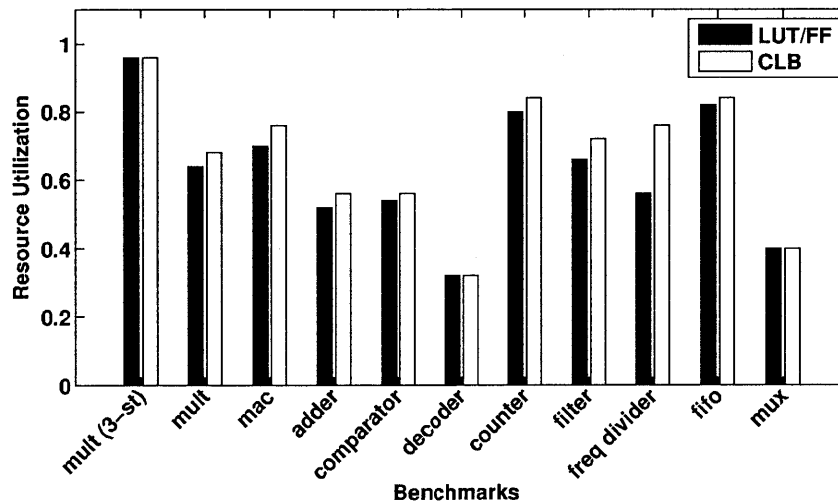
Figure 6-17: Resource utilzation of various benchmarks.
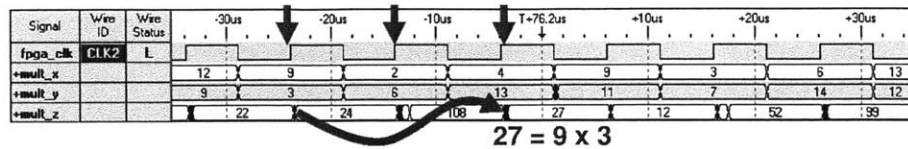
(a)
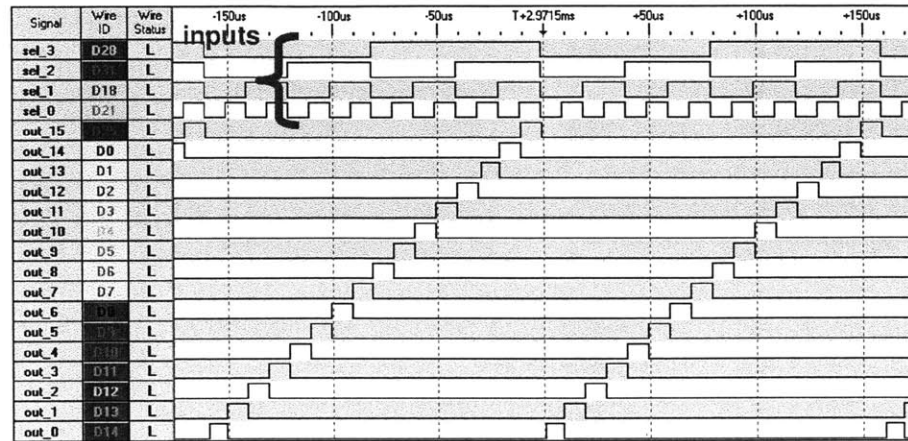
(b)

Figure 6-18: Measured waveforms from logic analyzer while running two example benchmarks: a (a) 3-stage pipelined multiplier and a (b) 4-to-16 decoder.

6.3 Measured Results

The FPGA test chip was fabricated in a 0.18 μm CMOS process. The test chip was packaged after completing the integration procedures outlined in Section 3.3.2. Figure 6-19 shows an image of the CMOS chip and the integrated graphene wires. We have successfully demonstrated reliable system operation using graphene interconnects.

6.3.1 TDC and Graphene Tester

Before we configure the FPGA, we run the graphene testers and retrieve the results for all the L=2 wires. Figure 6-20 shows the measured delay from the testers for both the graphene and reference M5 wires. The delay increases as the supply voltage scales down due to lower current levels. In general, the graphene wires have a larger variation than the reference M5 wires. The material non-uniformity contributes to this large variation in device performance. Nonetheless, as also seen in Figure 6-21,

109
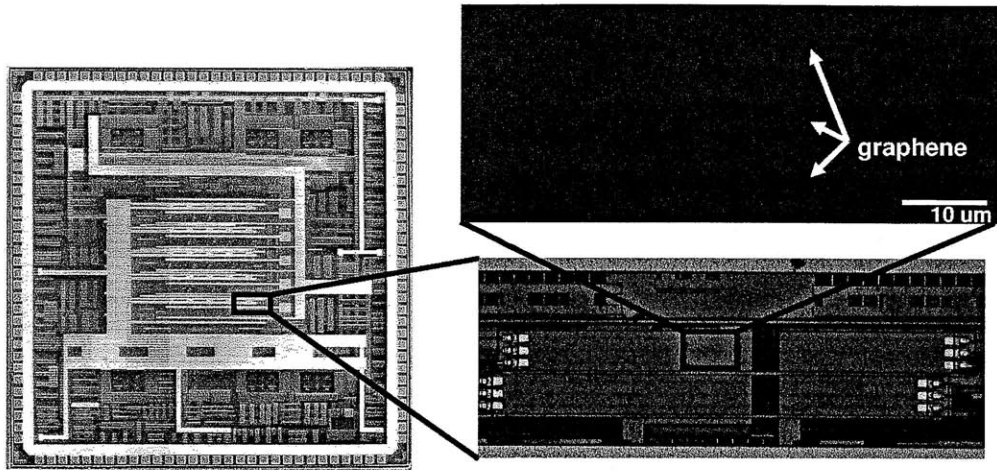
Figure 6-19: Image of 0.18 μm CMOS test chip and integrated graphene wires.

at low V_{DD}, the average delay of the graphene wires is slightly smaller than the M5 wires.
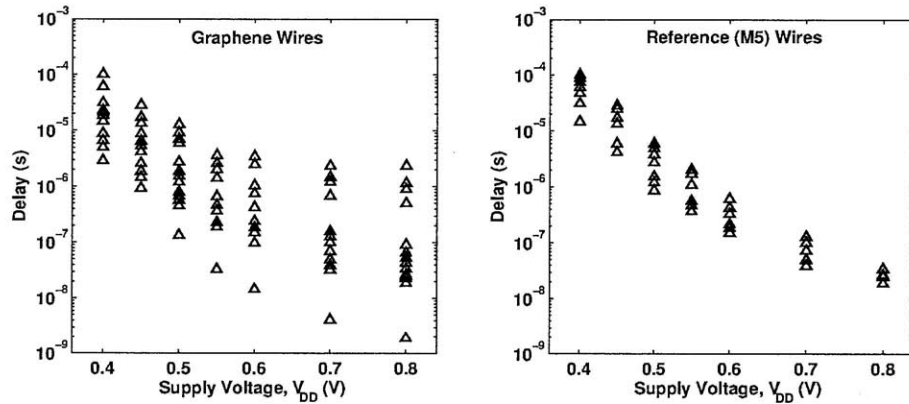
Figure 6-20: Measured results (data link delay) from the TDC-based tester for (a) graphene and (b) reference (M5) wires.

We have also estimated the delay of the low-swing data link. Using a commercially available field-solver, QuickCap, we have estimated the capacitance of the graphene wires using the known dimensions and material properties of the interlayer dielectrics. Nominally, the estimated capacitance of the graphene and M5 wire is 18.7fF and 52.5fF, respectively. The sheet resistance R_{SH} of the graphene wire is estimated to be roughly 270 Ω/sq, although the effective R_{SH} may have a large range between 100 and 1000 Ω/sq. Even if we assume an extremely large range of R_{SH}=100-1000 Ω/sq

110
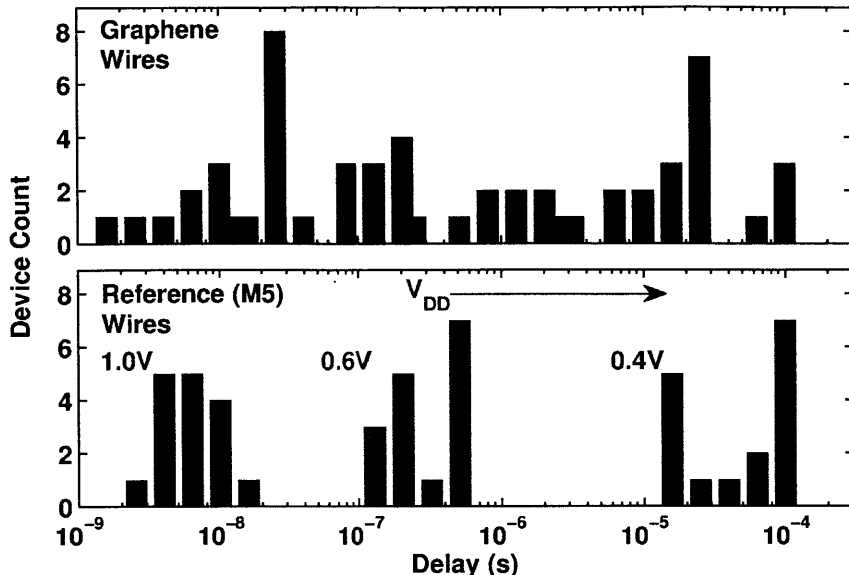
Figure 6-21: Histogram of L=2 wire delay measured from testers at various supply voltages.

and C_{GR}=10-100 fF, the simulated data lines do not adequately explain the wide variation in the measured results especially at low supply voltages.

Note that the measured delay is a combination of the wire and the transistors from the driver/receiver. Since the delay of the M5 wires (Figure 6-20(b)) show significantly less variation, we can rule out the driver/receiver as the main source for this variation. Furthermore, at low supply voltages, the resistance of the wire has almost negligible effect on the wire delay as the transistor resistance dominates. We suspect that the contact resistance or variation in the TDC (when V_{CTL} is small) is mainly responsible for the large variation in Figure 6-20.

