Fabrication and Characterization of Silicon Nanowire Field-Effect Sensors

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
David Aaron Routenberg
Dissertation Director: Prof. Mark A. Reed
December 2009
Soli Deo Gloria
Acknowledgments

Most successful research efforts require significant help from a lot of people. I spent the last 6 years learning that lesson despite my best efforts to deny the truth of it. Even when I was imagining myself to be “going it alone” there were a number of people whose support was instrumental to keeping me heading toward the goal of producing this dissertation. Some of these people contributed directly to the science and engineering, while other helped me to properly place my experiences in lab within the larger context of my life. These were the people who kept me from throwing up my hands in despair and abandoning the lab for seemingly greener pastures.

I want to thank my advisor, Mark, for his guidance and support, which began several years before I came to graduate school. He has graciously put up with me for far too many years and has given me the freedom to explore a lot of territory. His example has shown me how to actually do good science, and not just make pretty figures for papers though hopefully the two coincide every now and then.

I spent more hours than I care to remember in cleanroom facilities over the past 6 years. In particular I want to thank the staff members who made my work possible and even enjoyable at times. The staff at the CNF is second to none and I’d like to thank them as a whole before acknowledging some who were particularly instrumental to this work. Thank you all! As for the individuals, Rob Ilic is truly unique and has inspired me with both his love of science and his excellent guitar chops. In addition to Rob, Daron Westly, Meredith Metzler, Mike Skvarla, Paul Pelletier and Jerry Drumheller all helped me significantly and got me out of some serious jams. Here at Yale I’ve had the pleasure of working with Mike Power, Chris Tillinghast and Jim Agresta, each of whom has made
my time in the cleanroom more productive and more enjoyable. Thanks for keeping it running and letting me do all kinds of crazy stuff in there.

Anyone who has worked in an office knows that the administrative staff have all the real power. Thanks Vivian, Lisa and Arlene for all your help, for feeding me, and for talking me down on a regular basis. You all made Becton a much happier place for me.

Our research group has been an odd collection of personalities over the years. I think I have learned a bit from each person in the group and hope that I have left each of them with something useful as well. In no particular order, thanks Alek, Nitin, Eric, Stan, Ryan, Jin, Wenyong, Weihua, Marleen, Menno, Takhee, and Guosheng.

Now I can take a few paragraphs to thank the family and friends who put up with me and my lab-induced melancholy for so long and made it possible not only for me to finish grad school, but to even get here in the first place. These people encouraged, loved, cajoled, and disciplined me into the person, student, and researcher that I am today.

Thanks Mom and Dad for teaching me to love learning and to never be finished with it. Thank you both for all you have done to get me here, even when I didn’t like it. Kerry and Chris, you showed me that it could be done and encouraged me when it didn’t come easily. I wouldn’t be in this position without your example and support. Peter and Robyn, thanks for your raising your daughter and for treating me like your son. Caroline, thanks for many years of laughter.

To my teachers and professors, thanks for 24 years of teaching me to ask questions. Thanks to Dr. James Ellenbogen for taking a chance on me even though I didn’t quite fit the mold. I wouldn’t have made it to Yale without you.
Friends at Christ Presbyterian, you have made New Haven feel like home. I will miss you all dearly and will see you all again sooner or later. Molly and Peter, Jake and Hillary you all know what it is like to slog through a PhD. Thanks for making the trek a little easier for me.

My final words of thanks are reserved for my wife and best friend, Katherine, who has been patiently waiting for this day to come. Because of her encouragement, love, forgiveness, patience I made it through without too many scars. I love you and while I’m no longer a student, I hope to study you for the rest of my life.
Contents

List of Figures ........................................................................................................................................... x
List of Tables ........................................................................................................................................... xviii

Chapter 1: Introduction ............................................................................................................................ 1
1.1 Motivation ........................................................................................................................................... 1
1.2 Historical Perspective .................................................................................................................... 4
1.3 Overview of Work Presented ...................................................................................................... 5

Chapter 2: Theoretical Background ..................................................................................................... 8
2.1 ISFET Theory of Operation .......................................................................................................... 8
2.2 Scaling from ISFETs to Nanowires ........................................................................................ 15
2.3 Chemical Binding and Fluidics Considerations ........................................................................... 22

Chapter 3: Silicon Nanowire FETs Design and Fabrication ............................................................... 31
3.1 Device Overview ............................................................................................................................ 31
3.2 Backgate Devices ........................................................................................................................... 32
3.3 Topgate Devices ............................................................................................................................. 40
   First Process: ............................................................................................................................................ 41
   Second Process: ....................................................................................................................................... 52

Chapter 4: DC Characterization of Silicon Nanowire FET ............................................................... 61
4.1 Experimental Methods ................................................................................................................ 61
4.2 Calculating Capacitance .............................................................................................................. 61
4.3 Backgate Device Characteristics ............................................................................................. 69
4.4 Top-gate Device Characteristics ............................................................................................... 78
4.5 Analysis of DC Measurements .................................................................................................. 83
Chapter 5: Additional Characterization Techniques.................................................................86
  5.1 pH Sensing as a characterization tool ...........................................................................86
  5.2 LF Noise Spectroscopy as a Diagnostic .....................................................................94
Chapter 6: Microfluidics for Sensing ................................................................................102
  6.1 Design Considerations .............................................................................................102
  6.2 Microfluidic probe designs .......................................................................................107
  6.3 Microfluidic probe fabrication ...................................................................................109
    Method 1: .....................................................................................................................109
    Method 2: ....................................................................................................................113
    Method 3: ....................................................................................................................115
  6.4 Microfluidic probe testing .........................................................................................118
  6.5 Microfluidic surface patterning ................................................................................120
Chapter 7: Conclusions ....................................................................................................127
Appendix I: HgFET characterization of SOI Wafers ..........................................................131
Appendix II: Low temperature measurements ....................................................................136
  A2.1 Hall effect and field-effect mobility versus temperature .....................................137
  A2.2 One-Dimensional Subbands .................................................................................139
Bibliography ......................................................................................................................142
List of Figures

Figure 1: One-gigabit DRAM chip, Samsung 50 nm process. [33] .................................................................3

Figure 2: Schematic model of surrounding-gate nanowire MOSFET used to derive drain current. ........................................................................................................................................................................................................17

Figure 3: Nanowire modeled as $n$ segments of length $L/n$ where each segment contains $m$ binding sites .......................................................................................................................................................................................... 29

Figure 4: Overview of fabrication process for back-gate nanowire FETs ..................................................33

Figure 5: Electron-beam lithography patterns (a) Mask file shows large pads (crosshatched) written at high current and wires and small triangular regions written at low current. (b) Exposed and developed HSQ mask shows good alignment between low- and high-current regions. .................................................................................................................................................................................. 37

Figure 6: Scanning electron micrographs of TMAH etched Nanowires. HSQ mask has been stripped with dilute HF to provide better view of surfaces. (a) Viewed from above, both (111) sidewalls are visible. (b) 30 degree tilt shows clear view of (100) and (111) facets........................................................................................................................................................................................................................................37

Figure 7: Minimum-sized nanowire device as determined by silicon thickness and angle of (111) etch planes. Bright line is top of triangular cross section where (111) planes meet. .................................................................................................................................................................................. 38

Figure 8: Backgated nanowire devices (a) Optical micrograph of several 3mm x 3mm dice (b) Optical micrograph of fan-in and device region (c) Optical micrograph showing 4 devices and platinum reference electrodes. Dark purple regions are exposed BOX after passivation and undercutting (d) Electron micrograph of suspended nanowire FET with 80 nm channel width ........................................................................................................................................................................................................................................................................................................40

Figure 9: Overview of first fabrication process for top-gated nanowire FETs .................................42

Figure 10: Alignment key for contact photolithography ........................................................................44
Figure 11: Electron-beam lithography patterns (a) Each die contains 4 device patterns of varying widths. (b) Each device includes chip alignment marks for JEOL 6400. Width of nanowire is determined by spacing between adjacent rectangular regions. (c) Exposed and developed PMMA electron-beam resist with residue revealing line-scans. Residue is removed by oxygen-plasma before transferring pattern into wafer. (d) Differential interference contrast image shows device after pattern transfer and resist stripping.

Figure 12: Device Mesa shown in DIC image after pattern has been transferred through gate stack by reactive ion etching.

Figure 13: Source and drain regions patterned and etched for implantation.

Figure 14: Measured etch rates of as-deposited polysilicon, boron-implanted polysilicon, and intrinsic crystalline silicon (100) in 25% tetramethylammonium hydroxide at 50°C.

Figure 15: Dark-field micrograph shows silicon nitride window with TMAH etched nanowire channel and contact vias.

Figure 16: Top-gated nanowire devices (a) Optical micrograph of several 3mm x 3mm dice (b) Optical micrograph of fan-in and device region (c) Electron micrograph showing single device window (d) Electron micrograph of nanowire FET with 45 nm gate width.

Figure 17: Scanning electron micrographs of suspended top-gated nanowire FET. Nitride sidewall spacers are clearly visible.

Figure 18: TEM images of top-gated nanowire FETs showing sidewall spacers (dark regions) on (a) 45 nm gate-width device (b) 500nm gate-width device.

Figure 19: Nanowire FET chip wire-bonded into 16-pin dual-inline package.

Figure 20: Overview of second top-gate nanowire FET fabrication process.

Figure 21: First electron-beam lithography step for second top-gated nanowire FET fabrication process.
Figure 22: Second electron-beam lithography step defines gate contact region, exposes source and drain regions. (a) DIC optical micrograph of multiple devices on a die (b) Scanning electron micrograph shows individual gate region. .......................................................... 57

Figure 23: (a) Windows opened in interlayer dielectric expose active silicon for TMAH etching. (b) Vias etched through interlayer dielectric allow contacting of source, drain and gate regions. .............................................................................................................................................. 58

Figure 24: (a) Wafer is metalized using a liftoff process. (b) SU-8 passivation layer exposes only device channels and reference electrode: the rectangle in the upper right corner. ................................................................................................................................................................... 59

Figure 25: Scanning electron micrograph of top-gated device after channel is undercut using buffered hydrofluoric acid. ........................................................................................................................................................................... 60

Figure 26: Silvaco simulation of suspended back-gated nanowire (a) Finite-element mesh (b) Accumulated hole density (c) Cross section shows hole density and presence of two channels at high-field (d) Hole density versus position at bottom surface of channel .................................................................................................................................................................. 63

Figure 27: (a) Capacitance versus voltage characteristic of simulated back-gate nanowire FET device. (b) Capacitance per micron of channel length for back-gated nanowire FETs of varying channel widths. Analytical approximation often used for back-gated nanowire transistors is shown not to be applicable to our devices. ............. 64

Figure 28: Capacitance per micron of channel length for suspended back-gated nanowire FETs of varying channel widths. Again, the analytical model drastically overestimates the capacitance................................................................. 65

Figure 29: Silvaco simulation of top-gated nanowire FET (a) Finite-element mesh (b) Accumulated hole density (c) Cross section through midpoint of channel shows accumulation of holes under the gate (d) Carrier density versus position for at bottom surface of channel (e) carrier density versus position at top surface of channel (f) Electric field is largely confined to region beneath gate and falls off rapidly in silicon due to screening. ....................................................................................................................................................... 66
Figure 30: Capacitance per micron of length of dual gate nanowire FETs (a) Top gate capacitance versus gate width (b) Back gate capacitance versus channel width.............67

Figure 31: Calculating series resistances (a) Inversion mode devices of various widths exhibit consistently low series resistance compared to channel resistance. (b) Accumulation mode devices have similarly low series resistance........................................70

Figure 32: Calculating transconductance from $I_d vs V_g$ .........................................................70

Figure 33: Maximum transconductance versus reciprocal channel length in back-gated nanowire FETs. Gm scales linearly with 1/L for long channels, rolls for very short channels due to series resistance in (a) inversion mode devices and (b) accumulation mode devices......................................................................................................................................................71

Figure 34: Field-effect mobility versus channel length in backgated FETs (a) Inversion mode devices (b) Accumulation mode devices ............................................................................................................72

Figure 35: Field-effect mobility versus channel width in back-gated nanowire FETs (a) Inversion mode devices (b) Accumulation mode devices............................................................................................................72

Figure 36: Subthreshold slope of back-gated nanowire FETs versus channel width. N-channel devices show strong dependence on width whereas p-channel devices do not. ..................................................................................................................................................................................74

Figure 37: Energy-band diagram for back-gated nanowire FET. (a) Device is in accumulation thus interface states near conduction band edge are not accessible (b) Device is in inversion so interface and surface states are near fermi level and are accessible.............................................................................................................................................................75

Figure 38: Subthreshold characteristics of nanowire FET devices (a) Suspended accumulation mode device (b) Unsuspended inversion mode device. ........................................77

Figure 39: Output characteristics of back-gated nanowire FETs (a) Suspended accumulation mode device (b) Unsuspended Inversion mode device........................................77

Figure 40: Subthreshold characteristics of unsuspended 45 nm gate-width inversion mode nanowire FET (a) Top-gated operation (b) Back-gated operation .........................80
Figure 41: Subthreshold characteristics of unsuspended 45 nm gate-width accumulation mode nanowire FET (a) Top-gated operation (b) Back-gated operation

Figure 42: Output characteristics of unsuspended 45 nm gate-width inversion mode nanowire FET (a) Top-gated operation (b) Back-gated operation

Figure 43: Output characteristics of unsuspended 45 nm gate-width accumulation mode nanowire FET (a) Top-gated operation (b) Back-gated operation

Figure 44: Subthreshold characteristic of suspended top-gated nanowire FET

Figure 45: Subthreshold characteristic of accumulation mode top-gated nanowire FET fabricated with second method

Figure 46: Sensor chip prepared with Tygon reservoir attached by cyanoacrylate glue

Figure 47: Drain current versus reference electrode voltage at varied back-gate voltages demonstrates that front channel response can be tuned by adjusting back-gate voltage to decouple front and back channels.

Figure 48: Current versus pH measurement (a) Time recording of current as solutions of varying pH are added to the reservoir (b) Characteristic curve for current versus pH

Figure 49: pH response of complementary devices

Figure 50: Nonideal effects in nanowire field-effect pH Sensors (a) Current sampling vs. time: pH varied from 5 to 9.5 and back in 0.5 pH steps (b) Current versus pH curve for same device shows clear hysteresis. (c) Time recording of current during repeated changes of 1 pH unit shows monotonic drift effect. (d) Typical devices exhibit both hysteresis and drift.

Figure 51: Noise versus drain current (a) Normalized noise PSD at varied $V_{ds}$ (b) Constancy of noise amplitude with respect to channel current

Figure 52: Normalized noise power spectral density versus frequency at varied $V_g$
Figure 53: Reciprocal noise amplitude versus gate voltage is proportional to $1/\alpha H$. ...

Figure 54: Hysteresis in devices used for noise spectroscopy (a) TMAH etched devices (b) Cl$_2$ ICP etched devices .......................................................................................................................... 100

Figure 55: Nanowire FET chip prepared for sensing with Tygon reservoir ........................................ 103

Figure 56: Microfluidic probe allows non-destructive microfluidic and electronic probing of individual dice on a wafer. ........................................................................................................................................ 105

Figure 57: Microfluidic probe for automatic probe station allow rapid fluidic screening at the wafer scale. ........................................................................................................................................ 106

Figure 58: Three approaches to microfluidic probing: (a) PDMS gasket surrounding recessed channel. (b) PDMS gaskets interface with ports on SU-8 buried channel. (c) flat bottomed PDMS seals against top of SU-8 walls. ......................................................................................... 107

Figure 59: Outline of fabrication process for method 1 microfluidic probe ........................................ 109

Figure 60: Mask patterns for microfluidic probe method 1 (a) First mask patterns gasket. (b) Second mask patterns channel. ................................................................................................................. 110

Figure 61: Two layer SU-8 mold for microfluidic probe method 1 ...................................................... 111

Figure 62: Method 1 microfluidic probe ready for use ........................................................................ 112

Figure 63: Overview of fabrication process for method 2 microfluidic probe ................................. 113

Figure 64: Mask patterns for buried microfluidic channels ............................................................... 114

Figure 65: Method 2 microfluidic probe (a) SU-8 buried channel (b) Gaskets that interface with buried channel ................................................................................................................................. 115

Figure 66: Overview of fabrication process for method 3 microfluidic probe ......................... 116

Figure 67: Mask pattern for SU-8 ridge ............................................................................................. 117

Figure 68: Method 3 microfluidic probe (a) SU-8 ridge surrounds devices. (b) Flat bottomed probe forms ceiling over ridge ................................................................. 117
Figure 69: Comparison of leakage in fluidic probes. Internal pressure required to cause leak is plotted versus downward force applied to surface. ............................................................... 119

Figure 70: Sampling measurement of drain current at various pH using microfluid probe method 1 ........................................................................................................................................................... 119

Figure 71: Fabrication process for microfluidic surface patterning probe tip ............... 121

Figure 72: Mask patterns for patterning probe device (a) Channel patterns with ports (shaded regions) (b) Fan-in allows access to closely spaced channels................................. 122

Figure 73: Channels and ports viewed from below............................................................. 123

Figure 74: Channel layer has been plasma bonded to fan-in layer to form completed flow-channel network......................................................................................................................... 124

Figure 75: Completed patterning device (a) Fan-in connects cored holes to channels. (b) Final assembly with tubing and bracket for mounting on micropositioner.............. 124

Figure 76: Glass slide selectively labeled with florescein isothiocyanate to demonstrate microfluidic selective surface patterning........................................................................................... 125

Figure 77: Schematic view of HgFET (a) Cross-section (b) Plan-view .............................. 131

Figure 78: PDMS probe for HgFET measurement....................................................................................... 132

Figure 79: Accumulation mode characteristics of SOI HgFET (a) Drain current versus backgate voltage (b) Field-effect mobility of holes ................................................................. 133

Figure 80: Inversion mode characteristics of SOI HgFET (a) Drain current versus backgate voltage (b) Field-effect mobility of electrons ................................................................. 134

Figure 81: Nanowire FET with Hall-bar geometry ............................................................... 137

Figure 82: Hall-effect and field-effect hole mobility versus temerature.......................... 138

Figure 83: Field-effect mobility versus temperature (a) Accumulation mode device (b) Inversion mode device ............................................................................................................................... 139
Figure 84: Oscillations in transconductance due to 1-dimensional sub-band filling.... 140
List of Tables

Table 1: Gate capacitance in farads per micron of all fabricated devices..............................68
Table 2: Parameters of back-gated nanowire FETs.................................................................75
Table 3: Effect of etching conditions on FET performance.....................................................76
Table 4: Dual-gate nanowire FET parameters.......................................................................78
Table 5: Suspended dual-gate FET parameters .................................................................79
Table 6: ISFET parameters for nanowire FET pH sensors.....................................................92
Table 7: Hooge parameter values for various nanowire FETs .............................................99
Table 8: SOI parameters extracted from HgFET measurements ...........................................135
Chapter 1: Introduction

1.1 Motivation

Chemical and biomolecular sensors will change the world around us. During the next few decades will likely see the development of technologies that allow us to better monitor our health with sensors that rapidly test our blood or saliva [1-6]. We may monitor our food, drinking water and even the air we breathe to detect rapidly emerging pathogenic threats [3, 7]. We will probably take drugs developed using new screening techniques that rely on biosensors [1, 2, 4-6, 8-12]. We will protect our borders, cities and airports by detecting explosives or bio-warfare agents [7]. These technologies will all be based upon a single principle, which is the ability to detect a vanishingly small amount of a specific molecule. It is presently impossible to predict all of the methods that human ingenuity will employ in an effort to achieve this task which is conceptually simple yet practically very difficult. Even surveying the historical trends and current state of bio-sensing would take more pages than could be included in this thesis. Thus we will limit ourselves to discussing one small corner of the vast scope of bio-sensing. This tiny corner comprised of electronic field-effect sensors has more than enough history and diversity to keep us interested and busy for quite some time if we were to provide a survey. It also happens to be one of the areas where significant progress has been made recently and yet much remains to be done before this technology can be put to use effectively.

There are many semiconductor material systems which would be suitable for fabrication of nanowire field-effect sensors. Quite a few have been demonstrated
successfully, including silicon, silicon-germanium, indium oxide, tin oxide, gallium-nitride, among others, [13-28] and yet silicon stands out as a clear favorite for the very same reason that the microprocessors in our computers are crafted from silicon: microfabrication. Silicon has proved over and over again to be a practical and versatile material for electronic devices. It does not have the high mobility or direct band-gap of III-V compound semiconductors but it can be grown cheaply, has a stable oxide, has reliable etchants, allows for good control over electronic properties and can be fabricated on scale of very-large wafers. Even in our own lab, we tried several other material systems and growth methods [29-32] before settling on silicon as our material of choice.

Another question which often arises is whether to grow or to fabricate nanowires. Bottom-up growth has enabled researchers to demonstrate a myriad of nanostructures of various material compositions and geometries. However, it remains a challenge to turn some of these materials, which are admittedly exciting and beautiful, into robust device technologies. Ultimately, nature seems to favor bottom-up organization and thus it is certainly a useful and noble task to pursue. By using top-down fabrication, on the other hand, the semiconductor industry has demonstrated the ability to create billions of nanostructures on a single square centimeter of silicon. Figure 1 shows a 1 gigabit DRAM chip fabricated by Samsung in 2006 using a 50 nanometer process. Today these chips can be purchased for a few dollars. This is clearly the path of least resistance in the near future and the one to which I have devoted several years of graduate study.
In any interdisciplinary engineering effort, of which bio-sensing is certainly an example, there is often the worry that we “cannot see the forest for the trees.” Researchers could become so focused on their tiny piece of the puzzle that they forget that the end goal is actually to produce a useful device. However, the opposite is even more common and perhaps more dangerous. In the rush to produce exciting end results, the details are lost and ultimately the results are less understandable and less compelling. In the case of bio-sensing, where we must combine aspects of solid-state device physics and electrical engineering with chemistry, biochemistry, and fluid mechanics it is difficult for one individual to be fully versed in all the details. And yet to fully understand and appreciate the complexity of the operation of nanowire sensors, it is important to have a detailed understanding of the solid-state physics, chemistry, and mechanics that combine to produce a measurable effect. To neglect any of these aspects necessarily leads to poor engineering decisions and to poor science. As a result, we have tried to fully understand our nanowires as semiconductor devices prior using them as sensors, and then make use of this understanding to inform the ways in which we use the sensors in light of what we know about biochemistry and fluid mechanics. The result is hopefully a more complete, robust and usable system for bio-sensing.
1.2 Historical Perspective

It is widely agreed upon that semiconductor biosensors share as a common ancestor the ion-sensitive field effect transistor (ISFET) proposed and fabricated by Bergveld almost 40 years ago [34, 35]. This device, a planar field-effect transistor (FET) capable of detecting ionic concentrations (most often protons) in a solution, has been the subject of extensive study [34-59] will be discussed theoretically in the next chapter as it forms a basis for all other field-effect sensing devices. Variations on the planar ISFETs, such as the enzyme-functionalized field effect transistor (EnFET) were used successfully to detect a variety of analytes such as penicillin, glucose, and urea [53, 60, 61]. In each of these cases, the basic structure of the ISFET was preserved and the modality of sensing remained the same. The immobilized enzyme reacted with the analyte to produce a local concentration of protons, which were then detected by the pH sensitive ISFET. Thus the analyte was detected by the ISFET indirectly.

The next major advance attempted to directly detect selectively-immobilized charged species. Direct immune-detection using an ISFET was introduced by Schenck in 1978 [62]. The ImmunoFET used an ISFET in which the gate insulator was functionalized with antibody molecules. Specific binding of charged analyte molecules were predicted to induce a localized electric potential which would modulate the conductance of the ISFET. This method was largely unsuccessful due to ionic screening, which shields the surface of the ISFET from charge at a distance greater than the Debye length. There were many variations on this theme proposed and tested throughout the 1980's and 1990's, yet none produced incontrovertible evidence of direct detection.
In 2001, Lieber's group reported the first use of a semiconductor nanowire as both an ISFET pH sensor as well as an ImmunoFET [16]. It was not clear from the paper how the Debye screening issue was circumvented and thus this article remains a source of controversy. However, what is clear is that this article set off an avalanche of research and results involving the use of nanowire based FET sensors, presumably excited by the following claim.

*Binding to the surface of a nanowire (NW) or nanotube (NT) can lead to depletion or accumulation of carriers in the "bulk" of the nanometer diameter structure (versus only the surface region of a planar device) and increase sensitivity to the point that single-molecule detection is possible.* [16]

While there was no physical basis provided for this statement, it has often been cited as a justification for using nanowires for sensing [1, 2, 8, 17, 63-65]. In chapter Chapter 2: we discuss in detail the theoretical implications of scaling ISFETs to the nanometer scale; however, the case is not as clear as the quote above would have us believe.

Since 2001 there have been many examples of using nanowire FET sensors both as indirect sensors [14, 15, 21] and as direct sensors to detect proteins [13, 16, 23, 24, 66, 67], small molecules [11, 19, 26, 67, 68], nucleic acids [17, 20, 28, 63, 65, 66, 69], as well as more complex biochemical signaling events such as neuronal action-potentials [70].

1.3 **Overview of Work Presented**

Before discussing the experimental work, Chapter 2 will explore some important theoretical aspects of electronic field-effect sensors, beginning with the historical case
of the ISFET and then examining some of the complex factors that combine to determine the ultimate limits of sensitivity for such sensors. This analysis reveals some important things about how sensors and sensing systems should be designed and how they should be used. Chapter 3 discusses the efforts we have undertaken to fabricate complementary nanowire FET sensors. Several designs and fabrication process are discussed. Chapter 4 discusses the DC characterization techniques we have employed to investigate the performance of the FET devices that we have fabricated. We present the results of thousands of measurements intended to help us understand the operation of the nanowire FETs themselves, as the sensor system is only as good as the devices that comprise it. In Chapter 5, we discuss two additional methods of characterization that we have used for determining the performance of the sensors. First we quantify the sensitivity of our devices using pH sensing as a standard reference. This gives us an objective measure of how well the devices respond to surface charge. Secondly, we investigate the quality of our devices based on low-frequency noise spectroscopy. Low-frequency noise is one of the fundamental limits to reliable ultra-sensitive detection and thus quantification of this noise is an important part of characterizing a sensor device. Chapter 6 discusses a practical new method that we have developed for performing fluid exchange on sensor chips. We demonstrate how this method can be extended to provide a high throughput, wafer scale screening system. Finally we demonstrate how similar methods can be used to pattern surfaces. Chapter 7 offers conclusions based on the work presented in the preceding chapters.

Two appendices are included that discuss related work. The first discusses a method employed for measuring silicon thin-film parameters using a mercury-contact FET. This lets us compare the characteristics of the fabricated devices to those of the
starting material to determine how much the process has affected the material quality and device performance. The second appendix discusses cryogenic measurements performed on our silicon nanowire FET devices.
Chapter 2: Theoretical Background

2.1 ISFET Theory of Operation

As the basis for understanding the operation of all semiconductor field-effect chemical sensors, we discuss the operation of the ion-sensitive field-effect transistor (ISFET) as a pH sensor. This discussion will then be extended to the detection of other ionic species. As stated by inventor, Bergveld [37], “The ISFET is in fact nothing else than a MOSFET with the gate connection separated from the chip in the form of a reference electrode inserted in an aqueous solution which is in contact with the gate oxide.” This suggests that we should begin the discussion with the operation of a conventional MOSFET, wherein the current through the device, assuming non-saturation, is described by the linear model:

\[
I_D = C_{ox} \mu \frac{W}{L} \left[ (V_{gs} - V_t)V_{ds} - \frac{1}{2}V_{ds}^2 \right]
\]

where \( C_{ox} \) is the capacitance per unit area, \( \mu \) is the electron mobility, and \( W \) and \( L \) are the width and length of the channel, respectively. In operation, the parameter \( \beta = C_{ox} \mu W / L \) as well as \( V_{ds} \) and \( V_t \) remain constant and only \( V_{gs} \) is varied, which yields a function for drain current whose only parameter is \( V_{gs} \). \( V_t \), which remains constant for a given device, is determined by the equation:

\[
V_t = \frac{\phi_M - \phi_{Si}}{q} - \frac{Q_{ox} + Q_{ss} + Q_B}{C_{ox}} + 2\phi_f
\]
wherein the first term accounts for the difference in work functions between the gate electrode \( \phi_m \) and the silicon \( \phi_{Si} \), the second accounts for the oxide charge, interface charge and depletion layer charge in the silicon, and the last term accounts for the difference in Fermi level between doped and intrinsic silicon.

In the ISFET, we define the metal terminal of the reference electrode as the gate (albeit a remote one), suggesting that we should define the interface potential due to the solution as a perturbation to \( V_t \) and not \( V_{gs} \) [34-37]. Assuming the fabrication conditions remain the same for both devices, the equation for threshold voltage retains all the same terms, with the addition of the reference electrode potential \( E_{ref} \), and the interface potential \( \psi_0 + \chi^{sol} \), which includes the chemical input parameter \( \psi_0 \), a function of pH, and the surface dipole moment of the solvent, \( \chi^{sol} \). Thus the threshold voltage for an ISFET becomes

\[
V_t = E_{ref} - \psi_0 + \chi^{sol} - \frac{\phi_{Si}}{q} - \frac{Q_{ox} + Q_{ss} + Q_B}{C_{ox}} + 2\phi_f
\]

Assuming a fixed reference potential, the threshold voltage, and consequently the drain current depend only on \( \psi_0 \).