Figure 6-22 shows the measured transfer curve of the TDC. Before the TDC is activated, the internal registers are reset to one. Thus, the total energy of the TDC depends on the delay of the graphene signal which triggers the internal registers to sample a zero or one. Figure 6-23 shows the measured TDC energy as a function of the digital code. When the code (or delay) is small, more registers are triggered to sample a value of one and thus the total energy is higher.

The minimum TDC energy is shown in Figure 6-24. The TDC achieves a minimum

111
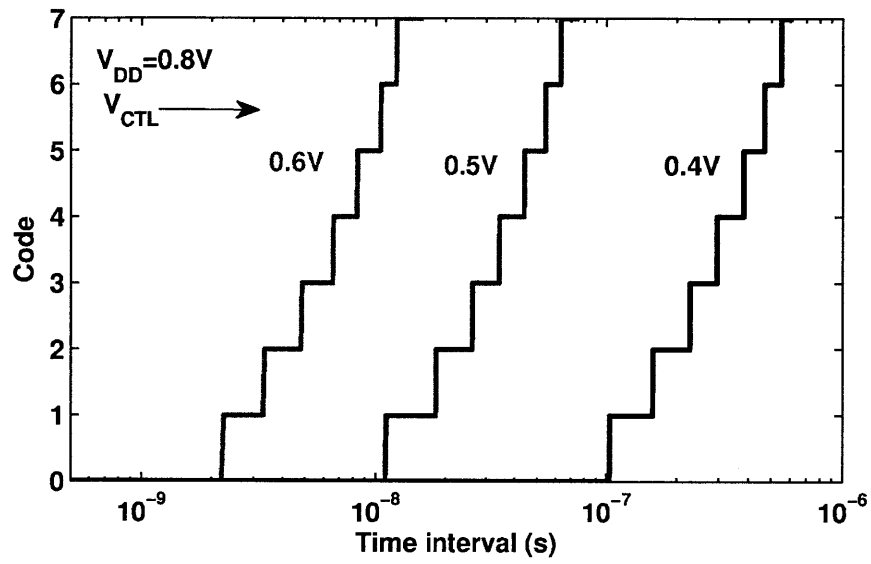
Figure 6-22: Measured TDC transfer curve.
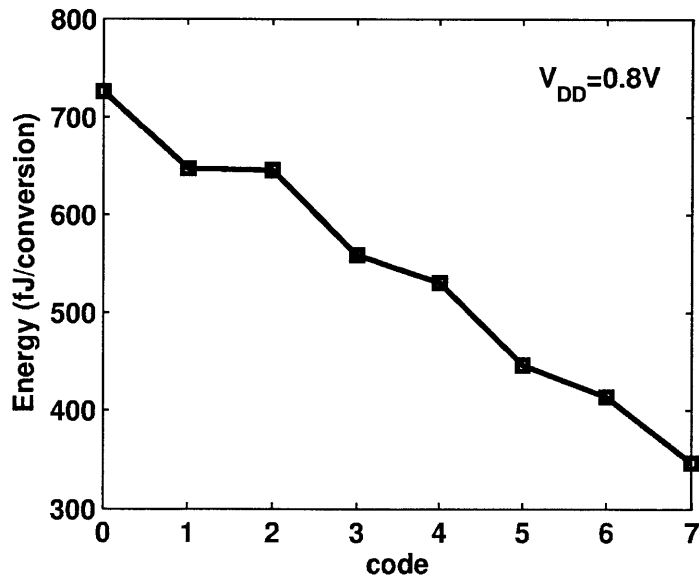
Figure 6-23: Measured TDC energy per conversion.

energy of 67.8 fJ/conversion around 350 mV supply.
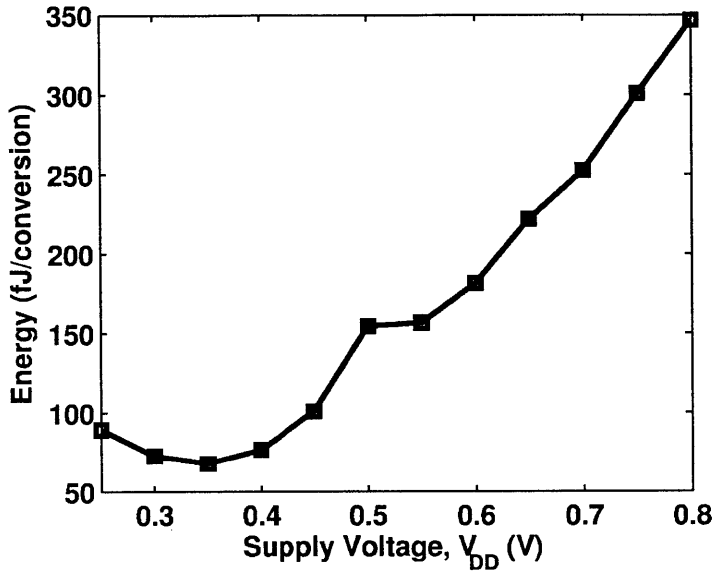
Figure 6-24: Measured TDC energy as a function of supply voltage.

6.3.2 System Energy

After the TDC-based tester characterizes each graphene wire, we re-generate the bitstream to activate only the desired graphene or M5 wires. The following figures show measured results while running a representative benchmark (3-stage pipelined multiplier). The impact of using graphene wires is discussed in detail in the following section.

Figure 6-25 shows the power consumption of the test chip. Running a benchmark consumes the most amount of power since the majority of the chip becomes active. A large part of this power appears to be related to the clock signal. Although clock gating is applied aggressively throughout the design, all core units have local clock buffers and become active regardless of the data signals.

Figure 6-25a shows the total power consumption of the system. Decreasing the supply voltage has a large impact on the energy efficiency since the total energy of the system is proportional to $C_{eff}V_{DD}^2$ where C_{eff} is the effective switching capacitance.

113
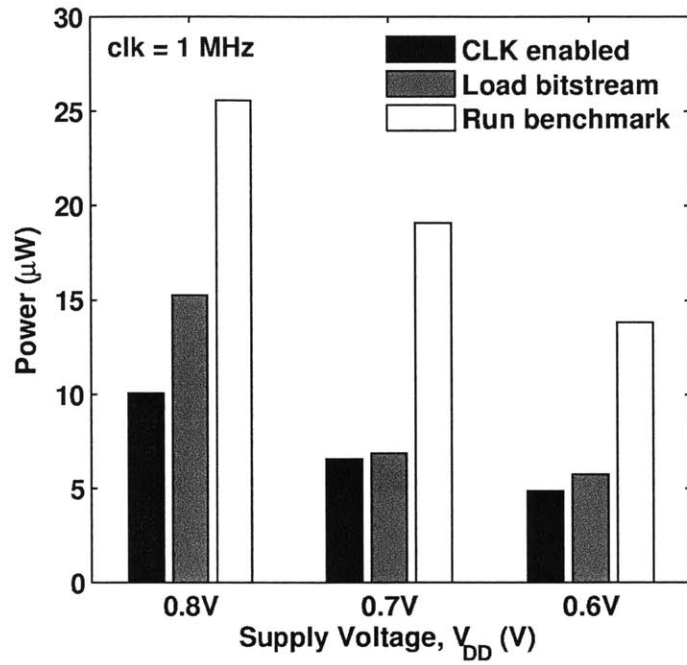
(a)
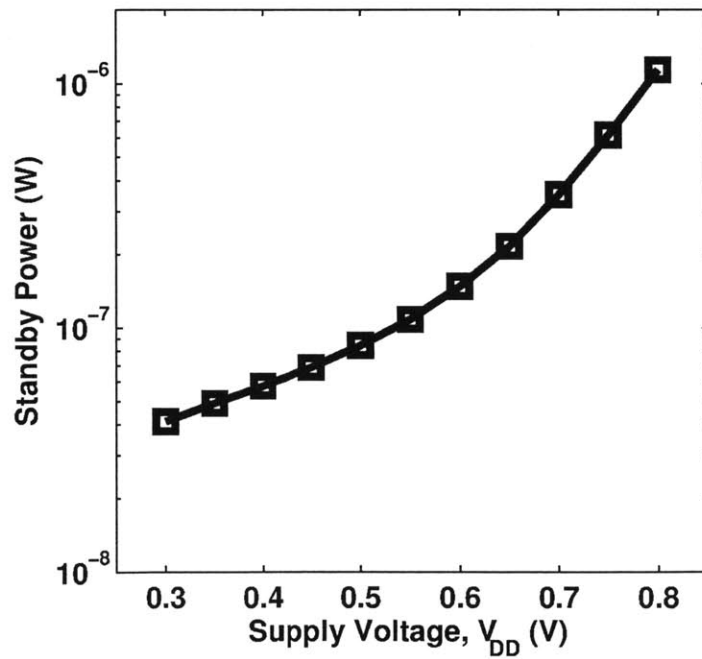
(b)

Figure 6-25: Measured power consumption of chip (a) at various run modes and when (b) system is idle.