At this point it will be useful to derive a general expression for pH sensitivity of an ISFET, defined as the change in \( \psi_0 \) for a given change in solution pH. This treatment was developed by van Hal et al. [40, 55, 56, 71]. This method is valid for a gate insulator made of any oxide whose charging mechanism is described by association and dissociation of a single amphoteric group [56]. We will discuss the effect of various insulator materials later in the chapter.
We begin with the site-binding model described by Yates [71] which describes the charging of an oxide in a solution as an equilibrium between the surface sites and H⁺ ions in the bulk. The equilibrium reactions are:

\[ AOH \Leftrightarrow AO^- + H_B^+ \text{ and } AOH_2^+ \Leftrightarrow AOH + H_B^+ \]  

with equilibrium conditions

\[
\frac{v_{AO^-} - \alpha_{Hs}^+}{v_{AOH}} = K_a \quad \text{and} \quad \frac{v_{AOH}^2 - \alpha_{Hs}^+}{v_{AOH}^2} = K_b
\]

where \( K_a \) and \( K_b \) are the dissociation constants, \( v_i \) is the density of binding sites per unit area, and \( \alpha_{Hs}^+ \) is the activity of protons at the oxide surface, which can be related to the bulk proton activity by the Nernst equation:

\[ \alpha_{Hs}^+ = a_{H_B^+} \exp \left( \frac{-q\psi_0}{kT} \right) \]

The surface charge density \( \sigma_0 \) can then be expressed as

\[
\sigma_0 = qN_s \left( \frac{\alpha_{Hs}^2 - K_a K_b}{K_a K_b + K_a \alpha_{Hs}^+ + \alpha_{Hs}^2} \right) = -q[B]
\]

where \( N_s \) is the total number of sites and \([B]\) is the density of negatively charged sites minus the number of positively charged sites per unit area. The values for \( N_s \), \( K_a \), and \( K_b \) are dependent on specific material properties of oxide and have been determined experimentally [42, 56]. Another important parameter of a given oxide is the pH at which the surface has zero net charge, the \( pH_{pzc} \) [42, 56]. The change in surface charge density for a given infinitesimal change in pH about the \( pH_{pzc} \) is defined as intrinsic buffer capacity, \( \beta_{int} \).
\[ \frac{\delta \sigma_0}{\delta pH_s} = -q \frac{\delta [B]}{\delta pH_s} \]

\[ = -q N_S \frac{K_b a^2_{H^+} + 4K_a K_b a_{H^+} + K_a K_b^2}{(K_a K_b + K_b a_{H^+} + a^2_{H^+})^2} 2.3a_{H^+} \]

\[ = -q \beta \text{int} \]

The electrolytic side of the double layer contains an equal but opposite charge. It can be treated separately using the Gouy-Chapman-Stern model for a diffuse layer of charge. The charge in the diffuse layer is given by:

\[ \sigma_{DL} = -(8kT \epsilon_0 n^0)^{1/2} \sinh\left(\frac{zq \phi_2}{2kT}\right) = -C_i \psi_0 = -\sigma_0, \]

where \( \phi_2 \) is the potential at the plane of closest approach between the centers of the ions and the surface, \( n^0 \) is the concentration of each ion in the bulk, \( z \) is the magnitude of the charge on each ion, and \( C_i \) is the integral capacitance. The ability of the electrolyte to store charge in response to a change in electrostatic potential is called the differential capacitance and is given by:

\[ \frac{\delta \sigma_{DL}}{\delta \psi_0} = \frac{\delta \sigma_0}{\sigma \psi_0} = - \frac{(2 \epsilon_0 \epsilon_0 z^2 q^2 n^0/kT)^{1/2} \cosh(zq \phi_2/2kT)}{1 + (x_{z/\epsilon_0})^2 (2 \epsilon_0 \epsilon_0 z^2 q^2 n^0/kT)^{1/2} \cosh(zq \phi_2/2kT)} \]

\[ = -C_{dif} \]

Finally, the two sides of the double layer can be combined to give

\[ \frac{\delta \psi_0}{\delta pH_s} = \frac{\delta \psi_0}{\delta \sigma_0} \frac{\delta \sigma_0}{\delta pH_s} = \frac{-q \beta \text{int}}{C_{dif}} \]

The relation between \( pH_s \) and \( pH_b \) was given previously by the Nernst equation. Substitution of this equation into equation 2.11 gives a general expression for the change in surface potential with respect to changes in bulk pH:
\[
\frac{\delta \psi_0}{\delta pH_B} = -2.3 \frac{kT}{q} \alpha, \quad 2.12
\]

where \( \alpha \), a dimensionless sensitivity parameter is given by:

\[
\alpha = \frac{1}{\left(2.3kTC_{dif}/q^2\beta_{int}\right) + 1}
\]

Values for \( \alpha \) can vary between 0 and 1 depending on the intrinsic buffer capacity and the differential capacitance\[42, 55, 56\]. Practically, the buffer capacity is a material property describing the ability of the oxide to take up or deliver ions, and the differential capacitance is mostly determined by the ion concentration in the solution. Only when buffer capacity is high and differential capacitance is low can we get a value for \( \alpha \) that is near 1 and thus approaching the ideal Nernstian sensitivity of -59.3 mV/decade at 25°C. For silicon dioxide, using the site-dissociation model and the Gouy-Chapman-Stern model yields a sensitivity parameter that ranges from close to zero at the \( \text{pH}_{\text{pzc}} \) of 2 to 0.8 at pH 10 \[56\]. At a useful physiocal pH of 7.4 this parameter has a value of \( \sim 0.7 \). Experiments agree with this prediction and SiO\(_2\) has been demonstrated to yield sensitivity up to -45mV/dec at pH 7, with nearly zero response at pH 2. In practice, only Ta\(_2\)O\(_5\), whose intrinsic buffer capacity is relatively insensitive to pH, has been demonstrated to have a buffer capacity high enough to yield Nernstian behavior with measured values of \( \alpha \) as high as 0.95, or 58mV/dec, over the entire range of pH. Si\(_3\)N\(_4\) and Al\(_2\)O\(_3\) are fairly close with \( \alpha \) around 0.9 \[42, 55, 56\].

We now return to the MOSFET model and combine equations 2.3 and 2.1 to give the current in the non-saturated ISFET:

\[
I_D = C_{ox}\mu \frac{W}{L} \left[ \left( V_{gs} - E_{\text{ref}} - \psi_0 + \chi^{\text{soI}} - \frac{\Phi_{Si}}{q} \frac{Q_{ox} + Q_{ss} + Q_B}{C_{ox}} + 2\phi_f \right) V_{ds} - \frac{1}{2} V_{ds}^2 \right] \quad 2.13
\]
Therefore, the change in drain current with respect to pH is:

\[ \frac{\delta I_D}{\delta pH_B} = \beta \left(-2.3 \frac{kT}{q} - \alpha \right) V_{ds} \]  \hspace{1cm} (2.14)

The preceding discussion describes the electrochemical sensitivity of a theoretical ISFET device. We now turn to a discussion of the electrical sensitivity of a MOSFET device. Sensitivity of FET sensors is usually defined as the change in drain current due to a specific stimulus normalized by the initial current. This leads to the conclusion that sensitivity of a given FET device varies depending on the bias conditions. Within the linear or triode regime the current varies linearly with overdrive voltage, \( V_o = V_{gs} - V_t \), as described in equation 2.1. At a given overdrive, the sensitivity is defined and has the unit of V\(^{-1}\):

\[ \frac{\delta I_D / \delta V_o}{I_{\text{initial}}} = \frac{\beta V_{ds}}{\beta \left[(V_o)V_{ds} - \frac{1}{2} V_{ds}^2 \right]} = \frac{1}{V_o - \frac{1}{2} V_{ds}} \]  \hspace{1cm} (2.15)

Assuming constant \( V_{ds} \) we see that sensitivity decreases with increasing \( V_{gs} \) in the linear region.

In the subthreshold region, the drain current is determined by the barrier height between the source and channel regions across which carriers must diffuse. The current across the barrier varies exponentially with barrier height and thus with surface potential in the channel region. The relationship between the surface potential and the applied gate voltage is denoted by a parameter \( \eta \), defined as:

\[ \eta = \frac{d\phi_s}{dV_{gs}} = \frac{1}{1 + (dV_{ox}/d\phi_s)} = \frac{1}{1 + \left( C_d/C_{ox} \right)} \]  \hspace{1cm} (2.16)

where \( C_d \) is the depletion layer capacitance \( \epsilon_{st}/x_d \), and \( C_{ox} \) is the gate oxide capacitance. The form of the drain current is given as:
A semilog plot of the drain current versus gate voltage yields a straight-line region, whose reciprocal slope, termed *inverse subthreshold slope*, or often *subthreshold-swing* is denoted by $S$:

$$S = \frac{1}{\eta q} \ln 10 = \frac{59.3}{\eta} \text{ mV/decade}$$

For an ideal device $\eta$ equals 1 and the subthreshold slope is 59.3 mV/decade at room temperature. In practice, due to the presence of interface states, surface potential depends more weakly on gate voltage, decreasing $\eta$ and causing an increase in subthreshold swing. Values of 70 mV/decade to 120 mV/decade are common for modern MOSFET devices.

The sensitivity of a FET based sensor biased in the subthreshold region, assuming constant $V_{ds}$ would be a constant equal to $1/S$ as shown in equation 2.19. Thus the subthreshold swing represents the upper bound for sensitivity for a given device.

$$\frac{dI_D/dV_g}{I_{D\text{initial}}} = \frac{q\eta}{kT} \frac{I_D}{I_0} \exp\left(\frac{qV_{gs}\eta}{kT}\right) = \frac{q\eta}{kT} = 1/S$$

This is in contrast with previously reported result that sensitivity is maximized at point of peak transconductance [66].

We can combine the results of the electrochemical sensitivity and electrical sensitivity to determine the change in drain current with respect to pH. We will term this the *total sensor sensitivity*, $S_{tot}$. Sensor sensitivity for a pH sensor is determined by
the change in surface potential due to the change in pH multiplied by the change in current due to surface potential as stated in equation 2.20.

\[
S_{tot} = S_{ec} \cdot S_{el} = -\ln 10 \frac{kT}{q} \alpha \cdot \frac{1}{1 - \frac{kT}{\eta q} \ln 10} = \alpha \eta \frac{dec}{pH}
\]

For an ideal nernstian chemical response and an ideal subthreshold slope, \(S_{tot}\) has the ideal value of 1 decade/pH. For a practical device using silicon dioxide the limit should be considerably lower.

2.2 Scaling from ISFETs to Nanowires

There has been no shortage of publicity touting the potential of nanometer scale devices to revolutionize the field of bio-sensing. Many references refer to a “scaling-law” suggesting that the sensitivity of a field-effect sensor should increase as its diameter decreases. The most commonly supplied rationale is a qualitative argument that as the cross sectional area of a device decreases, the ratio of surface area to volume increases. This explanation usually proceeds as follows [72]:

Consider a cylindrical nanowire of diameter \(d\) with uniform doping density \(N_D\) and length \(L\). The conductance of this wire is given by:

\[
G_0 = q \mu N_D \pi d^2 / 4L
\]

Where \(\mu\) is the electron mobility. In this case, the potential is assumed to be uniform throughout the wire (ignoring whatever is outside the wire), hence the mobile charge in the channel equals the doping density. The wire may be considered as an n-channel depletion mode FET in the linear region of operation. If we now introduce a constant
surface charge density, $\sigma$, to approximate the charge due to molecular conjugation, the nanowire will be accumulated or depleted by an equal amount of charge:

$$\Delta Q = \sigma \pi dL \quad 2.22$$

The change in conductance is given by:

$$\Delta G = \pi d\mu \sigma / L \quad 2.23$$

therefore

$$S = \frac{\Delta G}{G_0} = \frac{4\sigma}{qdN_D} \quad 2.24$$

This suggests that sensitivity is inversely proportional to diameter and doping concentration. This approximation doesn’t, however, describe the situation completely accurately because it makes the following assumptions:

1. Constant potential in the radial direction
2. Current is only due to drift, diffusion current is zero (not valid in subthreshold or saturation)
3. No surface states or screening.

We have already demonstrated that sensitivity of an FET will be maximized in the subthreshold region so we require a model that can account for subthreshold conduction. Several analytical solutions have been developed for current in a cylindrical surrounding-gate nanowire by solving the Poisson-Boltzmann equation to determine the potential distribution throughout the nanowire channel and then using the Pao-Sah current formulation [73] to express the current in terms of mobile carrier concentrations [74-78]. This method has been used to model dual-gate as well as
surrounding gate MOSFETS with zero and finite body doping. For simplicity and because it yields the theoretical upper limit on sensitivity we will consider the undoped case.

Following the method developed by Jimenez [77], we consider an un-doped (or lightly doped) n-type, surrounding gate, long-channel MOSFET as shown in Figure 2.

![Schematic model of surrounding-gate nanowire MOSFET used to derive drain current.](image)

In cylindrical coordinates, the Poisson equation has the form:

\[
\frac{d^2\psi}{dr^2} + \frac{1}{r} \frac{d\psi}{dr} = \frac{kT}{q} \delta \frac{g(\psi - V)}{e^{kt^r}}
\]

where \(\delta = q^2 n_i/kT\varepsilon_{si}\), \(\psi(r)\) is the electrostatic potential, and \(V\) is the electron quasi-Fermi potential and \(\varepsilon_{si}\) is the permittivity of silicon. This equation is subject to the boundary conditions:

\[
\frac{d\psi}{dr}\bigg|_{r=0} = 0, \quad \psi\bigg|_{r=R} = \psi_S
\]

where \(\psi_S\) is the surface potential. We assume that \(V\) is constant in the radial direction (the gradual channel approximation). The Poisson equation can be solved analytically to yield:
\[
\psi(r) = V + \frac{kT}{q} \ln \left( \frac{-8B}{\delta(1 + Br^2)^2} \right)
\]

where the constant B is related to \( \psi_s \) by the second boundary condition in equation 2.26. The total mobile charge is given by:

\[
Q = C_{ox}(V_{gs} - \Delta \phi - \psi_s)
\]

where \( C_{ox} = \varepsilon_{ox}/(R \ln (1 + t_{ox}/R)) \) and \( \Delta \phi \) is the work-function difference between the gate electrode and the intrinsic silicon. Applying Gauss’s law gives the following relation:

\[
C_{ox}(V_{gs} - \Delta \phi - \psi_s) = Q = \varepsilon_{st} \frac{d\psi}{dr} \bigg|_{r=R}
\]

Combining equations 2.27 and 2.29, yields:

\[
\frac{q(V_{gs} - \Delta \phi - V)}{kT} - \ln \left( \frac{8}{\delta R^2} \right) = \ln(1 - \beta) - \ln(\beta^2) + \eta \left( \frac{1 - \beta}{\beta} \right)
\]

where \( \beta \) is defined as \( 1 + BR^2 \), constant with respect to \( r \), and determined from equation 2.30 and \( \eta = 4\varepsilon_{st}/C_{ox}R \) is a structural parameter. For a given \( V_{gs} \), \( \beta \) can be solved as a function of \( V \) from equation 2.30, where \( V = 0 \) at the source and \( V = V_{ds} \) at the drain. The functional dependence of \( V \) and \( \beta \) on \( y \) is determined by the current continuity equation which requires that \( I_{ds} = \mu(2\pi R)Q \frac{dV}{dy} \) is constant with respect to \( V \) and \( y \). Following the Pao-Sah current formulation, we integrate \( I_{ds} dy \) over the length of the channel, expressing \( dV/dy \) as \( (dV/d\beta)(d\beta/dy) \):

\[
I_{ds} = \frac{2\pi R}{L} \int_0^{V_{ds}} Q(V) dV = \frac{2\pi R}{L} \int_{\beta_s}^{\beta_d} Q(\beta) \frac{dV}{d\beta} d\beta
\]
Where $\beta_d$ and $\beta_s$ are solutions to equation 2.30 corresponding to $V = V_{ds}$ and $V = 0$. The total mobile charge per unit gate area, in terms of $\beta$ is given by combining equations 2.27 and 2.27 to yield

$$Q(\beta) = (2\varepsilon_s)(\frac{2kT}{q})(\frac{1}{\beta R})(1 - \beta)$$  \hspace{1cm} 2.32

while $dV/d\beta$ is given as a function of $\beta$ by differentiating eq. 2.30. Substituting these into equation 2.31 and integrating analytically yields $I_{ds}$ in terms of $\beta$

$$I_{ds} = \mu \frac{4\pi\varepsilon_s L}{q} (\frac{2kT}{q})^2 \left[ \eta \frac{1 - \eta}{2} + \frac{\ln \beta}{2} \right]_{\beta_d}^{\beta_s}$$  \hspace{1cm} 2.33

To simplify the algebra, we define two functions to represent the RHS of equations 2.30 and 2.33:

$$f(\beta) = \ln(1 - \beta) - \ln(\beta^2) + \eta \left( \frac{1 - \beta}{\beta} \right)$$  \hspace{1cm} 2.34

$$g(\beta) = \frac{\eta}{4\beta^2} + \frac{1 - \eta}{\beta} + \frac{\ln \beta}{2}$$  \hspace{1cm} 2.35

For any given $V_{gs}$ and $V_{ds}$, we can calculate $\beta_s$ and $\beta_d$ from the conditions

$$f(\beta_s) = \frac{q(V_{gs} - V_0)}{kT}$$  \hspace{1cm} 2.36

$$f(\beta_d) = \frac{q(V_{gs} - V_0 - V_{ds})}{kT}$$  \hspace{1cm} 2.37

where

$$V_0 = \Delta \phi + \frac{kT}{q} \ln \left( \frac{8}{\delta R^2} \right)$$  \hspace{1cm} 2.38
The drain current is then compute by inserting $\beta_s$ and $\beta_d$ back into equation 2.33. Using this model we can now derive the current for the SRG FET in each region of operation.