We have verified that the test chip operates down to V_{DD}=0.3 V, well below the threshold voltage. However, the system achieves minimum energy operation around V_{DD}=0.45 V. When V_{DD} <0.45 V, the leakage energy begins dominate and increases the total energy.
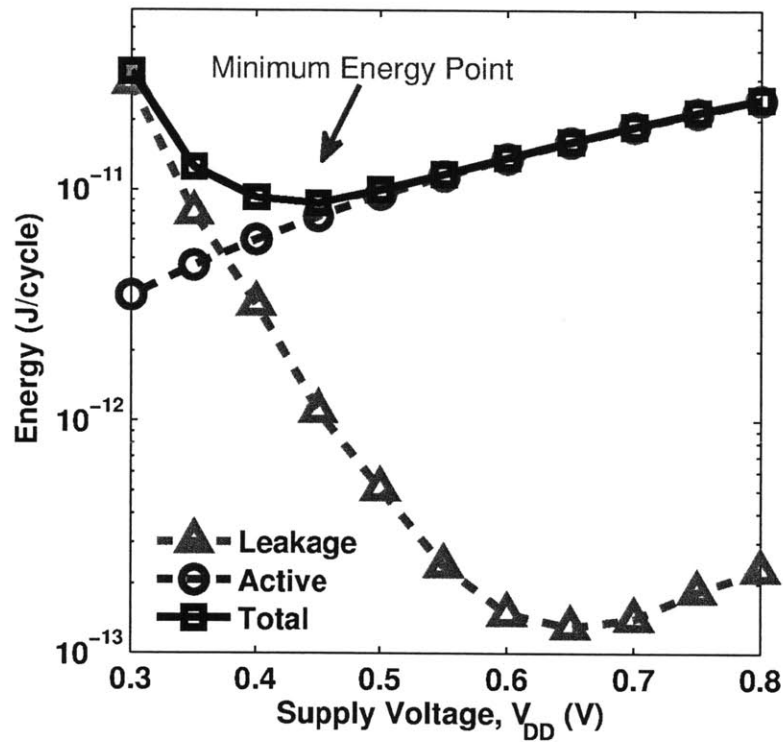
Figure 6-26: Measured total energy of FPGA chip.

Although we have chosen a representative benchmark, the actual frequency and energy profile will be slightly different for each benchmark application. The 3-stage pipelined multiplier uses the most number of resources in the FPGA. Other benchmarks use less resources and some are combinational logic functions that do not require a clock signal (Figure 6-17). Figure 6-27 shows the energy and frequency of various benchmark applications.
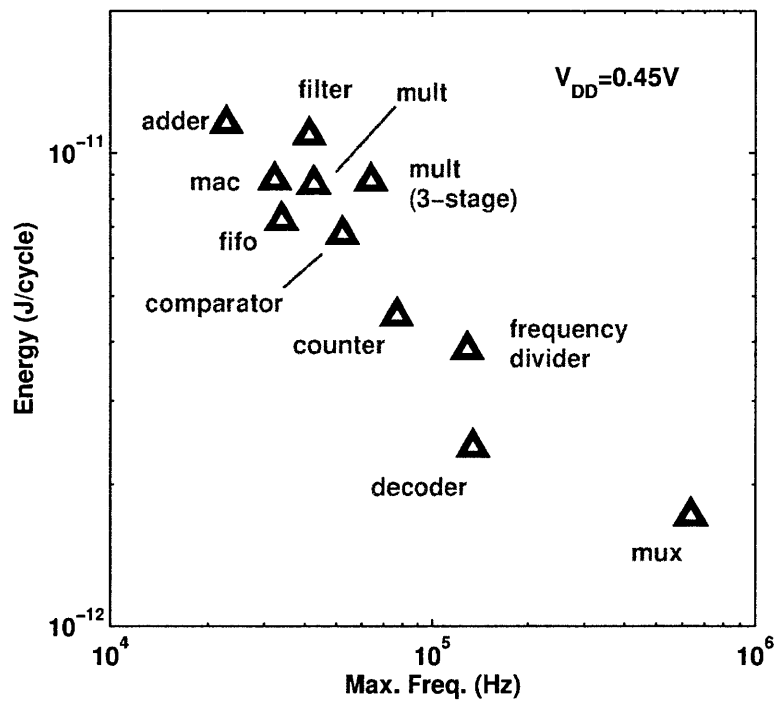
Figure 6-27: Performance of various benchmarks at V_{DD}=0.45 V.

6.3.3 Comparison of Graphene and Metal Wires

The test chip can be programmed to route the signals on one of the 4 wires for each L=2 segment. To understand the effect of using graphene wires, we generated two sets of bitstreams that either activated the graphene or M5 wires. We enabled the M5 wire if a particular segment did not have any working graphene devices. The following figures are measured results of the test chip while enabling the graphene wires and running a representative benchmark (3-stage pipelined multiplier).

Figure 6-28 shows the maximum operating frequency and (L=2 wire) energy of the test chip while enabling the graphene wires. Rather than plotting the total system energy, we only plot the energy dissipated from the secondary supply V_{REF} which is the only source that directly affects the graphene or M5 wire. Furthermore, the total system energy is dominated by all other logic blocks and thus comparing the L=2 wire energy gives us a more accurate picture of the performance difference between graphene and M5 wires. The delay of the L=2 wire, however, affects the critical path delay and has a direct impact on the frequency of the system. Figure 6-29 plots the same data as in Figure 6-28 except that the data points are normalized to the case when the M5 wires are enabled.
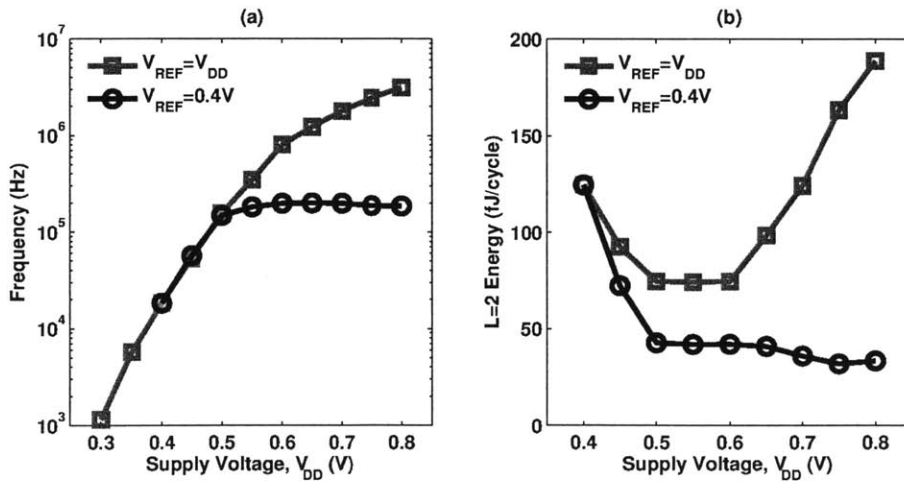
Figure 6-28: Measured (a) maximum operating frequency and (b) L=2 wire energy (from V_{REF}) while running a representative benchmark (3-stage pipelined multiplier) and enabling the graphene wires. Labels indicate relative system performance when M5 wires are enabled (not shown in figure).

117
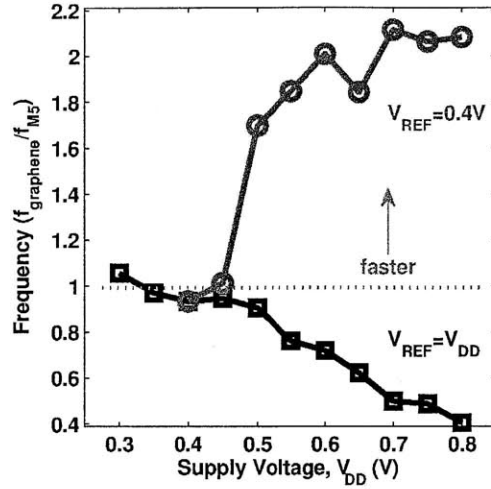
(a)
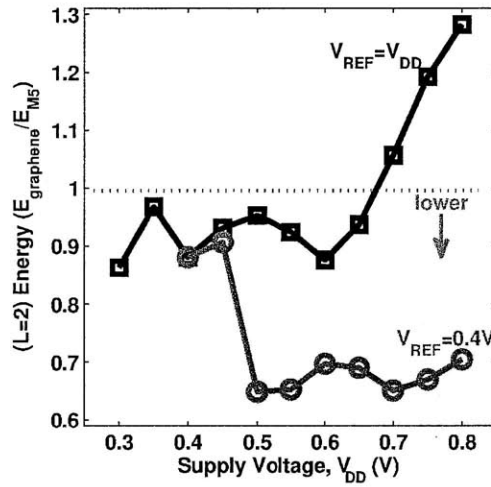
(b)

Figure 6-29: Measured (normalized) (a) maximum operating frequency and (b) L=2 wire energy of the system when graphene wires are enabled. The data points are normalized to the case when M5 wires are enabled.

We can scale V_{DD} while maintaining a constant $V_{REF}(=0.4$ V) or while $V_{REF}=V_{DD}$. Each plot in Figure 6-28 and Figure 6-29 show both cases. Since V_{REF} set the voltage swing on the L=2 wires, the frequency and L=2 energy is roughly constant when V_{REF} is constant. As expected, when V_{DD} approaches $V_{REF}=0.4$ V, the data lines quickly converge to the case when $V_{REF}=V_{DD}$.

On the other hand, when V_{REF} (or equivalently V_{DD}) scales down, the energy also

118

decreases and reaches a minimum energy of approximately 74 fJ/cycle at 0.5 V supply. As the supply voltage is further decreased, leakage energy begins to dominate and increases the overall energy. At the minimum energy point, the L=2 wires dissipate 2.5x less energy than when $V_{REF}=V_{DD}=0.8$ V. If we consider the fact that the nominal supply voltage is 1.8 V in this process technology, the actual energy savings is likely much higher. Transistors have much lower current drives in subthreshold operation and this manifests in the slower frequency as V_{DD} decreases.

From Figure 6-29, when $V_{REF}=0.4$V, we observe that the system performs up to 2.11x faster and consumes 1.54x less (L=2 wire) energy when using the graphene wires over the M5 wires. The optimal design point is around $V_{REF}=0.4$ V and $V_{DD}=0.5$-0.6 V. Few-layer graphene wires are extremely thin and thus have low capacitance but very high resistances compared to similarly sized metal wires. When the graphene wires are driven by a lower voltage swing, the delay and the energy can be improved because the wire performance is dominated by the large resistances of the transistors and the wire capacitance [76]. This result was previously explained in detail in Section 2.3.3. At high supply voltages, the wire RC dominates the total delay and thus graphene wires have much slower delays. However, at low supply voltages, the large resistance of the transistors begins to dominate and the wire capacitance becomes important, resulting in a lower delay and energy for the graphene wire. We have demonstrated the potential of using few-layer graphene wires to increase performance in subthreshold circuits. The following section discusses the limitations of this work.

6.3.4   Limitations

Previously, we compared the frequency and L=2 energy of the system when graphene or M5 wires were enabled. A summary of the measured results is presented in Table 6.2. In subthreshold operation, graphene wires have many advantages over metal wires. However, this work serves as a technology demonstrator and the overall impact of using graphene is somewhat limited here.

When the system is actively running a benchmark, Figure 6-30 shows that slightly more than half the power is used for interconnects, including the clock network. The

Table 6.2: Summary of FPGA test chip.

Chip Summary		
Technology	0.18 μm CMOS	
Active Area	2.55 mm^2	
Graphene Wires	W=2 μm, L=300 μm 4-layers ($\sim$ 270 Ω/sq)	
TDC / Tester	INL	+0.59/-0.36 LSB
	DNL	+0.77/-0.52 LSB
	Energy	67.8 fJ/conv (@ 0.35V)
Min. Functional Point	300 mV	
Min. Energy Point	450 mV	
V$_{DD}$	**Energy/cycle**	**Frequency**
450 mV	8.7 pJ	53 kHz
500 mV	10.0 pJ	156 kHz
600 mV	13.8 pJ	800 kHz
Performance Improvement (graphene vs. reference M5) @ **V$_{REF}$=0.4V**		
Operating Frequency	2.11x faster	
(L=2 wire) Energy	1.54x lower	

L=2 wires only comprise a small portion ($<$ 3%) of the total energy. Therefore, although the graphene wires may have lower energy, the overall impact of that on the total energy is limited in this work. Furthermore, only half of the L=2 segments had working graphene devices in the example discussed in Section 6.3.3. Therefore, if the device yield improves and all active L=2 wires only use graphene, we expect the amount of energy savings to increase. In addition, the physical length of the L=2 wires was limited to 300 μm. If the wire length is prolonged, the capacitance difference between the graphene and M5 wires increases and thus the energy difference also increases. More importantly, we need to use graphene wires more extensively throughout the FPGA to have a big system-wide impact. In larger and more complex FPGAs, the global interconnect has been shown to dominate the delay and energy of the system [5,6,114]. Replacing the clock mesh or global wires with graphene presents an opportunity for future designs.
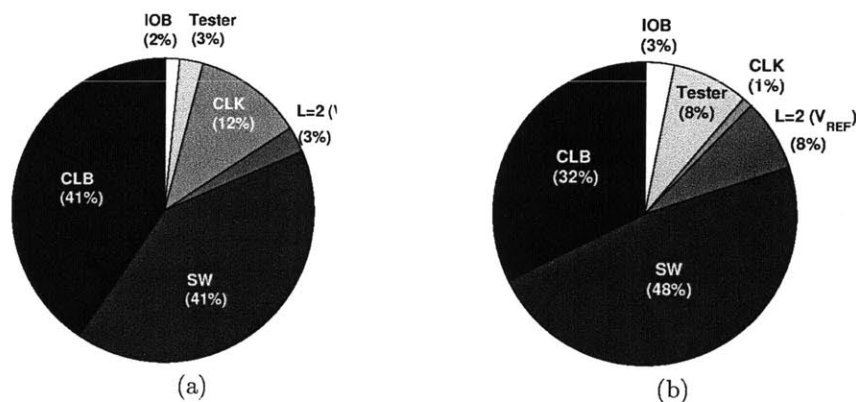
Figure 6-30: Breakdown of measured power consumption when system is (a) running a benchmark or (b) idle. V_{DD}=0.45 V.

6.4 Summary

We have demonstrated reliable signal routing on 4-layer graphene wires integrated on a custom-designed FPGA. Graphene sheets are synthesized using Cu foils and then integrated on top of the CMOS chip. The FPGA test chip implements a 5x5 logic array and includes a local tester at each L=2 wire segment to monitor the delay of each

graphene wire. The graphene wires have a (2.8x) lower capacitance than the metal wires, resulting in up to 2.11x faster speeds and 1.54x lower (L=2 wire) energy when driven by a low-swing voltage at 0.4V. This work demonstrates the potential of using low capacitance graphene wires for ultra-low power electronics. As larger systems are designed, we must use longer graphene wires and use such wires more extensively to achieve acceptable energy savings and justify the complexity of integrating graphene with CMOS.

Chapter 7

Conclusions

7.1 Summary of Results

In summary, this thesis characterized the performance of both multilayer and few-layer graphene interconnects and presented two prototype graphene/CMOS test chips.

First, a physics-based circuit model was used to compare the performance of graphene and Cu wires. When line-edge roughness is accounted for, graphene is expected to outperform Cu in terms of resistance when the line width falls below 10 nm. However, this requires very thick graphene films or graphene wires with very smooth edges. Alternatively, we can take advantage of the small capacitance of graphene wires. Roughly 10 or fewer number of graphene layers is needed to have 2x smaller capacitance than Cu wires. Although this results in large wire resistances, this is advantageous in subthreshold circuits where the large transistor resistance dominates.

We also investigated the growth of multilayer graphene wires fabricated from sheets grown by Ni CVD. The Ni CVD process results in an average thickness of 10-20 nm with a sheet resistance between 500-1000 Ω/sq. Maximum current densities up to 4×10^7 A/cm^2 have been measured in ambient air. The variation in the number of graphene layers directly affects the sheet resistance and current carrying capacity. This limits the performance and use of Ni-grown graphene sheets although they yield better manufacturing throughput than Cu CVD.

Next, we characterized the performance of these multilayer graphene wires when integrated with CMOS. Multilayer graphene sheets were integrated on a 0.35 μm CMOS chip and were fabricated into wires up to 1 mm in length. Reliable operation of end-to-end data communication on these graphene wires was demonstrated, achieving bit error rates below 2×10^{-10}. A low-swing signaling technique was used to achieve a transmitter energy of 0.3-0.7 pJ/bit/mm and a total energy efficiency of 2.4-5.2 pJ/bit/mm. Despite the high sheet resistivity of as-grown graphene ($> 600\Omega$/sq), integrated graphene links running at 50 Mbps was demonstrated. On-chip measurements were in good agreement with simulated results, and recent work shows promise as the reported sheet resistivity is an order of magnitude smaller than the graphene devices used in this work. Ultimately, cleaner processing steps and optimized CVD growth conditions are necessary to produce higher quality graphene films.

Finally, we demonstrated reliable signal routing on 4-layer graphene wires integrated on a custom-designed FPGA. Graphene sheets are synthesized using Cu foils and then integrated on top of a 0.18 μm CMOS chip. The FPGA test chip implements a 5x5 logic array and includes a local tester at each double-length (L=2) wire segment to monitor the delay of each graphene wire. The graphene wires have a 2.8x lower capacitance than the metal wires, resulting in up to 2.11x faster speeds and 1.54x lower (L=2 wire) energy when driven by a low-swing voltage at 0.4 V. This work presents the first graphene-based system application and demonstrates the potential of using low capacitance graphene wires for ultra-low power electronics.

7.2 Concluding Thoughts

Despite the concern of many, CMOS scaling has advanced far beyond the predicted barriers. Innovative solutions such as strain enhancers and high-K metal gates were developed to extend the transistor scaling roadmap. However, transistor dimensions are now approaching the atomic limit which poses an even greater threat. Similarly, copper resistivity increases at smaller line widths due to increased scattering effects. The problem is exacerbated as heat dissipation, leakage current, and thermal noise

become more difficult to cope with.

Nonetheless, CMOS technology is not likely to fade away any time soon. Many chip makers continue to make major investments in the next transistor generations and CMOS remains one of the most robust and scalable technologies. Most technologies that look beyond CMOS are premature and manufacturability is a big concern that is not adequately addressed. This is especially true for most nano-materials. While a more gradual transition towards nanotechnology is expected, this thesis has demonstrated that the merge between graphene and CMOS is not only feasible but can compete with CMOS in the near future for subthreshold systems.

To date, the most successful commercial applications of carbon-based materials have taken advantage of their mechanical or material properties. For example, carbon nanotubes are extremely strong and lightweight and have been widely used in sporting equipment such as hockey sticks, baseball bats, and bicycle frames [90]. While these applications are facilitated by bulk growth of carbon-based materials, integrated circuits require a much higher level of performance and reliability over millions of devices.

Thus, improving the manufacturing reliability is a very important area for future research. Although CVD methods generally yield large graphene sheets with excellent area coverage, preserving this throughout the integration process is a big challenge that deserves more attention especially for integrated circuit applications. Today, highly uniform graphene sheets can be grown in virtually any shape or size. However, the process of transferring graphene sheets to another substrate can cause the sheets to fold up or tear apart. Furthermore, the extremely small thickness of the film makes it very vulnerable throughout the subsequent fabrication steps. One way to improve this process is to reduce the number of transfers or completely eliminate the transfer step. Directly growing graphene on the CMOS chip is more desirable but back-end processes have lower thermal budgets. A possible solution is to develop a low-temperature process that uses Cu or Ni seed layers to directly grow graphene on the CMOS chip. Although this would result in a hybrid Cu/graphene or Ni/graphene stack, this method may improve both the performance and manufacturing reliability.

More importantly, a dielectric capping layer may be needed to physically protect the graphene sheets during the various process steps.