In the subthreshold region, $f(\beta_s)$ and $f(\beta_d) \ll 1$, so $\beta_s$ and $\beta_d \sim 1$, therefore $f(\beta_{s,d}) \sim \ln(1 - \beta_{s,d})$ and $g(\beta_{s,d}) \sim \eta/4\beta_{s,d}^2 + (1 - \eta/2)/\beta_{s,d}$. The current can then be expressed as:

$$I_{ds} = \mu \frac{\pi R^2 L}{n_i kT} e \frac{q(V_{gs} - \Delta \phi)}{kT} \left(1 - e^{-q(V_{ds})/kT}\right)$$

We can calculate the transconductance in the subthreshold region by differentiating with respect to $V_{gs}$:

$$G_m = q\mu \frac{\pi R^2 L}{n_i e} \frac{q(V_{gs} - \Delta \phi)}{kT} \left(1 - e^{-q(V_{ds})/kT}\right)$$

Sensitivity can be expressed as $G_m/I_{ds}$, so combining equations 2.39 and 2.40, yields

$$S = \frac{G_m}{I_{ds}} = \frac{q}{kT}$$

This may seem like an unexpected result because there is no dependence on $R$, however it is actually quite intuitive when considering transistors (and ISFETs) achieved nearly 60mV/dec subthreshold slope several decades ago (for the long-channel case), and their continual shrinkage hasn’t improved upon this physical limit.

In the linear region of operation, $f(\beta_s)$ and $f(\beta_d) \gg 1$, so $\beta_s$ and $\beta_d \sim 0$, therefore $f(\beta_{s,d})$ is dominated by the term $\eta (1 - \beta_{s,d})/\beta_{s,d}$ and $g(\beta_{s,d})$ is dominated by the term $\eta/4\beta_{s,d}^2$. The current can then be expressed as:

$$I_{ds} = 2\mu C_{ox} \frac{\pi R}{L} \left(V_{gs} - V_t - \frac{V_{ds}}{2}\right) V_{ds}$$
where \( V_t = V_0 - \eta (kT/q) \) represents the threshold voltage. \( C_{ox} \), \( V_0 \) and \( \eta \) are each dependant on R so the size dependence of the drain current isn’t immediately clear from this expression. \( V_t \) may be expressed as a function of R:

\[
V_t = \Delta \phi + \frac{kT}{q} \ln \left( \frac{8}{\delta} \right) - \frac{2kT}{q} \ln \left( R \left( 1 + \frac{t_{ox}}{R} \right)^{2e_{sl}/e_{ox}} \right)
\]

We can calculate the transconductance in the linear region by differentiating with respect to \( V_{gs} \):

\[
G_m = 2 \mu C_{ox} \frac{\pi R}{L} V_{ds}
\]

Thus the sensitivity can be calculated as:

\[
S = \frac{G_m}{I_{ds}} = \frac{1}{\left( V_{gs} - V_t - \frac{V_{ds}}{2} \right)}
\]

Rewriting this in terms of R gives the much more unwieldy equation

\[
S = \frac{1}{\left( V_{gs} - \Delta \phi - \frac{kT}{q} \ln \left( \frac{8}{\delta} \right) + \frac{2kT}{q} \ln \left( R \left( 1 + \frac{t_{ox}}{R} \right)^{2e_{sl}/e_{ox}} \right) - \frac{V_{ds}}{2} \right)}
\]

For devices with very thin oxide compared to the radius of the silicon channel, the 4th term in the denominator reduces to \( \ln(R) \), giving the sensitivity a roughly \( 1/\ln(R) \) dependance.

Overall, we can conclude that making a sensor smaller will improve the sensitivity in the linear region but not in the subthreshold region, where the sensitivity is highest. There may be other considerations for making sensors smaller such as
binding kinetics, or fluid dynamics; however, electrical sensitivity is not necessarily a valid reason for shrinking devices to the nanometer scale.

### 2.3 Chemical Binding and Fluidics Considerations

There are at least three other factors that should be considered when discussing sensitivity of scaled-down FET sensors. These are the effects of chemical binding, fluidic transport, ionic screening on the sensor sensitivity. These three factors are intimately connected and cannot truly be separated though we will try to discuss them individually.

Let us consider a sensor in a static solution containing a specific analyte. For simplicity, we will assume first-order Langmuir kinetics and thus the concentration of conjugated receptors on the surface of the sensor is given by

\[
\frac{dN}{dt} = k_{\text{on}} (N_0 - N) \rho_s - k_{\text{off}} N \tag{2.47}
\]

where \(k_{\text{on}}\) and \(k_{\text{off}}\) are the kinetic capture and dissociation constants, \(N_0\) is the areal density of surface receptors, \(N\) is the density of conjugated receptors and \(\rho_s\) is the concentration of analyte molecules present at the sensor surface at time \(t\). If the molecules are supplied to the solution much more quickly than they can be bound, then the conjugation is in the “reaction-limited” regime. The surface concentration can be taken to equal the bulk concentration \(\rho_0\), and equation 2.47 can be solved as [79]:

\[
\frac{N(t)}{N_0} = \frac{\rho_0 / K_D}{1 + \rho_0 / K_D} \left( 1 + e^{-(k_{\text{on}} \rho_0 + k_{\text{off}})t} \right) \tag{2.48}
\]

where \(K_D = k_{\text{on}} / k_{\text{off}}\). The fraction of bound receptors in equilibrium is given by
For a sensor of a given area \(A\), we can calculate a critical concentration \(\rho_1\) at which only a single molecule binds the sensor in equilibrium:

\[
\rho_1 = \frac{K_D}{N_0A}
\]  \(\text{2.50}\)

This represents a lower bound on the sensitivity imposed by the binding of the analyte, regardless of the electrostatics of the sensor. Using an example given by Squires [79], of an antibody with \(K_D = 1 \text{ nM}\), \(\rho_1\) may be estimated for sensors of various sizes. Assuming the surface modification is optimized, a sensor surface may have a receptor density as high as \(N_0 = 2 \times 10^{12}\). A nanowire sensor with a 10 nm diameter and 2 um length would have about \(10^3\) binding sites on its surface yielding a critical concentration \(\rho_1 = 1 \text{ pM}\). In comparison, a planar sensor 2um x 2um would have about \(10^5\) binding sites and thus \(\rho_1 = 10 \text{ fM}\). Thus, in the reaction-limited regime it would seem that maximizing the surface area would allow for a decrease in minimum detectible concentration.

Now let us consider the case where the surface binding occurs more quickly than the molecules can be delivered to the surface. The concentration at the surface is again given by equation 2.47 and by the diffusion of molecules across the concentration gradient near the surface due to the depletion of molecules from the solution by binding:

\[
\frac{d\rho}{dt} = D\nabla^2\rho
\]  \(\text{2.51}\)
where $D$ is the diffusion coefficient of the analyte molecules in solution. The flux of molecules to the surface of the sensor is given by

$$I = D \int_{\text{Area}} \nabla \rho ds \quad 2.52$$

The density of conjugated receptors as a function of time would be given by simultaneously solving equations 2.47 - 2.49. There is no closed form solution to these equations. We can, however make an approximation which renders the problem more tractable. For most relevant sensing systems, such as DNA-DNA, antibody-protein, or specific protein-protein interactions, $k_{off}$ is very low compared to the $k_{on}$ and therefore at least initially, the conjugation of the surface depends only on the forward reaction. We can view this as irreversible binding or we can just look at it as a system that is far from equilibrium.

This problem has been solved for a number of different surface geometries including planar, linear, and spherical surfaces [72, 79-81]. Based on the approach taken by Nair and Alam we see that for a large planar sensor, the conjugation as a function of time is given by:

$$N(t) = \rho_0 t \left( \frac{1}{D/\sqrt{2Dt} + \frac{1}{k_F N_0}} \right)^{-1} \quad 2.53$$

whereas for a cylindrical nanowire sensor, the conjugation is given by

$$N(t) = \rho_0 t \left( \frac{2\pi R}{D/\ln\left[\frac{\sqrt{4Dt} + R}{R}\right] + \frac{1}{k_F N_0}} \right)^{-1} \quad 2.54$$

These equations represent the interplay between diffusion, represented by the first term and binding kinetics, represented by the second term. They are used to calculate
the density of conjugated receptors after a given time for a particular initial concentration of analyte molecules. Comparing a planar sensor to a 1-D nanowire sensor, assuming $k_F \to \infty$, which represents purely diffusion-limited conjugation, suggests that the nanowire sensor we considered in the reaction limited case should achieve 3-4 orders of magnitude higher levels of conjugation than planar sensors. However, even for the nanowire sensors, the time required to capture a detectible number of molecules with an initial concentration of 100fM is calculated to be around 1 hour for a diffusion constant of corresponding to a 12 base-pair molecule of DNA [80, 81].

To circumvent the bottleneck imposed by the diffusion-limit many have proposed adding convective transport using microfluidic channels to increase the flux of analyte molecules to the surface of the sensor. This has been analyzed both numerically [81, 82] and analytically [79, 81]. Each of these studies assumes pressure-driven Poisseuille flow in an incompressible fluid, which is described by the Navier-Stokes equation. The approximate analytic solution for the analyte flux in terms of flow rate and physical dimensions is given by [83] and further verified by [81] as:

$$J = D L N_D \rho_0 \left[ \frac{2\pi}{4.885 - \ln(P_s)} \left(1 - \frac{0.09266P_s}{4.885 - \ln(P_s)} \right) \right]$$

$$P_s = \frac{6QW^2}{Dwh^2}$$

where $L$ and $W$ are the sensor length and width, respectively, $w$ and $h$ are the channel width and height, respectively, and $Q$ is the volumetric flow rate. We can use this equation to calculate the flux versus flow rate for the nanowire and planar ISFET sensors that we have been considering assuming a channel with a height of 50 nm.
Revisiting the 10 nm x 2 um nanowire and 2um x 2um ISFET sensors that we discussed earlier, and \( D = 150 \text{um}^2 \text{s}^{-1} \) corresponding to a 20-mer of ssDNA we calculate \( P_e \approx 3 \times 10^{-3}Q \) for the nanowire sensor and \( P_e \approx 30Q \) for the planar sensor. Using equation 2.55, we calculate that increasing the flow rate from 1 microliter per second to 1000 microliters per second only increases the flux by about a factor of 2 for the nanowire, while for the planar sensor the increase is about a factor of 10. This result is in good agreement with the results of numerical methods employed by others [79, 82]. It shows that even for extremely high flow rates, diffusion limits cannot be overcome by convective transport in laminar flow because the fluid velocity at the channel boundaries is essentially zero.

It has been demonstrated that for fully developed turbulent flow in a micro-channel, as in a macro-scale duct, the velocity is very-nearly constant throughout the channel, in contrast with the parabolic flow profile observed in laminar flow [84]. As such, the transport of analyte molecules to the nano-sensor should be much more efficient in a micro-channel under turbulent flow conditions than under laminar flow conditions. The transition has been observed to occur in micro-scale channels with Reynolds numbers at or below the Reynolds numbers corresponding to the transition in macroscale ducts. Above Reynold number 2500-3000, the turbulent flow should be fully developed[84].

Reynolds number is defined as:

\[
Re = \frac{QD_H}{\nu A}
\]

where \( D_H \) is the hydraulic diameter, defined for a rectangular duct as \( D_H = 4A/P \), with \( A \) being the cross sectional area of the duct, \( P \) being the wetted perimeter, \( Q \) is the
volumetric flow rate, and \( \nu \) is the viscosity of the fluid. For the typical micro-channels used in the studies referenced above, hydraulic diameters were in the range in 50-100 microns, leading to maximum Reynolds numbers of around 500. If, however, we can increase the hydraulic diameter to 500 um, Reynolds numbers up to 2500 can be achieved and we can take advantage of turbulent flow to improve the sensor response.

There are a few other practical considerations to discuss that will affect the ultimate sensitivity and usefulness of field effect sensors as they are scaled aggressively. The first of these is the discrete nature of charge. Both the dopant atoms in the nanowire sensor channel and the analyte molecules themselves have been considered as a uniform continuous charge density; however, for a nanowire sensor with a doping density of \( 1 \times 10^{15} \), which is actually high compared to most of our nanowire sensors, there is roughly 1 dopant atom per micron of length. This can lead to stochastic fluctuations in sensor performance, such as threshold voltage from device to device. It was shown in [72] that such fluctuations can affect the response to a single molecule by as much as an order of magnitude.

We have fabricated nanowire sensors from high-resistivity SOI wafers with doping levels below \( 10^{13} \) to attempt to mitigate this effect. The fact that the analyte itself is comprised of discrete, charged molecules presents a different challenge. The charge from a single analyte molecule is localized and thus can only affect a volume of the channel that is within the Debye length in the silicon. For a very narrow channel, a single bound molecule may be sufficient to deplete or accumulate a region that is comparable to the entire channel width. The nanowire may, however, be many times longer than the Debye length, and thus require a large number of bound analyte molecule to affect the entire length of the channel.
We can illustrate this point in a simple way by considering the wire of length \( L \) as a collection of \( N \)-segments of length \( L/N \) as shown in Figure 3. The number of segments is chosen so that \( L/N \) represents the maximum length over which a single bound molecule on the surface of the nanowire may accumulate or deplete the wire, i.e. the Debye length. Each segment contains \( M \)-sites which can be conjugated by an analyte molecule and each segment is considered off unless one of its \( M \)-sites is conjugated. At equilibrium, and at a concentration equal to the dissociation constant, \( K_D \), the probability of a single site \( S_{m,n} \) being occupied is \( P_{m,n} = \frac{1}{2} \). We will consider the population of each site to be uncorrelated to all the other sites, a valid approximation in the reaction limited case. Then for a nanowire that is being accumulated by conjugation the probability that each segment will contain at least 1 conjugated site, the minimum condition to turn the wire “on,” will be described by

\[
(1 - (1 - P_{m,n})^M)^N
\]

Let us consider, for example, a 10 nm x 2 um long nanowire with a Debye length of 50 nm, and a 2 x 10^{12} cm\(^2\) surface receptors. This wire will be divided into 40 segments, each having about 30 surface receptors. At a concentration of \( K_D/5 \), the probability of a given site being occupied is 0.1, thus the probability of the wire being “on” would be \( (1 - (1 - 0.1)^{30})^{40} = 17.7\% \). At \( K_D/10 \), the probability of the wire being on is less than 0.01 percent. We can compare this to the device which is depleted by conjugation where only a single segment must contain a single conjugated device to turn the wire off, thus the probability of the wire turning “off” is given by

\[
1 - (1 - P_{m,n})^{1200}
\]
Even for a concentration of $K_D/500$, where the individual site binding probability is 0.001, the probability of the wire turning “off” would be $1 - (1 - 0.001)^{1200} = 70\%$. This very simple model that we have developed suggests that the minimum detectable concentration for depletion mode devices may be much lower than for accumulation mode devices. Zhou and Wei [85] demonstrate a similar effect by modeling the nanowire and planar sensors as a random network of resistors.

The importance of this finding shouldn’t be underestimated as it emphasizes the necessity for complementary sensor elements. The reason for this is the fact that at a given pH, a particular analyte molecule may have a positive or negative charge depending on the isoelectric point of the molecule. It would be convenient to set the pH to an extreme value so that all molecules of interest would be of the same polarity; however, this would likely denature the molecules and therefore isn’t an option. Thus, for a wire with a particular carrier type, in this case holes, conjugation of a positively charged molecule would deplete the wire, while conjugation of a negatively charged molecule would accumulate the wire. To ensure that we can sense all molecules of interest by depletion of the wire, thus taking advantage of the lower threshold of detection, would require that we have both n-channel and p-channel devices.

![Figure 3: Nanowire modeled as $n$ segments of length $L/n$ where each segment contains $m$ binding sites](image)

We have attempted to address the main considerations that affect the sensitivity of nanowire field-effect sensors. We have shown that based on the electrostatics, field-effect sensors are most sensitive when biased into the subthreshold region, and in this
region, the sensitivity is independent of size, so that it may not be worth pursuing the aggressive scaling of sensor devices on the basis of electrical sensitivity alone. We have demonstrated that in the reaction limited regime, sensors should become less responsive as they are scaled down, however in the diffusion limited regime the opposite should be true, although the time scales required to sense very low concentrations are predicted to be prohibitively long. Even with the introduction of high flow rates, in laminar flow the response is still limited by diffusion and only increases a few-fold. We have calculated that by using large channel diameter and very high flow rates, we should be able to push into the fully-turbulent flow regime and achieve much higher mass-transport than by diffusion alone. Finally, we have seen that the discrete nature of charge requires that we consider individual stochastic events in nanowire sensors. We have demonstrated that depletion mode sensors should be more sensitive than accumulation mode devices. Together, all of these factors represent a fraction of the complexity inherent in nanowire sensing systems. Clearly good design, modeling and experimentation must be brought together if we are to fully understand these systems.
Chapter 3: Silicon Nanowire FETs Design and Fabrication

3.1 Device Overview

During the course of this work, we went through many different design iterations to improve both the performance and manufacturability of the nanowire FET sensors. One of the main considerations was the need for complementary devices for the reason proposed at the end of the last chapter: that the threshold of detection should be considerably lower for depletion-mode sensing devices than accumulation-mode sensing. Further, to be able to put both n-channel and p-channel FETs on a single chip and operate both simultaneously requires localized-gate control of each nanowire FET. Another important consideration was minimizing the effects of process-induced damage on the performance of the nanowire FETs. Process-induced damage often manifests itself as interface or surface trap states which can reduce mobility, and thus drive current, reducing signal to noise ratio (SNR), as well as producing low-frequency noise, which also decreases SNR.

This chapter will focus on two types of devices. The first of these was a globally back-gated nanowire FET fabricated on silicon-on-insulator (SOI) wafers. We will refer to these from now on simply as “back-gated devices.” These devices are an extension of the work done previously in our lab [66]. The design improves manufacturability by decreasing the number of processing steps and has enabled fabrication of devices with critical feature sizes down to 20 nm. Both accumulation- and inversion-mode devices
are demonstrated on high-resistivity p-type substrates. We have also produced freely-suspended nanowire FETs which increases the available surface area and relaxes the diffusion constraints of a surface bound wire to possibly improve sensitivity.

The second type of device is a locally-gated nanowire FET which we will refer to as “top-gated devices”. These devices have individually-addressable top-gates which enables simultaneous complementary device operation. They are fabricated using a method that yields very little mobility-degradation and excellent noise characteristics. Adding the top-gate has further implications for sensing sensitivity as the carrier population can be moved away from the back surface where it is relatively insensitive to charged species on the surface. These devices also have usefulness for low-dimensional physics as they represent a well-controlled physically-confined quasi-1-dimensional material. Unlike the back-gated devices, the top-gated devices have a degenerately-doped gate enabling low-temperature measurements down to liquid helium temperatures. High mobility and low noise are two critical features for these types of measurements. This will be discussed further in Appendix II: Low temperature measurements. Finally, we have produced free-standing suspended top-gated devices. Because the top-gate shields much of the channel from the solution during sensing, we expose the bottom surface of the channel to the solution by undercutting and releasing it from the substrate making it available for conjugation by analyte molecules.

3.2 Backgate Devices

An outline of the fabrication process for the back-gated devices is shown in Figure 4. All of the process steps prior to step 5 were performed at the Cornell
Nanoscale Science & Technology Facility cleanroom (CNF). The remaining steps were performed in the Yale cleanroom.

Figure 4: Overview of fabrication process for back-gate nanowire FETs
The masks for this process were written on the TRE Electromask CC-251 optical pattern generator/image repeater in the Yale cleanroom. 7 inch master contact masks were generated by a combination of the image-repeater for repeated features and the pattern generator for unique features.

We started with 6” SOI wafers from SOITEC. The wafers had a 205 nm thick high-resistivity boron-doped active layer (> 2000 ohm-cm) and a 405 nm buried oxide layer (BOX). The high resistivity wafers were chosen to minimize ionized impurity scattering during low temperature measurements. The handle wafers were boron-doped with a resistivity of 14-22 ohm-cm. Both silicon layers were of (100) orientation.

Prior to processing, the wafers were cleaned using RCA Standard Clean-1 (ammonium hydroxide : hydrogen peroxide : water 1:1:6), RCA Standard Clean-2 (hydrochloric acid : hydrogen peroxide : water 1:1:6) and a 50:1 water : hydrofluoric acid dip to strip the native oxide. In Step 1 the active layer was thinned down to 25 nm by a series of thermal oxidations and sacrificial hydrofluoric acid etches. The furnace tubes were cleaned by flowing HCl gas for at least 2 hours prior to loading wafers.

Alignment marks were etched into the wafer in Step 2 to allow alignment of the electron-beam lithography system as well as all subsequent contact lithography steps. These zero-level alignment marks were patterned by contact lithography. The pattern was aligned to the flat of the wafer as closely as possible to ensure that all lithographic patterns are aligned to the crystallographic axes of the wafer. This is extremely important later in the process because of an orientation-dependent etch. The alignment pattern was transferred into the wafer by reactive-ion etching. First, the pattern was etched through the active layer using a CF₄ RIE process in an Oxford Plasmalab 80+
Second, the pattern was etched through the BOX using a CHF$_3$/Ar RIE process in the same tool. Finally, the pattern was transferred into the handle wafer to a depth of 3-4 microns using a silicon deep-RIE (Bosch process, Unaxis 660 ICP, CNF). These alignment marks can readily be seen during all the subsequent processing steps for ease of alignment.

The source and drain regions of the devices were doped by ion implantation in step 3. The wafers were patterned with positive photoresist (Shipley 1813) and contact lithography to expose only the source and drain regions. These regions were doped by low-energy ion implantation on an Eaton NV-6200AV ion implanter at CNF. Boron difluoride ions were implanted at 10 KeV to a dose of $5 \times 10^{15}$ for p-type contacts and arsenic ions were implanted at the same energy and dose for n-type contacts. These species had a projected range of 10-12 nm at the specified energy, as calculated by SRIM [86] suggesting most of the dopant would remain in the active layer. The dopant was distributed throughout the thickness of the active layer and electrically activated by rapid-thermal annealing at 1000°C for 1 minute in argon (AG Heatpulse 610, CNF).

The wafers were patterned using a JEOL 9300FS electron-beam lithography system in step 4. This step is required to produce the critical dimensions of the nanowire FET channels. By using a negative-tone electron-beam resist, we can pattern the entire mesa structure including the source and drain pad regions and the channels with a single lithographic step. This is compared to the previous process [66, 87], whereby we would first use contact lithography to pattern the rough mesa geometry and then further trim down the channels after patterning with positive-tone electron beam resist. While this would not be a wise move for production-scale fabrication as it
increases the e-beam write time significantly, it saves a mask level and several process steps per wafer and is thus very favorable for our level of production. Because of the flexible nature of this step, we have smoothly spanned a range of 2 orders of magnitude in channel width to investigate the transition from planar micron-scale FETs down to quasi-1D nanowire FETs.

Hydrogen Silsequioxane (HSQ) (Dow-Corning XR-1514) was applied to the wafer by spin-coating to yield a 135 nm thick layer. The HSQ was exposed at 100 KeV using doses ranging from 500-2500 uC/cm² depending on the size of the feature and proximity to other large features. The critical features (device channels and small triangular regions in Figure 5) were exposed at a beam current of 2 nA. These were drawn with various widths ranging from 20 nm up to 2 um wide. To save time, large features (pads and large trapezoidal regions in Figure 5) were exposed at 20 nA. The unexposed HSQ was developed in a tetramethylammonium hydroxide based developer (AZ 300-MIF) until visibly cleared, around 4 minutes. The HSQ was then annealed by RTA at 900 °C for 5 minutes in an oxygen atmosphere. This treatment drives any remaining hydrogen out of the HSQ and yields a dense SiO₂-like material with good wet- and dry-etch resistance.

In step 5 the electron-beam lithography patterns are transferred into the silicon active layer using tetramethylammonium hydroxide (TMAH), an orientation-dependant silicon etchant. This etchant etches in the <100> crystallographic direction of silicon nearly 100 times faster than it does in the <111> direction, leaving smooth sidewalls defined by the (111) planes. Figure 6 shows an electron micrograph of two devices after TMAH etching.
Figure 5: Electron-beam lithography patterns  (a) Mask file shows large pads (crosshatched) written at high current and wires and small triangular regions written at low current.  (b) Exposed and developed HSQ mask shows good alignment between low- and high-current regions.

Figure 6: Scanning electron micrographs of TMAH etched Nanowires. HSQ mask has been stripped with dilute HF to provide better view of surfaces. (a) Viewed from above, both (111) sidewalls are visible. (b) 30 degree tilt shows clear view of (100) and (111) facets.

The TMAH etched sidewalls ideally have a lower density of surface states or bulk trap states than surfaces etched by a plasma etching process which can cause surface damage and shallow implanted ions which can serve as traps. The TMAH does not etch silicon dioxide or HSQ at an appreciable rate, thus, prior to etching, the native oxide must be removed from the surface of the silicon. For large channels widths, this can be
achieved by a brief dip in dilute hydrofluoric acid. However, the HF will also attack the HSQ, which is more porous than thermal oxide and is thus etched more quickly. For very narrow channels, the native oxide was removed using a brief CF$_4$ RIE process in the Oxford Plasmalab 80+ (Yale). The devices were then etched for 10 minutes in TMAH at 50°C (25% by weight, in water). This produced a device whose bottom dimension was roughly equal in width to the top dimension originally drawn by the electron beam due to some undercutting of the HSQ mask. The bottom width of the smallest device that can be patterned is determined by the thickness of the silicon active layer and the sidewall angle and is given by $w_{\text{min}} = 2t / \tan(54.7°)$. For our devices which have a thickness of 25 nm, the minimum width is 35.4 nm. A minimum-width device is shown in the electron micrograph in Figure 7.

![Electron micrograph of nanowire device](image)

**Figure 7:** Minimum-sized nanowire device as determined by silicon thickness and angle of (111) etch planes. Bright line is top of triangular cross section where (111) planes meet.

The wafer was patterned for metal lift-off in step 6 using a bi-layer resist process with positive photoresist (Shipley 1808) and lift-off resist (Microchem LOR 5A). 20 nm titanium and 100 nm gold were deposited by electron-beam evaporation and lifted-off
to produce the metal pads, fan-in and contacts. This produced ohmic contacts without any additional annealing. The backside of the wafer was also metalized with the same deposition process.

In step 7, the devices were passivated with a layer of negative-tone epoxy-based resist (SU-8 2002, Microchem). A 2um thick layer of SU-8 was patterned via contact lithography to open windows exposing the contact pads as well as the device channels and reference electrodes. This purpose of this passivation layer is to isolate the metal layer during aqueous sensing experiments and prevent short circuits through the solution.

At this point, some of the devices were undercut using buffered hydrofluoric acid, shown as step 8. This etchant acts isotropically on the silicon dioxide, thus it undercuts the wire by the same amount as the etch depth. As such, channels as wide as 500 nm can be released from the buried oxide while preserving enough of the BOX to electrically isolate the channel from the back-gate. Some of the devices were not undercut for comparison. This step was sometimes performed between steps 6 and step 7 with similar results.

An additional, optional processing step was performed on some devices between steps 6 and 7. A second liftoff process was used to define a platinum pseudo-reference electrode as shown. An 80 nanometer thick layer of platinum was deposited by electron-beam evaporation and lifted off with the same bi-layer resist process as the first metallization. The platinum reference electrode was later omitted as gold was found to be sufficiently inert to serve as a pseudo reference in most cases. The finished back-gate devices are shown in Figure 8.
3.3 **Top-Gate Devices**

Two separate processes were developed for fabricating top-gated nanowire FETs. The resulting devices are fairly similar; however the fabrication processes are quite distinct. We took lessons learned from fabricating the first design and applied them to the second to speed up the fabrication process and increase the yield. In both cases, we were able to produce high-quality nanowire-FETs with near-bulk mobility indicating low levels of plasma damage. Both structures were undercut with HF yielded
fully suspended individually gated nanowires, suggesting their utility for complementary sensing applications.

**First Process:**

An outline of the fabrication process for the first top-gate device process is shown in Figure 9. All of the process steps except for the TMAH-etch were performed at the Cornell NanoScale Science & Technology Facility cleanroom. The TMAH etching was performed in the Yale cleanroom. All photomasks were either purchased from Compugraphics USA or written on the TRE Electromask CC-251 in the Yale Cleanroom.

Eight inch diameter ultrathin SOI wafers were purchased from SOITEC having 55 nm thick active silicon and 145 nm buried oxide (BOX). Both the active layer and handle wafer were boron doped to 14-22 ohm-cm. The wafers were laser cut to 4” (Silicon Quest) for compatibility with some of the processing tools in the CNF and Yale cleanrooms. Prior to beginning processing and prior to all furnace steps the wafers were cleaned using the RCA clean. This included 10 minutes in RCA Standard Clean-1 (ammonium hydroxide : hydrogen peroxide : water 1:1:6), 10 minutes in RCA Standard Clean 2 (hydrochloric acid : hydrogen peroxide : water 1:1:6) and a 50:1 water : hydrofluoric acid dip to strip the native oxide. Each cleaning step was followed by a rinse in DI water to at least 16 Mohms.
Figure 9: Overview of first fabrication process for top-gated nanowire FETs
The active silicon layer was thinned down to 30 nm by a series of thermal oxidations and sacrificial hydrofluoric acid etches. Prior to the first oxidation step, the furnace was purged with Cl\textsubscript{2} gas for at least 2 hours to minimize any metal contamination. A 12 nm thick gate oxide was grown by dry thermal oxidation at 900°C. A heavily boron-doped polysilicon gate layer was deposited 45 nanometers thick by LPCVD at 600°C. This layer was further implanted with boron at 10 KeV to a dose of 5 x 10\textsuperscript{15} (Core Technologies) to increase its etch-resistance to tetramethylammonium hydroxide.