Another aspect that is not well understood is the issue of multilayer stacking. Growing more than 10 layers with high uniformity and performance is very difficult to demonstrate. Growing high-quality monolayer graphene films has already been demonstrated but extreme care is required when transferring to another substrate. Multilayer graphene films are more robust but no clear path exists to grow high-quality multilayer films. Furthermore, unless the stacking order can be precisely controlled, it will be difficult to prevent thick multilayer graphene from turning in to graphite. Understanding the exact nature of multilayer graphene sheets and synthesizing them is an important problem especially if one is interested in low resistance graphene wires at the local level. An alternative solution is to use insulating 2D sheets, such as boron nitride (BN), inserted in between high-quality few-layer graphene sheets. BN has a similar lattice structure as graphene. Experimental work has shown that graphene devices on BN have superior performance to graphene devices on SiO_2 [148]. Inserting BN not only increases the performance of graphene but allows an alternate route for fabricating high-quality and electrically decoupled multilayer graphene stacks.

Graphene also has other remarkable properties, namely high current capacity and thermal conductivity, which may prove to be another compelling reason for adopting graphene. Graphene and carbon nanotubes are known to be extremely stable and robust. Reliability concerns such as process variations and aging effects in silicon technology continues to be an increasing problem. Investigating the reliability limit of graphene in terms of current capacity and heat dissipation is an interesting topic for future research.

In addition to material innovation, identifying and developing graphene-based applications is also very important. The most important contribution of this thesis is that we have demonstrated a complete system using multiple graphene devices as low-capacitance wires. Even though the 4-layer graphene films presented here have a relatively high sheet resistance (~ 300 Ω/sq), we have demonstrated that using

few-layer graphene wires have smaller capacitance than metal wires and can be advantageous in energy-constrained systems. The FPGA test chip also includes local testers to monitor the quality of graphene wires. One limitation is that the process of determining the delay of each graphene wire and selecting one wire is performed off-chip and only once right before the bitstream is uploaded to the FPGA test chip. In future designs, this selection algorithm can be included in the chip to provide real-time device monitoring and adaptive switching based on the system workload. Given the wide variation of graphene devices when integrated on CMOS, developing adaptive control and monitoring methods is very important for such systems. Furthermore, another area of future research is to develop larger and more complex systems that rely more heavily on low-capacitance graphene wires at the global level. FPGAs are ideal because of their regular array structure and inherently interconnect-centric architecture. Other examples include processor/memory architectures or on-chip networks that require dense interconnect structures for cross-chip communication. This work demonstrated a prototype FPGA using 48 graphene devices. Developing larger systems is necessary to understand the full impact and benefits of using graphene devices for long global wires.

Ultimately, the focus of graphene research needs to extend beyond material synthesis and device performance in order to establish graphene as a viable interconnect technology. This work establishes a platform to evaluate graphene wires and provides an important step towards realizing a truly integrated graphene/CMOS system.

Bibliography

[1] M. T. Bohr, "Interconnect scaling-the real limiter to high performance ulsi," in *IEDM Tech. Dig.*, Dec 1995.

[2] J. Meindl, "Interconnect opportunities for gigascale integration," *IEEE Micro*, vol. 23, no. 3, pp. 28–35, May-June 2003.

[3] J. Meindl, R. Venkatesan, J. Davis, J. Joyner, A. Naeemi, P. Zarkesh-Ha, M. Bakir, T. Mule, P. Kohl, and K. Martin, "Interconnecting device opportunities for gigascale integration (gsi)," in *IEDM Tech. Dig.*, 2001, pp. 23.1.1–23.1.4.

[4] R. Ho, K. Mai, and M. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, no. 4, pp. 490–504, Apr. 2001.

[5] B. Calhoun, J. Ryan, S. Khanna, M. Putic, and J. Lach, "Flexible circuits and architectures for ultralow power," *Proc. IEEE*, vol. 98, no. 2, pp. 267–282, Feb. 2010.

[6] F. Li, Y. Lin, and L. He, "Vdd programmability to reduce fpga interconnect power," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, Nov. 2004, pp. 760–765.

[7] *International Technology Roadmap for Semiconductors (ITRS)*. Semicond. Ind. Assoc., 2007.

[8] U. Nawathe, M. Hassan, L. Warriner, K. Yen, B. Upputuri, D. Greenhill, A. Kumar, and H. Park, "An 8-core 64-thread 64b power-efficient sparc soc," in *ISSCC Tech. Dig.*, February 2007, pp. 108–590.

[9] D. Pham, T. Aipperspach, D. Boerstler, M. Bolliger, R. Chaudhry, D. Cox, P. Harvey, P. Harvey, H. Hofstee, C. Johns, J. Kahle, A. Kameyama, J. Keaty, Y. Masubuchi, M. Pham, J. Pille, S. Posluszny, M. Riley, D. Stasiak, M. Suzuoki, O. Takahashi, J. Warnock, S. Weitzel, D. Wendel, and K. Yazawa, "Overview of the architecture, circuit design, and physical implementation of a first-generation cell processor," *IEEE J. Solid-State Circuits*, vol. 41, no. 1, pp. 179–196, January 2006.

[10] T.-Y. Chiang, S. Souri, C. O. Chui, and K. Saraswat, "Thermal analysis of heterogeneous 3d ics with various integration scenarios," in *IEDM Tech. Dig.*, 2001, pp. 31.2.1–31.2.4.

[11] J. Joyner and J. Meindl, "Opportunities for reduced power dissipation using three-dimensional integration," in *Proc. IEEE Intl. Interconnect Technol. Conf.*, 2002, pp. 148–150.

[12] R. Chang, N. Talwalkar, C. Yue, and S. Wong, "Near speed-of-light signaling over on-chip electrical interconnects," *IEEE J. Solid-State Circuits*, vol. 38, no. 5, pp. 834–838, May 2003.

[13] A. Jose, G. Patounakis, and K. Shepard, "Pulsed current-mode signaling for nearly speed-of-light intrachip communication," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 772–780, April 2006.

[14] M. Flynn and J. Kang, "Global signaling over lossy transmission lines," in *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, November 2005, pp. 985–992.

[15] R. Ho, K. Mai, and M. Horowitz, "Efficient on-chip global interconnects," in *Symp. VLSI Circuits*, June 2003, pp. 271–274.

[16] R. Ho, I. Ono, F. Liu, R. Hopkins, A. Chow, J. Schauer, and R. Drost, "High-speed and low-energy capacitively-driven on-chip wires," in *ISSCC Tech. Dig.*, Feb. 2007.

[17] E. Mensink, D. Schinkel, E. Klumperink, E. van Tuijl, and B. Nauta, "A 0.28pj/b 2gb/s/ch transceiver in 90nm cmos for 10mm on-chip interconnects," in *ISSCC Tech. Dig.*, Feb. 2007.

[18] H. Zhang, V. George, and J. Rabaey, "Low-swing on-chip signaling techniques: effectiveness and robustness," *IEEE Trans. VLSI Syst.*, vol. 8, no. 3, pp. 264 –272, Jun. 2000.

[19] C. Chen, R. Parsa, N. Patil, S. Chong, K. Akarvardar, J. Provine, D. Lewis, J. Watt, R. T. Howe, H.-S. P. Wong, and S. Mitra, "Efficient fpgas using nano-electromechanical relays," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 2010, pp. 273 – 282.

[20] F. Chen, M. Spencer, R. Nathanael, C. Wang, H. Fariborzi, A. Gupta, H. Kam, V. Pott, J. Jeon, T.-J. K. Liu, D. Markovic, V. Stojanovic, and E. Alon, "Demonstration of integrated micro-electro-mechanical switch circuits for vlsi applications," in *ISSCC Tech. Dig.*, Feb. 2010, pp. 150 –151.

[21] Q.-T. Jiang, M.-H. Tsai, and R. H. Havemann, "Line width dependence of copper resistivity," in *Proc. IEEE Intl. Interconnect Technol. Conf.*, 2001.

[22] W. Steinhögl, G. Schindler, G. Steinlesberger, and M. Engelhardt, "Size-dependent resistivity of metallic wires in the mesoscopic range," *Phys. Rev. B*, vol. 66, no. 7, p. 075414, Aug 2002.

[23] C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud, D. Mayou, T. Li, J. Hass, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. De Heer, "Electronic confinement and coherence in patterned epitaxial graphene," *Science*, vol. 312, no. 5777, pp. 1191 – 1196, 2006.

[24] X. Du, I. Skachko, A. Barker, and E. Y. Andrei, "Approaching ballistic transport in suspended graphene," *Nature Nanotech.*, vol. 3, no. 8, pp. 491–495, 2008.

[25] K. Bolotin, K. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. Stormer, "Ultrahigh electron mobility in suspended graphene," *Solid State Communications*, vol. 146, no. 9-10, pp. 351 – 355, 2008.

[26] A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. Lau, "Superior thermal conductivity of single-layer graphene," *Nano Lett.*, vol. 8, no. 3, pp. 902–907, 2008.

[27] Q. Shao, G. Liu, D. Teweldebrhan, and A. A. Balandin, "High-temperature quenching of electrical resistance in graphene interconnects," *Appl. Phys. Lett.*, vol. 92, no. 20, p. 202108, 2008.

[28] R. Murali, Y. Yang, K. Brenner, T. Beck, and J. D. Meindl, "Breakdown current density of graphene nanoribbons," *Appl. Phys. Lett.*, vol. 94, no. 24, p. 243114, 2009.

[29] T. Yu, E.-K. Lee, B. Briggs, B. Nagabhirava, and B. Yu, "Bilayer graphene system: Current-induced reliability limit," *IEEE Electron Device Lett.*, vol. 31, no. 10, pp. 1155 –1157, Oct. 2010.