Using contact lithography, zero-level alignment marks were patterned on the wafer in positive photoresist. These marks were transferred through the gate, gate oxide, active and BOX layers using a CF\textsubscript{4} inductively coupled plasma (ICP) etch process in an Oxford Plasmalab 100 RIE/ICP system (CNF). The pattern was further etched 4 microns into the handle wafer using a deep silicon ICP process (Bosch Process) in a Unaxis 660 ICP. This pattern included alignment marks for all subsequent contact lithography steps as well as for the JEOL 9300FS electron beam lithography system at CNF and the JEOL 6400/NPGS system in the Reed Lab at Yale. Using the alignment keys shown in Figure 10, I was able to achieve better than one-half micron layer-to-layer alignment error on the EVG 620 contact aligner.
Figure 10: Alignment key for contact photolithography

The electron beam lithography was the first step performed after patterning alignment marks because the lack of topography allowed for more uniform coating of electron beam resist and thus more reproducible line-widths. A 150 nm thick layer of PMMA (950 Å) was applied by spin coating and exposed using the JEOL 9300FS electron beam lithography system at CNF. The acceleration voltage was set to 100 KV with a beam current of 2nA. The exposed wafers were developed for 70 seconds in 1:3 methyl isobutyl ketone (MIBK) : isopropanol and then rinsed with isopropanol. PMMA ridges that would define the FET device channels were left behind by writing pairs of closely spaced rectangles. A dose array was written and the optimal dose was determined to be 1400 uC/cm² by measuring the developed patterns in a scanning electron microscope (SEM). As shown in Figure 11, each die contains 4 devices, and the channel width of each device is defined by the spacing of the pair of rectangles written during this electron beam lithography step. Each die was patterned with a 50 nm, 100 nm 200 nm and 500 nm wide device. The length of the channels was varied from die to die and ranged from 1 micron to 50 microns. The open rectangles were transferred into the underlying gate stack using a CF₄ RIE process to etch through the gate and a
CHF$_3$/Ar RIE process to etch through the gate oxide. Both processes were performed in an Oxford Plasmalab 80+ RIE system (CNF).

Figure 11: Electron-beam lithography patterns  (a) Each die contains 4 device patterns of varying widths. (b) Each device includes chip alignment marks for JEOL 6400. Width of nanowire is determined by spacing between adjacent rectangular regions. (c) Exposed and developed PMMA electron-beam resist with residue revealing line-scans. Residue is removed by oxygen-plasma before transferring pattern into wafer. (d) Differential interference contrast image shows device after pattern transfer and resist stripping.
The wafers were then patterned with contact lithography to produce mesas that define the source and drain regions of each device as shown in Figure 12. The mesa patterns were transferred through the gate stack and active silicon layers by CF₄ RIE in an Oxford Plasmalab 80+ (CNF).

![Figure 12: Device Mesa shown in DIC image after pattern has been transferred through gate stack by reactive ion etching.](image)

The gate stack was removed from the source and drain regions of the mesas by lithographically patterning windows over these regions and etching through the gate poly with CF₄ and then through the gate oxide with CHF₃/Ar using the Oxford Plasmalab 80+. The etched device is shown in Figure 13. The same patterned photoresist was then used as a mask during ion implantation to doped the source and drain regions. This yielded a device with very little gate over- or underlap. P-channel devices were doped with BF₂⁺ at 10 KeV to a dose of 5 x 10¹⁵ (Core Technologies). The lithography was repeated for the n-channel devices on the wafer which were implanted with As⁺ at 10 KeV to a dose of 5 x 10¹⁵. The dopant was electronically activated by rapid thermal annealing at 1000°C for 60 seconds in argon using an AG Heatpulse 210 RTA (Yale).
A 250 nanometer thick stoichiometric silicon nitride (Si₃N₄) layer was deposited by LPCVD at 700°C over the entire wafer to act as an interlayer dielectric. This layer electrically isolates the gate contacts from the source and drain contacts. A 50 nanometer thick silicon dioxide layer was deposited on top of the nitride using a Group Sciences (GSI, CNF) plasma enhanced CVD (PECVD) system to act as a hard mask for patterning the nitride layer. Using contact lithography, a window was opened over the channel of each device. This window was transferred through the PECVD oxide and through the top 200 nanometers of the silicon nitride using a CHF₃/O₂ RIE process in the Oxford 80+ (CNF). The final 50 nanometers of silicon nitride was wet-etched using phosphoric acid at 160°C. This etchant exhibits excellent etch selectivity for silicon nitride over silicon dioxide, crystalline silicon and polysilicon and thus the nitride was removed from the window without damaging the underlying device channel or gate stack.

The exposed channel region of the device was then etched in TMAH (25% by weight, in water) at 50°C. Intrinsic polycrystalline silicon typically etches in TMAH at a rate somewhere between the rates of intrinsic (111) and (100) silicon; however, at very high levels of p-type doping, the etch rate of silicon is known to slow down very significantly. We measured the etch rates of the as-deposited and ion-implanted
polysilicon versus undoped (100) silicon as shown in Figure 14. The etch rate slowed by about a factor of 50 as a result of the doping.

![Etch Rate Graph](image)

**Figure 14:** Measured etch rates of as-deposited polysilicon, boron-implanted polysilicon, and intrinsic crystalline silicon (100) in 25% tetramethylammonium hydroxide at 50°C

The next step is to etch vias through the nitride layer to access the source, drain and gate contacts. Various sized vias, down to 1 micron, were patterned using contact lithography. The patterns were transferred through the oxide and nitride layers using the same plasma-etch process used to open the windows over the channel. The devices are shown after opening contact-vias in a dark-field optical micrograph in Figure 15.

![Dark-field Micrograph](image)

**Figure 15:** Dark-field micrograph shows silicon nitride window with TMAH etched nanowire channel and contact vias.
Since some of the devices will be used as bio-sensors in aqueous environments, it was desirable to protect the surfaces of the nanowire channels with silicon nitride to prevent the diffusion of metal ions into the silicon or oxide layers. To minimize the density of interface states, a thin thermal oxide was first grown. Three to four nanometers of oxide were grown at 900°C in dry oxygen. The wafer was immediately transferred into the LPCVD furnace and 6 nm of silicon nitride were deposited.

Before depositing metal for the probe pads, lead-ins and contacts, the wafers were lithographically patterned for liftoff using a bi-layer resist process with positive photoresist (Shipley 1808) and lift-off resist (Microchem LOR 5A). The nitride and thermal oxide deposited in the previous step were etched using a CHF3/O2 RIE process in the Oxford 80+ (CNF). 30 nanometer of titanium/tungsten 10%/90% were deposited by electron beam evaporation (Denton, Yale) followed by 250nm aluminum. Liftoff was performed at room temperature in n-methylpyrrolidinone (NMP). Completed devices structure is shown in Figure 16.
We were concerned about the contact resistance for the smallest vias so a test wafer was fabricated to verify that the series resistance due to the contacts would be low enough. The device wafers were annealed in argon until the contact resistance was less than 5 Kohms. This required annealing at 425°C for 10 minutes.

The device wafers were protected with photoresist and diced using a K&S 7100 wafer saw (CNF). Devices to be suspended were etched briefly in CHF₃/O₂ in the oxford 80+ at Yale to remove the thin nitride layer. The buried oxide was then etched in 10:1...
buffered hydrofluoric acid to undercut the device channels. The undercut device is shown in Figure 17.

![Figure 17: Scanning electron micrographs of suspended top-gated nanowire FET. Nitride sidewall spacers are clearly visible.](image)

It was possible to undercut the devices without attacking the gate oxide due to the presence of a silicon nitride sidewall spacer. This spacer was formed on some of the devices due to an incomplete etch of the LPCVD nitride interlayer dielectric. Because of the conformal nature of the deposition and the anisotropic nature of the RIE etching, a region of nitride was left behind on either side of the gate stack. This region is clearly visible in the transmission electron micrographs shown in Figure 18. A similar technique is often used in CMOS processing to form insulating sidewall spacers.
Figure 18: TEM images of top-gated nanowire FETs showing sidewall spacers (dark regions) on (a) 45 nm gate-width device (b) 500nm gate-width device

Devices were glued into DIP packages with low temperature conducting epoxy (stycast) and wirebonded with a Westbond ultrasonic wedge bonder. A packaged device is shown in Figure 19.

Figure 19: Nanowire FET chip wire-bonded into 16-pin dual-inline package

Second Process:

A second process was developed specifically for fabricating suspended top-gated nanowire FETs. This process draws upon many of the lessons learned when fabricating the back-gated devices and uses the same mask set with the addition of a few extra masks. The overview of this process is shown in Figure 20.
Figure 20: Overview of second top-gate nanowire FET fabrication process
We started with the same 6" SOI as we used for the back-gate process. Again, the wafers were cleaned using RCA Standard Clean-1, RCA Standard Clean and an HF dip. The oxidation furnace tube was cleaned by flowing HCl gas for at least 2 hours prior to loading wafers. The active layer was thinned down to 25 nanometers by a series of thermal oxidations and sacrificial hydrofluoric acid etches. This time, a 25 nanometer thick stoichiometric silicon nitride gate dielectric layer was deposited by LPCVD at 800°C. Given its higher dielectric constant, the equivalent-oxide-thickness was nearly the same as the 12 nanometer silicon dioxide gate we used before. We predicted that this nitride layer would have a high density of interface states and would thus yield poor device characteristics, however, we required a material that would be resistant to etching in buffered hydrofluoric acid (BHF) to allow undercutting of the nanowires without etching the gate dielectric. Given more time, this layer would ideally be replaced by a high-K ALD deposited layer such as Hafnium oxide or aluminum oxide but LPCVD silicon nitride was sufficient for proof of concept. A 30 nanometer thick layer of Boron doped polysilicon was deposited by LPCVD at 600°C to serve as the gate conductor. This layer was implanted with BF2 at 10 KeV to a dose of 5 x 10¹⁵ using the Eaton ion-implanter at CNF and annealed for 60 seconds at 950°C in the AG Heatpulse 610 at CNF to electrically activate the dopant atoms.

Alignment marks were etched into the wafer to allow alignment of the electron-beam lithography system as well as all subsequent contact lithography steps. As in both the previous processes, the zero-level alignment marks were patterned by contact lithography. This pattern was transferred into the wafer by reactive-ion etching. First, the pattern was etched through the active layer using a CF₄ RIE process in an Oxford Plasmalab 80+ (CNF). Second, the pattern was etched through the BOX using a CHF₃/Ar
RIE process in the same tool. Finally, the pattern was transferred into the handle wafer to a depth of 3-4 microns using a silicon deep-RIE (Bosch process, Unaxis 660, CNF).

The wafer was patterned using a JEOL 9300FS electron-beam lithography system to define the nanowire FET channels and source/drain pad structures. As with the back-gate process, we used HSQ negative electron beam resist. The device channels, ranging from 20nm to 2 um were exposed at 2nA beam current. To save time, large non-critical features were exposed at 20 nA and small features were exposed at 2 nA. The unexposed HSQ was developed in AZ 300-MIF developer until visibly cleared, around 4 minutes. The HSQ was then annealed by RTA at 900 °C for 5 minutes in an oxygen atmosphere to drive out remaining hydrogen and densify it.

This pattern was transferred through the gate layer using a CF$_4$ reactive-ion etch process in an Oxford Plasmalab 80+ RIE system (CNF). Because of the lack of endpoint detection and poor selectivity of this etch chemistry, it was necessary to rely on timing to set the etch depth and stop within the gate dielectric layer. Ideally, a Cl$_2$ RIE process such as is used in Yale's Oxford 100 RIE/ICP would yield better results because of its excellent etch selectivity of silicon versus most dielectrics. The gate dielectric layer was etched using a CHF$_3$/O$_2$ RIE process which yields approximately 4:1 selectivity for silicon nitride versus crystalline silicon and allows for stopping the etch on the silicon active layer. The HSQ mask was removed by wet etching with BHF leaving the patterned gate and gate dielectric layers shown in Figure 21.
The gate layer was further patterned by a second round of electron-beam lithography. This step could easily have been replaced by a contact or projection lithography step as the minimum feature size was 2 microns. Electron-beam lithography was chosen simply because of ease of alignment and quick turnaround of new design prototypes. The wafers were coated with HSQ and exposed on the JEOL-9300FS at 100KeV with a beam current of 5nA. The unexposed HSQ was developed in AZ 300-MIF for 4 minutes. This pattern was transferred into the gate stack using the same etching processes as the previous step. During the CF$_4$ RIE process, the active layer surrounding the source/drain pad structure was etched down to the BOX in addition etching the polysilicon gate-layer as shown in Figure 22.

At this point in the process, the source and drain regions need to be doped. The HSQ remaining from the previous gate-trimming step was used for a self-aligned source/drain doping process. A shallow ion-implantation of BF$_2^+$ at 7KeV to $5 \times 10^{15}$ was used for p-channel devices (Eaton, CNF). N-channel devices were not completed due to an error in the As$^+$ ion-implantation step. The dopant was activated by rapid-thermal annealing for 60 seconds at 1000 °C with a ramp rate of 50°/sec in an argon
environment (SSI, Yale). After implantation, the HSQ was removed by etching in 10:1 BHF.

Figure 22: Second electron-beam lithography step defines gate contact region, exposes source and drain regions. (a) DIC optical micrograph of multiple devices on a die (b) Scanning electron micrograph shows individual gate region.

An interlayer dielectric was deposited to electrically isolate the gate electrode contact from the source and drain contacts. A 200 nanometer thick PECVD silicon dioxide layer was deposited over the entire wafer surface (GSI, Yale). The wafer was patterned by contact lithography to open a small window over the nanowire portion of each device. The SiO₂ layer exposed by the windows was etched 80% of the way through using a CHF₃/Ar RIE process (Oxford 80+, Yale) and then the final 20% was done using 10:1 BHF to allow etch stopping on the exposed active-silicon and polysilicon gate. The devices with patterned and etched windows are shown in Figure 23 (a). The silicon active layer exposed by the window was etched in TMAH (25% in Water) at 50°C for 10 minutes. This etchant does not etch the doped polysilicon gate, nitride gate dielectric, or buried oxide layers.
Figure 23: (a) Windows opened in interlayer dielectric expose active silicon for TMAH etching. (b) Vias etched through interlayer dielectric allow contacting of source, drain and gate regions.

Vias were patterned as shown in Figure 23 (b) using contact lithography to allow contacting of the source, drain and gate. The vias were etched 80% of the way through the oxide using a CHF$_3$/Ar RIE process (Oxford 80+, Yale) and then the final 20% was done using 10:1 BHF to allow etch stopping on the silicon or polysilicon contacts.

The wafer was patterned for metal lift-off using a bi-layer process with positive photoresist and Lift-off resist (LOR). 20 nanometers of titanium and 100 nanometers of gold were deposited and lifted-off to produce the metal pads, fan-in and contacts. Ohmic contacts were produced without any additional annealing. The metalized devices are shown in Figure 24 (a).

The surface of the wafer was passivated with a layer of negative-tone epoxy-based resist (SU-8, Microchem), which was patterned via contact lithography to open windows exposing the contact pads as well as the device channels. This purpose of this passivation layer is to isolate the metal layer during sensing experiments and to protect the silicon dioxide interlayer dielectric during the final channel undercutting etch. One
of the metal wires has an exposed region near the device which acts as a pseudo-reference electrode during sensing experiments. This region is visible in Figure 24 (b).

![Figure 24: (a) Wafer is metalized using a liftoff process. (b) SU-8 passivation layer exposes only device channels and reference electrode: the rectangle in the upper right corner.](image)

Lastly, the channels were undercut using an isotropic wet etch. BHF 10:1 was used to partially etch the buried oxide layer exposed by the window. Because the etchant acts isotropically on the silicon dioxide, it undercuts the wire by the same amount as the thickness etched. Thus, devices having channels as wide as 500 nm can be released from the buried oxide while preserving enough of the BOX to electrically isolate the channel from the back-gate. Because we have chosen a gate dielectric that is not etched by BHF, the polysilicon gate and gate dielectric are unaffected by the etchant. A suspended device is shown in Figure 25.
Figure 25: Scanning electron micrograph of top-gated device after channel is undercut using buffered hydrofluoric acid.
Chapter 4: DC Characterization of Silicon Nanowire FETs

4.1 Experimental Methods

The DC electrical characteristics of the fabricated devices were measured using an Agilent 4156B semiconductor parameter analyzer, Agilent 5250E switching mainframe, and a Cascade Microtech Summit 12K automated probe station. This setup was controlled by a custom Labview program to allow automatic screening of every device on a wafer. The Cascade probe station was also used manually for individual measurements with the 4156B. It provided a well shielded, light-tight environment for device measurements.

4.2 Calculating Capacitance

It is important to accurately determine the gate-capacitance of a field effect transistor because gate-capacitance is an important parameter in all of the models used to describe FET operation. The accuracy of the gate-capacitance estimate affects the accuracy for all calculated parameters, such as field-effect mobility. Unfortunately, we cannot easily measure the gate capacitance as we could for large planar FETs because it is so small compared with the parasitic capacitance of the fan-in and contacts. Most previously reported efforts to calculate nanowire device parameters rely on analytical calculations for gate-capacitance based on the device geometry [64, 88]. While these methods may be plausible for certain geometries at very high doping levels, they have often been stretched past their limits of applicability. Our geometry is one which would
not be well suited to the typical approximation of a metal-cylinder over a metal plane, due to its geometrical complexity, low level of doping, and small size that leads to volume accumulation and inversion. As such, we have employed numerical finite element analysis to calculate the gate-capacitance.

The gate capacitance of each nanowire FET was determined by modeling with Silvaco, a physics-based finite element analysis platform for simulating semiconductor devices. Silvaco DEVEDIT 3D was used to define a mesh corresponding to a 3 dimensional model of the device. The meshes were then simulated using the ATLAS DEVICE3D solver to simultaneously solve the potential and carrier concentrations in the wire as well as to extract the capacitance of the gate to the channel. The gate-channel capacitance is extracted at 1 MHz.

To calculate the capacitance per unit length for each device-width, we simulated several devices with lengths from 0.2um to 1um. We plotted the calculated capacitance versus length and found the zero-length intercept which represented the extra capacitance due to the ends of the wire and due to the contacts. We then divided the calculated capacitances by the length of the segments and subtracted the intercept value to find the actual unit gate-capacitance. Some of the simulations were performed with a small drain-source bias because having a small drift-current aided convergence.

For the back-gate only nanowires, we simulated the wires before and after etching the buried oxide. An example of the back-gate mesh with etched oxide is shown in Figure 26 (a). Figure 26 (b) shows the hole density in the accumulated wire. Note that there are two quasi 1-d channels that form along the bottom edges of the wire due to the high electric field. A cross section shown in Figure 26 (c) through the midpoint of
the wire confirms this and we see the carrier density versus position along the bottom interface in Figure 26 (d).

**Figure 26:** Silvaco simulation of suspended back-gated nanowire (a) Finite-element mesh (b) Accumulated hole density (c) Cross section shows hole density and presence of two channels at high-field (d) Hole density versus position at bottom surface of channel

The capacitance versus voltage curve for one of the devices is plotted in Figure 27 (a). As we would expect, the accumulation capacitance asymptotically approaches a constant value which is the capacitance of the gate insulator, $C_{ox}$. Figure 27 (b) shows the extracted value for $C_{ox}$ versus device width for each of the unsuspended back-gated devices that we fabricated. We also plot the calculated value using the cylindrical
approximation[88] for comparison. It is clear from this plot that this analytical model is not valid for our FETs, due to the low doping and non-cylindrical geometry.

Figure 27: (a) Capacitance versus voltage characteristic of simulated back-gate nanowire FET device. (b) Capacitance per micron of channel length for back-gated nanowire FETs of varying channel widths. Analytical approximation often used for back-gated nanowire transistors is shown not to be applicable to our devices.

The suspended back-gated nanowires were simulated by removing the top 120 nanometers of the buried oxide and replacing it with air. The extracted values for $C_{ox}$ versus device width for each size of device that we fabricated are shown in Figure 28. The extracted capacitance values for the suspended devices are about one third of the value for the corresponding unsuspended devices.
Figure 28: Capacitance per micron of channel length for suspended back-gated nanowire FETs of varying channel widths. Again, the analytical model drastically overestimates the capacitance.

For the dual-gate devices, the top-gate capacitance was solved with the back-gate grounded and likewise for the back-gate capacitance the top-gate was grounded. The mesh for a dual-gate device in top-gate only operation is shown in Figure 29 (a). The accumulated hole concentration is shown in Figure 29 (b). The gate has been hidden in the figure to reveal the high charge-density region beneath it. The cross section is shown in Figure 29 (c) and the cutlines through this cross section shown at the two interfaces, in figures Figure 29 (d) and Figure 29 (e), illustrate the carrier density as a function of position. Note that the entire volume of the nanowire is somewhat accumulated, even though the carriers are predominantly concentrated under the gate. This is in contrast to the inversion mode operation where inversion layer is strictly limited to the region directly under the gate. Because of this difference, the inversion mode capacitance should be somewhat lower than the accumulation mode capacitance.
Figure 29: Silvaco simulation of top-gated nanowire FET (a) Finite-element mesh (b) Accumulated hole density (c) Cross section through midpoint of channel shows accumulation of holes under the gate (d) Carrier density versus position for at bottom surface of channel (e) carrier density versus position at top surface of channel (f) Electric field is largely confined to region beneath gate and falls off rapidly in silicon due to screening.
The suspended dual-gate nanowires were simulated by removing the top 100 nanometers of the buried oxide and replacing it with air. The extracted value for top- and back-gate capacitance versus device width for each size of device that we fabricated is shown in Figure 30.

![Figure 30: Capacitance per micron of length of dual gate nanowire FETs (a) Top-gate capacitance versus gate width (b) Back-gate capacitance versus channel width](image)

For field-effect mobility calculations, the extracted capacitance at the onset of conduction should be used as this is where the effective mobility is highest. This corresponds to the flatband voltage for accumulation-mode devices or the threshold voltage for inversion-mode devices. This is sometimes estimated by taking the point one third of the way up the slope between the minimum value and the oxide capacitance [89]. In our case, we used knee of the curve to give an upper bound on the capacitance as we weren’t entirely confident in our ability to estimate the flatband voltage from the modeled current versus gate-voltage characteristic.
Table 1 gives a summary of the capacitance values calculated using Silvaco for each of the device geometries that we fabricated. These values will be used to calculate field-effect mobility in this chapter as well as Hooge’s parameter in the next chapter.

**Table 1: Gate capacitance in farads per micron of all fabricated devices**

<table>
<thead>
<tr>
<th>Back-gate only</th>
<th>Channel top width (nm)</th>
<th>Unsuspended Capacitance (F/µm)</th>
<th>Suspended Capacitance (F/µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>3.363 x 10^{-17}</td>
<td>1.6157 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>3.596 x 10^{-17}</td>
<td>1.74883 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>3.818 x 10^{-17}</td>
<td>1.876 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>4.014 x 10^{-17}</td>
<td>1.99017 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>4.239 x 10^{-17}</td>
<td>2.101 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>4.412 x 10^{-17}</td>
<td>2.20208 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>5.276 x 10^{-17}</td>
<td>2.70092 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>5.958 x 10^{-17}</td>
<td>3.165 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>6.564 x 10^{-17}</td>
<td>3.52583 x 10^{-17}</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>7.139 x 10^{-17}</td>
<td>3.87583 x 10^{-17}</td>
</tr>
<tr>
<td>Dual-gate</td>
<td>Top-gate width (nm)</td>
<td>Unsuspended Capacitance (F/µm)</td>
<td>Suspended Capacitance (F/µm)</td>
</tr>
<tr>
<td>Top-gated</td>
<td>45</td>
<td>1.81 x 10^{-16}</td>
<td>1.855 x 10^{-16}</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>3.28 x 10^{-16}</td>
<td>3.39 x 10^{-16}</td>
</tr>
<tr>
<td></td>
<td>195</td>
<td>6.105 x 10^{-16}</td>
<td>6.05 x 10^{-16}</td>
</tr>
<tr>
<td></td>
<td>495</td>
<td>1.415 x 10^{-15}</td>
<td>1.445 x 10^{-15}</td>
</tr>
<tr>
<td>Back-gated</td>
<td>45</td>
<td>9.135 x 10^{-17}</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>1.224 x 10^{-16}</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>195</td>
<td>1.5795 x 10^{-16}</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>495</td>
<td>2.5155 x 10^{-16}</td>
<td>-</td>
</tr>
</tbody>
</table>
4.3 Backgate Device Characteristics

All devices on the back-gated FET wafers were screened using the Cascade automatic probe station to determine their $I_d \text{ vs } V_g$ characteristics. Each wafer had roughly $10^4$ devices so, for the sake of time, we had to select representative devices for more thorough measurements and analysis.

We first measured the series resistance of the devices due to the contacts, fan-in and source and drain resistances to verify that it was low compared to the channel resistance. The series resistance was determined by measuring the on-current through a series of devices of different lengths at a fixed bias condition. The measured current divided by drain-source bias gives the channel resistance plus the series resistance. By plotting this total resistance versus length and extrapolating to find the y-intercept we can deduce the series resistance. This was done for devices of both types with a variety of channel widths as is shown in Figure 31. The measurements were taken with the devices biased fully-on to minimize channel resistance. The measured series resistance was about 50 kOhm, which is low enough to be neglected in most cases considering that the typical channel resistance is one megohm or greater.

We calculated values of field effect mobility, threshold or flatband voltage, and subthreshold slope for a number of devices of each length, width and type based on the screening data. Field effect mobility is typically calculated for planar FETs using the equation:

$$\mu_{FE} = \frac{L}{W C_{ox}} \frac{g_m}{V_{ds}}$$

4.1
Figure 31: Calculating series resistances (a) Inversion mode devices of various widths exhibit consistently low series resistance compared to channel resistance. (b) Accumulation mode devices have similarly low series resistance.

where $C_{ox}$ is the gate capacitance per unit area and $g_m$ is the maximum value of transconductance, determined from the numerical derivative of the $I_d$ vs $V_g$ curve as shown for one of our devices in Figure 32. We will redefine $C_{ox}$ in equation 4.1 as a capacitance per unit length calculated for a given device width extracted using Silvaco and we replace the geometric factor $L/W$ with simply the channel length, $L$.

Figure 32: Calculating transconductance from $I_d$ vs $V_g$
This method is widely accepted for estimating low-field carrier mobility; however, it yields the effective mobility at a gate voltage above the threshold voltage, where the mobility is degraded due to the high transverse electric-field. Thus, the calculated field-effect mobility represents a lower bound on the actual low-field carrier mobility.

Figure 33 illustrates the relationship between transconductance and channel length and width.

**Figure 33:** Maximum transconductance versus reciprocal channel length in back-gated nanowire FETs. Gm scales linearly with $1/L$ for long channels, rolls for very short channels due to series resistance in (a) inversion mode devices and (b) accumulation mode devices.

As expected, the transconductance scales linearly with reciprocal length for devices with channels longer than a few microns. For shorter channel devices, the series resistance becomes significant and thus the measured transconductance is limited. As a result, the calculated mobility is also limited for devices with short channels. We plot the electron and hole field-effect mobility versus length in Figure 34.
to illustrate this effect. The measured mobility approaches the actual value asymptotically as the length increases.

Figure 34: Field-effect mobility versus channel length in backgated FETs (a) Inversion mode devices (b) Accumulation mode devices

Taking the average mobility of the long-channel devices, we plot mobility versus channel width in Figure 35.

Figure 35: Field-effect mobility versus channel width in back-gated nanowire FETs (a) Inversion mode devices (b) Accumulation mode devices
Threshold and flat-band voltages were calculated by extrapolation of the linear portion of the transfer characteristic. These values are, however, very dependent on the thickness of the silicon film. Our back-gate device wafers have a high degree of thickness variation in the silicon film due to the fact that we started with wafers having 205 nm active and thinned these to 20-30 nm. An acceptable variation of two percent in the starting thickness, would translate into a variation of 20 percent in the final wafer thickness. This thickness variation yields large variations in threshold and flat-band voltage for nominally identical devices across the wafer. Therefore, it wasn’t very useful to compare these values between devices of different sizes unless they were from the same location on the wafer. This suggests that to be able to integrate these devices at a large scale would require a much tighter tolerance on the wafer thickness. In the top-gated wafer where we started with a thinner active silicon layer, we see much lower variation of threshold and flat-band voltages.

Inverse subthreshold-slope, usually simply referred to as subthreshold swing or S, was calculated from a linear fit of the subthreshold portion of the transfer characteristic on a semilog plot. Using the subthreshold swing we can estimate the density of interface states for the buried interface according to the equation:

\[
S = 2.3 \frac{kT}{q} \frac{C_{ox} + C_{it} + C_d}{C_{ox}}
\]

We observed different subthreshold behavior in the n-channel and p-channel devices. In the n-channel, the subthreshold swing was very dependent on the device width, whereas in the p-channel devices, the subthreshold swing was much lower and was relatively constant with respect to channel width as shown in Figure 36.
Figure 36: Subthreshold swing of back-gated nanowire FETs versus channel width. N-channel devices show strong dependence on width whereas p-channel devices do not.

The fact that the subthreshold swing scales with length can be understood when we recall the plot of oxide capacitance versus width from the previous section. The capacitance due to the interface traps scales linearly with channel width, while the oxide capacitance scales less than linearly due to the volume effect so the ratio of $C_{it}/C_{ox}$ decreases as the channel width decreases, and thus the subthreshold swing also decreases with channel width. The fact that only the n-channel devices exhibit significant $C_{it}$ gives us some insight into the nature of the interface trap states. It suggests that we have a high density of surface or interface states near the conduction band edge as shown in the energy band diagram in Figure 37. These states are easily accessible when the channel is inverted giving rise to a large $C_{it}$. When the device is in accumulation, the states are far from the Fermi level and thus they do not contribute to the capacitance. Because of the volume inversion effect in very thin active SOI, the surface states at the air/silicon surface are also accessible and may play a role as well.
Figure 37: Energy-band diagram for back-gated nanowire FET. (a) Device is in accumulation thus interface states near conduction band edge are not accessible (b) Device is in inversion so interface and surface states are near fermi level and are accessible.

Table 2: Parameters of back-gated nanowire FETs shows a summary of the average mobility and subthreshold slope for the various channel widths. We also calculated the same values for devices prepared by different etching methods as shown in Table 3. It is noteworthy that the TMAH etched devices have higher mobility and lower subthreshold slope than either set of dry-etched devices. The fabrication of these test devices is discussed further in the next section.