[30] K.-J. Lee, A. P. Chandrakasan, and J. Kong, "Breakdown current density of cvd-grown multilayer graphene interconnects," *IEEE Electron Device Lett.*, vol. 32, no. 4, pp. 557 –559, Apr. 2011.

[31] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, "Energy band-gap engineering of graphene nanoribbons," *Phys. Rev. Lett.*, vol. 98, no. 20, p. 206805, May 2007.

[32] X. Wang, Y. Ouyang, X. Li, H. Wang, J. Guo, and H. Dai, "Room-temperature all-semiconducting sub-10-nm graphene nanoribbon field-effect transistors," *Phys. Rev. Lett.*, vol. 100, no. 20, p. 206803, May 2008.

[33] A. Naeemi and J. Meindl, "Conductance modeling for graphene nanoribbon (gnr) interconnects," *IEEE Electron Device Lett.*, vol. 28, no. 5, pp. 428–431, May 2007.

[34] R. Murali, K. Brenner, Y. Yang, T. Beck, and J. Meindl, "Resistivity of graphene nanoribbon interconnects," *IEEE Electron Device Lett.*, vol. 30, no. 6, pp. 611–613, June 2009.

[35] S. Ghosh, I. Calizo, D. Teweldebrhan, E. P. Pokatilov, D. L. Nika, A. A. Balandin, W. Bao, F. Miao, and C. N. Lau, "Extremely high thermal conductivity of graphene: Prospects for thermal management applications in nanoelectronic circuits," *Appl. Phys. Lett.*, vol. 92, no. 15, p. 151911, 2008.

[36] K. S. Kim, Y. Zhao, H. Jang, S. Y. Lee, J. M. Kim, K. S. Kim, J.-H. Ahn, P. Kim, J.-Y. Choi, and B. H. Hong, "Large-scale pattern growth of graphene films for stretchable transparent electrodes," *Nature*, vol. 457, pp. 706–710, Feb. 2009.

[37] A. Reina, X. Jia, J. Ho, D. Nezich, H. Son, V. Bulovic, M. S. Dresselhaus, and J. Kong, "Large area, few-layer graphene films on arbitrary substrates by chemical vapor deposition," *Nano Lett.*, vol. 9, no. 1, pp. 30–35, 2009.

[38] X. Chen, K.-J. Lee, D. Akinwande, G. F. Close, S. Yasuda, B. Paul, S. Fujita, J. Kong, and H.-S. P. Wong, "High-speed graphene interconnect monolithically integrated with cmos ring oscillators operating at 1.3 ghz," in *IEDM Tech. Dig.*, 2009.

[39] J. Kedzierski, P.-L. Hsu, A. Reina, J. Kong, P. Healey, P. Wyatt, and C. Keast, "Graphene-on-insulator transistors made using c on ni chemical-vapor deposition," *IEEE Electron Device Lett.*, vol. 30, no. 7, pp. 745 –747, July 2009.

[40] H. Wang, A. Hsu, J. Wu, J. Kong, and T. Palacios, "Graphene-based ambipolar rf mixers," *IEEE Electron Device Lett.*, vol. 31, no. 9, pp. 906 –908, sept. 2010.

[41] Y. Lee, S. Bae, H. Jang, S. Jang, S.-E. Zhu, S. H. Sim, Y. I. Song, B. H. Hong, and J.-H. Ahn, "Wafer-scale synthesis and transfer of graphene films," *Nano Lett.*, vol. 10, no. 2, pp. 490–493, 2010.

[42] X. Li, W. Cai, J. An, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo, and R. S. Ruoff, "Large-Area Synthesis of High-Quality and Uniform Graphene Films on Copper Foils," *Science*, vol. 324, no. 5932, pp. 1312–1314, 2009.

[43] Y. Yang and R. Murali, "Impact of size effect on graphene nanoribbon transport," *IEEE Electron Device Lett.*, vol. 31, no. 3, pp. 237–239, Mar. 2010.

[44] A. Naeemi and J. Meindl, "Compact physics-based circuit models for graphene nanoribbon interconnects," *IEEE Trans. Electron Devices*, vol. 56, no. 9, pp. 1822 –1833, Sept. 2009.

[45] A. Naeemi and J. D. Meindl, "Performance benchmarking for graphene nanoribbon, carbon nanotube, and cu interconnects," in *Proc. IEEE Intl. Interconnect Technol. Conf.*, June 2008, pp. 183–185.

132

[46] P. Burke, "Luttinger liquid theory as a model of the gigahertz electrical properties of carbon nanotubes," *IEEE Trans. Nanotechnol.*, vol. 1, no. 3, pp. 129 – 144, sep 2002.

[47] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, "Energy band-gap engineering of graphene nanoribbons," *Phys. Rev. Lett.*, vol. 98, no. 20, p. 206805, May 2007.

[48] H. M. Wang, Y. H. Wu, Z. H. Ni, and Z. X. Shen, "Electronic transport and layer engineering in multilayer graphene structures," *Appl. Phys. Lett.*, vol. 92, no. 5, p. 053504, 2008.

[49] S. Datta, *Quantum transport: atom to transistor.* Cambridge Univ Press, 2005.

[50] Y.-W. Son, M. L. Cohen, and S. G. Louie, "Energy gaps in graphene nanoribbons," *Phys. Rev. Lett.*, vol. 97, no. 21, p. 216803, Nov 2006.

[51] X. Li, X. Wang, L. Zhang, S. Lee, and H. Dai, "Chemically derived, ultrasmooth graphene nanoribbon semiconductors," *Science*, vol. 319, no. 5867, pp. 1229–1232, 2008.

[52] C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud, D. Mayou, T. Li, J. Hass, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. de Heer, "Electronic confinement and coherence in patterned epitaxial graphene," *Science*, vol. 312, no. 5777, pp. 1191–1196, 2006.

[53] P. Avouris, Z. Chen, and V. Perebeinos, "Carbon-based electronics," *Nature Nanotech.*, vol. 2, no. 10, pp. 605–615, 2007.

[54] S. Y. Zhou, G. H. Gweon, A. Fedorov, P. N. First, W. A. de Heer, D.-H. Lee, F. Guinea, A. H. Castro Neto, and A. Lanzara, "Substrate-induced bandgap opening in epitaxial graphene," *Nature Materials*, vol. 6, no. 10, pp. 770–775, 2007.

[55] S. Bae, H. Kim, Y. Lee, X. Xu, J. Park, Y. Zheng, J. Balakrishnan, T. Lei, H. Kim, Y. Song, Y.-J. Kim, K. Kim, B. Ozyilmax, J.-H. Ahn, H. Byung, and S. Iijima, "Roll-to-roll production of 30-inch graphene films for transparent electrodes," *Nature Nanotech.*, vol. 5, no. 8, pp. 574–578, 2010.

[56] L. X. Benedict, V. H. Crespi, S. G. Louie, and M. L. Cohen, "Static conductivity and superconductivity of carbon nanotubes: Relations between tubes and sheets," *Phys. Rev. B*, vol. 52, no. 20, pp. 14 935–14 940, Nov. 1995.

[57] A. AlZahrani and G. Srivastava, "Gradual changes in electronic properties from graphene to graphite: first-principles calculations," *J. Phys.: Conden. Matter*, vol. 21, p. 495503, 2009.

[58] P. N. Nirmalraj, T. Lutz, S. Kumar, G. S. Duesberg, and J. J. Boland, "Nanoscale mapping of electrical resistivity and connectivity in graphene strips and networks," *Nano Lett.*, vol. 11, no. 1, pp. 16–22, 2011.

[59] C. Faugeras, A. Nerrière, M. Potemski, A. Mahmood, E. Dujardin, C. Berger, and W. A. de Heer, "Few-layer graphene on sic, pyrolitic graphite, and graphene: A raman scattering study," *Appl. Phys. Lett.*, vol. 92, no. 1, p. 011914, 2008.

[60] C. Xu, H. Li, and K. Banerjee, "Modeling, analysis, and design of graphene nano-ribbon interconnects," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1567 –1578, Aug. 2009.

[61] S. Tanachutiwat and W. Wang, "Exploring multi-layer graphene nanoribbon interconnects," in *Nano-Net*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, O. Akan, P. Bellavista, J. Cao, F. Dressler, D. Ferrari, M. Gerla, H. Kobayashi, S. Palazzo, S. Sahni, X. S. Shen, M. Stan, J. Xiaohua, A. Zomaya, G. Coulson, and M. Cheng, Eds. Springer Berlin Heidelberg, 2009, vol. 3, pp. 49–53.

[62] L. Ci, Z. Xu, L. Wang, W. Gao, F. Ding, K. Kelly, B. Yakobson, and P. Ajayan, "Controlled nanocutting of graphene," *Nano Research*, vol. 1, pp. 116–122, 2008.

[63] D. Gunlycke, H. M. Lawler, and C. T. White, "Room-temperature ballistic transport in narrow graphene strips," *Phys. Rev. B*, vol. 75, no. 8, p. 085418, Feb 2007.

[64] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, no. 5696, pp. 666–669, 2004.

[65] Y.-W. Tan, Y. Zhang, K. Bolotin, Y. Zhao, S. Adam, E. H. Hwang, S. Das Sarma, H. L. Stormer, and P. Kim, "Measurement of scattering rate and minimum conductivity in graphene," *Phys. Rev. Lett.*, vol. 99, no. 24, p. 246803, Dec 2007.

[66] Y. Zhang, Y. Tan, H. Stormer, and P. Kim, "Experimental observation of the quantum hall effect and berry's phase in graphene," *Nature*, vol. 438, no. 7065, pp. 201–204, 2005.

[67] J.-H. Chen, C. Jang, S. Xiao, M. Ishigami, and M. S. Fuhrer, "Intrinsic and extrinsic performance limits of graphene devices on sio2," *Nature Nanotech.*, vol. 3, no. 4, pp. 206–209, 2008.

[68] *International Technology Roadmap for Semiconductors (ITRS).* Semicond. Ind. Assoc., 2010.

[69] W. Steinhoegl, G. Schindler, G. Steinlesberger, M. Traving, and M. Engelhardt, "Scaling laws for the resistivity increase of sub-100 nm interconnects," in *Simulation of Semiconductor Processes and Devices, International Conference on*, sept. 2003, pp. 27 – 30.

[70] E. Sondheimer, "The mean free path of electrons in metals," *Advances in Physics*, vol. 1, no. 1, pp. 1–42, 1952.

[71] A. F. Mayadas, M. Shatzkes, and J. F. Janak, "Electrical resistivity model for polycrystalline films: The case of specular reflection at external surfaces," *Appl. Phys. Lett.*, vol. 14, no. 11, pp. 345–347, 1969.

[72] G. Lopez, J. Davis, and J. Meindl, "A new physical model and experimental measurements of copper interconnect resistivity considering size effects and line-edge roughness (ler)," in *Proc. IEEE Intl. Interconnect Technol. Conf.*, June 2009, pp. 231 –234.

[73] "Mosis wafer electrical test data." [Online]. Available: http://mosis.com/test/

[74] S. De and J. N. Coleman, "Are there fundamental limitations on the sheet resistance and transmittance of thin graphene films?" *ACS Nano*, vol. 4, no. 5, pp. 2713–2720, 2010.

[75] C.-H. Jan, M. Agostinelli, M. Buehler, Z.-P. Chen, S.-J. Choi, G. Curello, H. Deshpande, S. Gannavaram, W. Hafez, U. Jalan, M. Kang, P. Kolar, K. Komeyli, B. Landau, A. Lake, N. Lazo, S.-H. Lee, T. Leo, J. Lin, N. Lindert, S. Ma, L. McGill, C. Meining, A. Paliwal, J. Park, K. Phoa, I. Post, N. Pradhan, M. Prince, A. Rahman, J. Rizk, L. Rockford, G. Sacks, A. Schmitz, H. Tashiro, C. Tsai, P. Vandervoorn, J. Xu, L. Yang, J.-Y. Yeh, J. Yip, K. Zhang, Y. Zhang, and P. Bai, "A 32nm soc platform technology with 2nd generation high-k/metal gate transistors optimized for ultra low power, high performance, and high density product applications," in *IEDM Tech. Dig.*, dec. 2009, pp. 1 –4.

[76] O. Jamal and A. Naeemi, "Evolutionary and revolutionary interconnect technologies for performance enhancement of subthreshold circuits," in *Proc. IEEE Intl. Interconnect Technol. Conf.*, June 2010, pp. 1 –3.

[77] J. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital integrated circuits*. Prentice Hall, 2002.

[78] "Predictive technology model." [Online]. Available: http://ptm.asu.edu

[79] H. P. Boehm, A. Clauss, G. O. Fischer, and U. Hofmann, "Das adsorptionsverhalten sehr dunner kohlenstoff-folien," *Zeitschrift fur anorganische und allgemeine Chemie*, vol. 316, no. 3-4, pp. 119–127, 1962.

[80] H. Boehm, R. Setton, and E. Stumpp, "Nomenclature and terminology of graphite intercalation compounds. report by a subgroup of the international

committee for characterization and terminology of carbon and graphite on suggestions for rules for the nomenclature and terminology of graphite intercalation compounds," *Synthetic metals*, vol. 11, no. 6, pp. 363–371, 1985.

[81] C. Berger, Z. Song, T. Li, X. Li, A. Y. Ogbazghi, R. Feng, Z. Dai, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. de Heer, "Ultrathin epitaxial graphite: 2d electron gas properties and a route toward graphene-based nanoelectronics," *J. Phys. Chem. B*, vol. 108, no. 52, pp. 19912–19916, 2004.

[82] H. Yamaguchi, G. Eda, C. Mattevi, H. Kim, and M. Chhowalla, "Highly uniform 300 mm Wafer-Scale deposition of single and multilayered chemically derived graphene thin films," *ACS Nano*, vol. 4, no. 1, pp. 524–528, 2010.

[83] K. K. Kim, A. Reina, Y. Shi, H. Park, L.-J. Li, Y. H. Lee, and J. Kong, "Enhancing the conductivity of transparent graphene films via doping," *Nanotechnology*, vol. 21, no. 28, p. 285205, 2010.

[84] J. Krupka and W. Strupinski, "Measurements of the sheet resistance and conductivity of thin epitaxial graphene and sic films," *Appl. Phys. Lett.*, vol. 96, no. 8, p. 082101, 2010.

[85] A. Reina, S. Thiele, X. Jia, S. Bhaviripudi, M. Dresselhaus, J. Schaefer, and J. Kong, "Growth of large-area single-and Bi-layer graphene by controlled carbon precipitation on polycrystalline Ni surfaces," *Nano Research*, vol. 2, no. 6, pp. 509–516, 2009.

[86] C. Xu, H. Li, and K. Banerjee, "Graphene nano-ribbon (gnr) interconnects: A genuine contender or a delusive dream?" in *IEDM Tech. Dig.*, Dec. 2008.

[87] P. Avouris, J. Appenzeller, R. Martel, and S. J. Wind, "Carbon nanotube electronics," *Proc. of the IEEE*, vol. 91, no. 11, pp. 1772 – 1783, 2003. [Online]. Available: http://dx.doi.org/10.1109/JPROC.2003.818338

[88] P. McEuen, M. Fuhrer, and H. Park, "Single-walled carbon nanotube electronics," *IEEE Trans. on Nanotechnology*, vol. 1, no. 1, pp. 78 – 85, 2002. [Online]. Available: http://dx.doi.org/10.1109/TNANO.2002.1005429

[89] H.-S. Wong, J. Appenzeller, V. Derycke, R. Martel, S. Wind, and P. Avouris, "Carbon nanotube field effect transistors - fabrication, device physics, and circuit implications," *ISSCC Tech. Dig.*, pp. 367 – 370, 2003. [Online]. Available: http://dx.doi.org/10.1109/ISSCC.2003.1234339

[90] "Zyvex technologies." [Online]. Available: http://zyvexpro.com/

[91] "Amroy europe ltd." [Online]. Available: http://www.amroy.fi/

[92] "Nanomix inc." [Online]. Available: http://www.nano.com/

[93] "Eikos inc." [Online]. Available: http://www.eikos.com/

[94] G. F. Close, S. Yasuda, B. Paul, S. Fujita, and H.-S. P. Wong, "A 1 ghz integrated circuit with carbon nanotube interconnects and silicon transistors," *Nano Lett.*, vol. 8, no. 2, pp. 706–709, 2008.

[95] X. Chen, D. Akinwande, K.-J. Lee, G. F. Close, S. Yasuda, B. C. Paul, S. Fujita, J. Kong, and H.-S. P. Wong, "Fully integrated graphene and carbon nanotube interconnects for gigahertz high-speed cmos electronics," *IEEE Trans. Electron Devices*, vol. 57, no. 11, pp. 3137 –3143, Nov. 2010.

[96] K.-J. Lee, M. Qazi, J. Kong, and A. P. Chandrakasan, "Low-swing signaling on monolithically integrated global graphene interconnects," *IEEE Trans. Electron Devices*, vol. 57, no. 12, pp. 3418 –3425, Dec. 2010.

[97] N. Liu, L. Fu, B. Dai, K. Yan, X. Liu, R. Zhao, Y. Zhang, and Z. Liu, "Universal segregation growth approach to wafer-size graphene from non-noble metals," *Nano Lett.*, vol. 11, no. 1, pp. 297–303, 2011.

[98] S. Lee, K. Lee, and Z. Zhong, "Wafer scale homogeneous bilayer graphene films by chemical vapor deposition," *Nano Lett.*, vol. 10, no. 11, pp. 4702–4707, 2010.

[99] Y. Y. Wang, Z. H. Ni, Z. X. Shen, H. M. Wang, and Y. H. Wu, "Interference enhancement of raman signal of graphene," *Appl. Phys. Lett.*, vol. 92, no. 4, p. 043121, 2008.

[100] I. Jung, M. Pelton, R. Piner, D. Dikin, S. Stankovich, S. Watcharotone, M. Hausner, and R. Ruoff, "Simple approach for high-contrast optical imaging and characterization of graphene-based sheets," *Nano Lett.*, vol. 7, no. 12, pp. 3569–3575, 2007.

[101] Z. Ni, H. Wang, J. Kasim, H. Fan, T. Yu, Y. Wu, Y. Feng, and Z. Shen, "Graphene thickness determination using reflection and contrast spectroscopy," *Nano Lett.*, vol. 7, no. 9, pp. 2758–2763, 2007.

[102] P. E. Gaskell, H. S. Skulason, C. Rodenchuk, and T. Szkopek, "Counting graphene layers on glass via optical reflection microscopy," *Appl. Phys. Lett.*, vol. 94, no. 14, p. 143101, 2009.

[103] S. Roddaro, P. Pingue, V. Piazza, V. Pellegrini, and F. Beltram, "The optical visibility of graphene: interference colors of ultrathin graphite on sio2," *Nano Lett.*, vol. 7, no. 9, pp. 2707–2710, 2007.

[104] P. Blake, E. W. Hill, A. H. C. Neto, K. S. Novoselov, D. Jiang, R. Yang, T. J. Booth, and A. K. Geim, "Making graphene visible," *Appl. Phys. Lett.*, vol. 91, no. 6, p. 063124, 2007.

[105] M. Suzuki, Y. Ominami, Q. Ngo, C. Y. Yang, A. M. Cassell, and J. Li, "Current-induced breakdown of carbon nanofibers," *J. Appl. Phys.*, vol. 101, no. 11, p. 114307, 2007.

[106] J. Y. Huang, S. Chen, S. H. Jo, Z. Wang, D. X. Han, G. Chen, M. S. Dresselhaus, and Z. F. Ren, "Atomic-scale imaging of wall-by-wall breakdown and concurrent transport measurements in multiwall carbon nanotubes," *Phys. Rev. Lett.*, vol. 94, no. 23, p. 236802, Jun 2005.

[107] H. Kitsuki, T. Yamada, D. Fabris, J. R. Jameson, P. Wilhite, M. Suzuki, and C. Y. Yang, "Length dependence of current-induced breakdown in carbon nanofiber interconnects," *Appl. Phys. Lett.*, vol. 92, no. 17, p. 173110, 2008.

[108] P.-C. Wang and R. G. Filippi, "Electromigration threshold in copper interconnects," *Appl. Phys. Lett.*, vol. 78, no. 23, pp. 3598–3600, 2001.

[109] Z. Liu, A. A. Bol, and W. Haensch, "Large-scale graphene transistors with enhanced performance and reliability based on interface engineering by phenylsilane self-assembled monolayers," *Nano Lett.*, vol. 11, no. 2, pp. 523–528, 2011.

[110] R. Ho, T. Ono, R. Hopkins, A. Chow, J. Schauer, F. Liu, and R. Drost, "High speed and low energy capacitively driven on-chip wires," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 52–60, Jan. 2008.

[111] H. Zhang, V. Prabhu, V. George, M. Wan, M. Benes, A. Abnous, and J. Rabaey, "A 1-v heterogeneous reconfigurable dsp ic for wireless baseband digital signal processing," *IEEE J. Solid-State Circuits*, vol. 35, no. 11, pp. 1697 –1704, nov 2000.

[112] J. Rabaey, A. Abnous, Y. Ichikawa, K. Seno, and M. Wan, "Heterogeneous reconfigurable systems," in *Signal Processing Systems, 1997. SIPS 97 - Design and Implementation., 1997 IEEE Workshop on*, nov 1997, pp. 24 –34.

[113] I. Kuon, R. Tessier, and J. Rose, "Fpga architecture: Survey and challenges," *Found. Trends Electron. Des. Autom.*, vol. 2, no. 2, pp. 135–253, 2008.

[114] J. Ryan and B. Calhoun, "A sub-threshold fpga with low-swing dual-vdd interconnect in 90nm cmos," in *CICC Tech. Dig.*, Sept. 2010, pp. 1–4.

[115] "Xilinx virtex datasheet." [Online]. Available: http://www.xilinx.com/support/documentation/data_sheets/ds003.pdf

[116] "Xilinx virtex-ii datasheet." [Online]. Available: http://www.xilinx.com/support/documentation/data_sheets/ds031.pdf

[117] E. Ahmed and J. Rose, "The effect of lut and cluster size on deep-submicron fpga performance and density," in *Proc. ACM/SIGDA Intl. Symp. FPGA*. New York, NY, USA: ACM, 2000, pp. 3–12.

[118] ——, "The effect of lut and cluster size on deep-submicron fpga performance and density," *IEEE Trans. VLSI Syst.*, vol. 12, no. 3, pp. 288 – 298, Mar. 2004.

[119] I. Kuon and J. Rose, "Area and delay trade-offs in the circuit and architecture design of fpgas," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 2008, pp. 149–158.

[120] H. Kaul, M. Anders, S. Mathew, S. Hsu, A. Agarwal, R. Krishnamurthy, and S. Borkar, "A 320 mv 56 μw 411 gops/watt ultra-low voltage motion estimation accelerator in 65 nm cmos," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 107 –114, jan. 2009.

[121] S. Hanson, M. Seok, Y.-S. Lin, Z. Y. Foo, D. Kim, Y. Lee, N. Liu, D. Sylvester, and D. Blaauw, "A low-voltage processor for sensing applications with picowatt standby mode," *IEEE J. Solid-State Circuits*, vol. 44, no. 4, pp. 1145 –1155, april 2009.

[122] J. Kwong, Y. Ramadass, N. Verma, and A. Chandrakasan, "A 65 nm sub-v_t microcontroller with integrated sram and switched capacitor dc-dc converter," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 115 –126, Jan. 2009.

[123] Y. Pu, J. Pineda de Gyvez, H. Corporaal, and Y. Ha, "An ultra-low-energy multi-standard jpeg co-processor in 65 nm cmos with sub/near threshold supply voltage," *IEEE J. Solid-State Circuits*, vol. 45, no. 3, pp. 668 –680, march 2010.

[124] B. Zhai, S. Pant, L. Nazhandali, S. Hanson, J. Olson, A. Reeves, M. Minuth, R. Helfand, T. Austin, D. Sylvester, and D. Blaauw, "Energy-efficient subthreshold processor design," *IEEE Trans. VLSI Syst.*, vol. 17, no. 8, pp. 1127 –1137, aug. 2009.

[125] B. Calhoun, A. Wang, and A. Chandrakasan, "Modeling and sizing for minimum energy operation in subthreshold circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 9, pp. 1778 – 1786, sept. 2005.

[126] D. Lewis, V. Betz, D. Jefferson, A. Lee, C. Lane, P. Leventis, S. Marquardt, C. McClintock, B. Pedersen, G. Powell, S. Reddy, C. Wysocki, R. Cliff, and J. Rose, "The stratix routing and logic architecture," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 2003, pp. 12–20.

[127] G. Lemieux, E. Lee, M. Tom, and A. Yu, "Directional and single-driver wires in fpga interconnect," in *Proc. IEEE Intl. Conf. FPT*, 2004, pp. 41 – 48.

[128] S. J. Wilton, "Architecture and algorithms for field-programmable gate arrays with embedded memory," Ph.D. dissertation, University of Toronto, 1997.

[129] V. George, H. Zhang, and J. Rabaey, "The design of a low energy fpga," in *Proc. ISLPED*, 1999, pp. 188 – 193.

[130] I. J. Chang, J.-J. Kim, K. Kim, and K. Roy, "Robust level converter for subthreshold/super-threshold operation:100 mv to 2.5 v," *IEEE Trans. VLSI Syst.*, vol. PP, no. 99, pp. 1 –9, 2010.

[131] S. Wooters, B. Calhoun, and T. Blalock, "An energy-efficient subthreshold level converter in 130-nm cmos," *IEEE Trans. Circuits Syst. II*, vol. 57, no. 4, pp. 290 –294, April 2010.

[132] H. Shao and C.-Y. Tsui, "A robust, input voltage adaptive and low energy consumption level converter for sub-threshold logic," in *Proc. European Solid State Circ. Conf.*, Sept. 2007, pp. 312 –315.

[133] G. Roberts and M. Ali-Bakhshian, "A brief introduction to time-to-digital and digital-to-time converters," *IEEE Trans. Circuits Syst. II*, vol. 57, no. 3, pp. 153 –157, march 2010.

[134] P. Chen, C.-C. Chen, C.-C. Tsai, and W.-F. Lu, "A time-to-digital-converter-based cmos smart temperature sensor," *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1642 – 1648, aug. 2005.

[135] M. Lee and A. Abidi, "A 9 b, 1.25 ps resolution coarse-fine time-to-digital converter in 90 nm cmos that amplifies a time residue," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 769 –777, april 2008.

[136] P. Dudek, S. Szczepanski, and J. Hatfield, "A high-resolution cmos time-to-digital converter utilizing a vernier delay line," *IEEE J. Solid-State Circuits*, vol. 35, no. 2, pp. 240 –247, feb 2000.

[137] "IEEE standard for verilog hardware description language," *IEEE Std 1364-2005 (Revision of IEEE Std 1364-2001)*, pp. 0_1 –560, 2006.

[138] J. Luu, I. Kuon, P. Jamieson, T. Campbell, A. Ye, W. M. Fang, and J. Rose, "Vpr 5.0: Fpga cad and architecture exploration tools with single-driver routing, heterogeneity and process scaling," in *Proc. ACM/SIGDA Intl. Symp. FPGA.* New York, NY, USA: ACM, 2009, pp. 133–142.

[139] V. Betz and J. Rose, "Vpr: A new packing, placement and routing tool for fpga research," in *Proc. 7th International Workshop on Field-Programmable Logic and Applications*, 1997, pp. 213–222.

[140] V. Betz, J. Rose, and A. Marquardt, *Architecture and CAD for deep-submicron FPGAs.* Kluwer Academic Publishers Norwell, MA, USA, 1999.

[141] P. Jamieson and J. Rose, "A verilog rtl synthesis tool for heterogeneous fpgas," in *Field Programmable Logic and Applications, International Conference on*, Aug. 2005, pp. 305 – 310.

[142] J. Pistorius, M. Hutton, A. Mischchenko, and R. Brayton, "Benchmarking method and designs targeting logic synthesis for fpgas," in *Proc. IWLS*, 2007, pp. 230–237.

[143] "Berkeley logic synthesis and verification group, abc: A system for sequential synthesis and verification, release 70731."

[144] V. Betz and J. Rose, "Cluster-based logic blocks for fpgas: area-efficiency vs. input sharing and size," in *CICC Tech. Dig.*, may 1997, pp. 551 –554.

[145] A. S. Marquardt, V. Betz, and J. Rose, "Using cluster-based logic blocks and timing-driven packing to improve fpga speed and density," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 1999, pp. 37–46.

[146] A. Marquardt, V. Betz, and J. Rose, "Timing-driven placement for fpgas," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 2000, pp. 203–213.

[147] V. Betz and J. Rose, "Automatic generation of fpga routing architectures from high-level descriptions," in *Proc. ACM/SIGDA Intl. Symp. FPGA*, 2000, pp. 175–184.

[148] C. Dean, A. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. Shepard *et al.*, "Boron nitride substrates for high-quality graphene electronics," *Nature Nanotech.*, vol. 5, no. 10, pp. 722–726, 2010.