Table 2: Parameters of back-gated nanowire FETs

<table>
<thead>
<tr>
<th>Channel Width (nm)</th>
<th>Hole Mobility (cm²/V-s)</th>
<th>Subthreshold Slope (V/dec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>42.0</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>83.5</td>
<td>0.72</td>
</tr>
<tr>
<td>80</td>
<td>107.9</td>
<td>0.75</td>
</tr>
<tr>
<td>100</td>
<td>148.3</td>
<td>0.62</td>
</tr>
<tr>
<td>200</td>
<td>155.5</td>
<td>0.59</td>
</tr>
<tr>
<td>300</td>
<td>259.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>
Inversion mode devices (n-channel)

<table>
<thead>
<tr>
<th>Channel Width</th>
<th>Hole Mobility (cm$^2$/V-s)</th>
<th>Subthreshold Swing (V/dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>267</td>
<td>6.22</td>
</tr>
<tr>
<td>40</td>
<td>237</td>
<td>5.38</td>
</tr>
<tr>
<td>60</td>
<td>204</td>
<td>4.1</td>
</tr>
<tr>
<td>80</td>
<td>113</td>
<td>3.14</td>
</tr>
<tr>
<td>100</td>
<td>125</td>
<td>2.71</td>
</tr>
<tr>
<td>200</td>
<td>165</td>
<td>2.33</td>
</tr>
<tr>
<td>300</td>
<td>187</td>
<td>2.26</td>
</tr>
<tr>
<td>400</td>
<td>203</td>
<td>2.12</td>
</tr>
<tr>
<td>500</td>
<td>176</td>
<td>1.74</td>
</tr>
</tbody>
</table>

Table 3: Effect of etching conditions on FET performance

100 nanometer wide inversion mode devices

<table>
<thead>
<tr>
<th>Etch Condition</th>
<th>Electron Mobility (cm$^2$/V-s)</th>
<th>Subthreshold Swing (V/dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMAH 50°C</td>
<td>144.8</td>
<td>2.27</td>
</tr>
<tr>
<td>Cl$_2$ ICP</td>
<td>92.3</td>
<td>3.84</td>
</tr>
<tr>
<td>CF$_4$ RIE</td>
<td>54</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Individual devices were measured more thoroughly to determine $I_d$ vs $V_{gs}$ and $I_d$ vs $V_{ds}$ characteristics. We plot representative characteristics for devices having the smallest fabricated channel width: 40 nm for unsuspended inversion mode and 20 nm for suspended accumulation mode. The subthreshold characteristics shown in Figure 38 exhibit $I_{on}/I_{off}$ of about $10^5$ and subthreshold slope of 0.60 V/decade for the inversion mode device and $I_{on}/I_{off}$ of $2 \times 10^4$ and subthreshold slope of 1.32 V/decade for the accumulation mode devices.
Figure 38: Subthreshold characteristics of nanowire FET devices (a) Suspended accumulation mode device (b) Unsuspended inversion mode device.

It should be noted that the low-frequency noise is significant in these devices. This can be clearly seen in the output conductance measurements shown in Figure 39. We see that the noise amplitude appears to increase with the gate voltage. This is discussed in detail in section 5.2. The drain current characteristic exhibits saturation with an Early voltage that increases with increasing channel length as expected [90].

Figure 39: Output characteristics of back-gated nanowire FETs (a) Suspended accumulation mode device (b) Unsuspended Inversion mode device
4.4 Top-gate Device Characteristics

Top-gated devices were similarly screened on the Cascade automatic probe station. They were screened once using only the back-gate and a second time using only the top-gate. Again, we calculated values of field-effect mobility and subthreshold slope, this time for both gates for a number of devices of each length, width and type based on the screening data using the same methods described for the back-gate devices.

Table 4: Dual-gate nanowire FET parameters

<table>
<thead>
<tr>
<th>p-Channel</th>
<th>Mean Mobility (cm²/V·s)</th>
<th>Max Mobility (cm²/V·s)</th>
<th>Mean SS (V/dec)</th>
<th>Min SS (V/dec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Gated</td>
<td>158</td>
<td>199</td>
<td>0.090</td>
<td>0.072</td>
</tr>
<tr>
<td>Back Gated</td>
<td>127</td>
<td>220</td>
<td>0.70</td>
<td>0.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n-Channel</th>
<th>Mean Mobility</th>
<th>Max Mobility</th>
<th>Mean SS</th>
<th>Min SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Gated</td>
<td>461</td>
<td>559</td>
<td>0.31</td>
<td>0.128</td>
</tr>
<tr>
<td>Back Gated</td>
<td>278</td>
<td>403</td>
<td>0.58</td>
<td>0.22</td>
</tr>
</tbody>
</table>

The values of these parameters for both the top-gated and back-gated operation are given in Table 4. Because of the asymmetric geometry, there is little coupling between the front and back channels [91] so we did not exhaustively investigate the dual-gated operation. We also tested devices after suspending and the results are shown in Table 5.
Table 5: Suspended dual-gate FET parameters

<table>
<thead>
<tr>
<th>45 nm gate-width Suspended Top-Gated Nanowire FETs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>p-Channel</strong></td>
</tr>
<tr>
<td>Before undercut</td>
</tr>
<tr>
<td>After undercut</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>n-Channel</strong></th>
<th>Mean Mobility</th>
<th>Max Mobility</th>
<th>Mean SS</th>
<th>Min SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before undercut</td>
<td>461</td>
<td>559</td>
<td>0.31</td>
<td>0.150</td>
</tr>
<tr>
<td>After undercut</td>
<td>419</td>
<td>625</td>
<td>0.281</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Again, individual devices were measured more thoroughly to determine $I_d$ vs $V_{gs}$ and $I_d$ vs $V_{ds}$ characteristics. The subthreshold characteristics are plotted for unsuspended inversion mode accumulation mode devices with gate widths of 45 nanometers in Figure 40 and Figure 41, respectively. The devices exhibit considerably lower threshold voltage and subthreshold swing when they are operated with the top-gate than with the back-gate. This is expected because the top-gate dielectric is much thinner and thus the gate capacitance is much higher, minimizing the effects of the depletion capacitance and interface-state capacitance as described by equation 4.2. The p-channel (accumulation) devices have better characteristics than the n-channel (inversion) devices with lower subthreshold swing and better on-to-off ratio. This is likely due to the presence of surface states near the conduction band edge as we discussed in the previous section.
Figure 40: Subthreshold characteristics of unsuspended 45 nm gate-width inversion mode nanowire FET (a) Top-gated operation (b) Back-gated operation

Figure 41: Subthreshold characteristics of unsuspended 45 nm gate-width accumulation mode nanowire FET (a) Top-gated operation (b) Back-gated operation
The output characteristics for both devices are shown in Figure 42 and Figure 43. The disparity in drive-current between the two devices shown is due to a difference in channel length. The devices behave according to the classic MOSFET model in the linear and saturation regions however the early effect is much more pronounced in the n-channel devices.

![Figure 42: Output characteristics of unsuspended 45 nm gate-width inversion mode nanowire FET (a) Top-gated operation (b) Back-gated operation](image)

![Figure 43: Output characteristics of unsuspended 45 nm gate-width accumulation mode nanowire FET (a) Top-gated operation (b) Back-gated operation](image)
After suspending the devices by undercutting with BHF, the top-gated electrical characteristics remained quite similar. The subthreshold characteristic for the accumulation and inversion mode devices are shown in Figure 44. The subthreshold swing of both types of devices has increased slightly due to the increase in surface state density at the now-exposed back surface of the channel. This also causes the threshold and flat-band voltages to shift outwards although not drastically. The on-to-off ratio for the inversion mode device actually increases significantly when the devices are suspended. This is likely due to the decreased coupling of the channel to the back gate, affording the front gate better control over the volume inversion of the channel.

![Figure 44: Subthreshold characteristic of suspended top-gated nanowire FET](image)

The devices from the second top-gated process were also characterized in the same way. The subthreshold characteristic of a representative device is shown in Figure 45. The subthreshold swing is much higher than the top-gated devices shown previously and the average hole-mobility is about 30, considerably lower than the previous set of top-gated devices. This is likely because the gate dielectric used was LPCVD silicon nitride, which we expect to have a high density of interface states at the
silicon/silicon nitride surface. The silicon nitride was chosen for reasons related to the fabrication as discussed in chapter 3. It will likely be replaced with an ALD high-K dielectric in future fabrication runs.

![Figure 45: Subthreshold characteristic of accumulation mode top-gated nanowire FET fabricated with second method](image)

4.5 Analysis of DC Measurements

One of the most important goals in the fabrication efforts was to avoid severe degradation of the quality of the silicon films during processing. This was assessed using the mobility measurements, and also by low-frequency noise measurements as will be discussed in the next section. To determine the amount of degradation we needed a method of assessing the mobility of the pristine SOI films before processing. Appendix I: HgFET characterization of SOI Wafers describes a method for doing so, which we developed based on the HgFET reported by Hovel [92]. Based on this method, we calculated the field-effect mobility in the unprocessed films to be 750 cm²/V-s for electrons and 250 cm²/V-s for holes. Comparing that to the electron mobility of 625
cm²/V-s and hole mobility of 199 cm²/V-s measured in our fabricated devices shows that we have been quite successful in minimizing the process-induced degradation that we set out to avoid. This can also be ascertained from the measured sub-threshold slope, which is very dependent on the density of interface states at the silicon surface. We have produced top-gated devices with subthreshold slopes as low as 75 mV/decade, something that is only possible with a very low defect density according to equation 2.1. We observe much higher values of S for the back-gated devices with thick buried oxides because of the low oxide capacitance relative to the depletion and interface-state capacitances. We can compare the S of our back-gated devices to that of the HgFET and see that this parameter is not significantly degraded. The suspended devices show S that is only somewhat higher than the non-suspended devices indicating a small increase in interface states at the back surface of the channel. The threshold voltage, however, is very much affected by the processing. We observed that the threshold voltage was a function of device width, something to be expected as the exposed surfaces cause additional surface band bending that shifts out the flatband and threshold voltages. We see relatively little effect due to channel length which is to be expected as we are well above the length scale where short channel effects should occur. Subthreshold slope should have no length dependence and in fact it does not. Mobility, which also shouldn’t have a length-dependence actually appears to increase with length, however this is due to the decreasing effect of the series resistance as the channel resistance increases with length.

In summary, we were able to produce complementary nanowire FET devices with global- and local-gating and demonstrate that the electrical performance of these
devices was excellent. They showed very little mobility degradation and high subthreshold slope, which is necessary for highly sensitive sensors.
Chapter 5: Additional Characterization

Techniques

Although the DC characterization of the previous chapter supports the claim that we have produced high-quality transistors, it says very little about their performance as biomolecular sensors. To speak to this issue we have used two other means of characterization. First we have used the devices as pH-sensors to determine their sensitivity to variations in surface potential induced by surface-bound ionic charge. Secondly we have developed a method for characterizing nanowire FETs using low-frequency noise spectroscopy to extract a figure-of-merit that allows us to compare FETs of different geometry, composition and processing history. Because low-frequency noise is one of the most important limiting factors in ultra-high sensitivity molecular detection, this characterization technique provides insight into the effects of sample design and sample processing on ultimate sensitivity.

5.1 pH Sensing as a characterization tool

As we discussed in Chapter 2 nanowire FETs can be used to detect changes in the pH of a solution. Using nanowire sensors to measure solutions of known pH can provide a good metric for characterizing the sensitivity of a sensor device to surface charge. A perfect sensor would exhibit a sensitivity, defined as \( S = \Delta I_{ds}/I_{ds} \), of 10 for a change in pH of unit, or 1 decade/pH. This would require a perfect FET device, having a subthreshold slope of 59mV/decade, and Nernstian gate insulator, having a response of -59mV/pH. In practice, a good sensor would exhibit a sensitivity well below this
theoretical maximum due to the presence of surface states at the silicon/insulator interface, ionic screening in the solution, mobile oxide charge and the intrinsic buffer capacity of the dielectric. For a nanowire FET with a silicon dioxide dielectric such as our back-gated devices, the subthreshold slope is worsened by the diffusion of mobile ions into the exposed dielectric [38-40, 45-47, 93]. In practice, 100 mV/dec subthreshold swing and an $\alpha$=0.7, where $\alpha$ represents the deviation from Nernstian behavior due to the intrinsic buffer capacity of the dielectric, are roughly the best we expect. This suggests that a response of about 0.42 decades/pH is the highest pH sensitivity we should expect when using a silicon dioxide gate dielectric.

We used back-gated FETs as pH sensors in order to characterize their sensitivity to surface charge. We compare the response of planar devices with a channel width of 1 um to nanowire devices with a width of 100 nm to determine whether the response is size dependent. We also compare p-channel accumulation mode devices with n-channel inversion mode devices. Finally, we use pH sensing to compare microfluidic solution transfer with our previous open-reservoir method for solution exchange, a result that is discussed in chapter Chapter 6:

The back-gated devices were prepared for sensing by gluing a piece of Tygon tubing onto the surface of the chip with cyanoacrylate glue to serve as a fluid reservoir, as shown in Figure 46.
This is discussed further in chapter Chapter 6:. The devices were contacted on a manual probe station and measured using an Agilent 4156B semiconductor parameter analyzer. The devices were first screened dry to verify their $I_d \text{ vs } V_g$. Satisfactory devices were then exposed to a 1X solution of Dulbecco’s phosphate buffered saline at pH 7 and the $I_d \text{ vs } V_{reference}$ was measured at $V_{ds} = 1V$. The back-gate voltage could then be adjusted to maximize the subthreshold slope and operating range as shown in Figure 47. The optimal back-gate bias condition in this device was determined to be -3V. These improvements in front channel performance were due to the asymmetric nature of the dual-gate SOI system [91, 94]. Under weak back-gate bias, the top-gate, or in this case the solution acting as a top-gate controls most of the carrier density in the channel. However, at high enough back-gate bias, a back-channel can form that the front-gate cannot control sufficiently well to fully turn it off. Once a suitable back-gate voltage was chosen, the reference electrode voltage was set to bias the device into the subthreshold
region where sensitivity is maximized. If higher currents are needed to improve signal to noise ratio, the reference voltage can be brought closer to the threshold voltage; however, the sensitivity will decrease.

Figure 47: Drain current versus reference electrode voltage at varied back-gate voltages demonstrates that front channel response can be tuned by adjusting back-gate voltage to decouple front and back channels.

The current as a function of pH was determined by measuring the current versus time using the 4156B in sampling mode. $V_{\text{reference}}$, $V_{\text{backgate}}$, and $V_{ds}$ were held constant over a period of 2000 seconds as the solutions were exchanged by manual pipetting or microfluidic flow. A typical time recording of $I_d$ is shown in Figure 48 (a). From this recording, the average current value at each pH level was determined and plotted to give the characteristic curve shown in Figure 48 (b).
The sensitivity was determined from these plots by dividing the relative current change by the pH change. The average sensitivity values that we calculate between pH=7 and pH=9 is 0.35 decades/pH, which is good compared to the other values reported in the literature, and compared to the maximum expected value of 0.42. We can determine whether the decreased sensitivity as compared to the theoretical maximum is due to degraded FET performance or decreased surface sensitivity by again observing the $I_d$ vs $V_{reference}$ characteristic. We can check the subthreshold slope at a given voltage to determine the FET sensitivity. We observed subthreshold slopes as low as 90mV/decade in solution with a mean value of 150mV/decade. 59mV divided by the actual subthreshold slope represents the FET contribution to the degradation, which we have termed $\alpha_f$. The actual change in conductance multiplied by the relative change in pH, divide by the measured subthreshold slope represents the change in surface potential due to the pH change. This surface potential change is divided by 59mV to
give a degradation factor corresponding to the deviation of the surface from perfectly Nernstian behavior, $\alpha$. The actual sensitivity of the device is equal to $\alpha_f \times \alpha$.

We also compared n-channel inversion mode devices to the p-channel accumulation mode devices already discussed. As we discussed in chapter Chapter 2:, an increase in pH causes a negative change in surface potential according to the Nernst equation. This can be viewed as a positive shift in threshold voltage, which is given by:

$$V_t = E_{\text{ref}} - \psi_0 + \chi^{\text{sol}} - \frac{\Phi_{\text{Si}}}{q} - \frac{Q_{\text{ox}} + Q_{\text{ss}} + Q_{\text{B}}}{C_{\text{ox}}} + 2\phi_f$$  \hspace{1cm} 5.1

For the n-channel FET, where $I_d$ is proportional to $V_{gs} - V_t$, this will result in a drain current that decreases with increasing pH. For the p-channel FET, where $I_d$ is proportional to $V_t - V_{gs}$, the opposite effect occurs and current increases with increasing pH. Figure 49 shows the pH response of an n-channel and a n-channel device to the same changes in pH. Note that the magnitude of the current changes are of comparable scale suggesting the sensitivity of the devices is similar; however, the sign of the current change is opposite for the two types of devices.

![Figure 49: pH response of complementary devices](image)

91
Table 6 lists the measured values for sensitivity, subthreshold slope as well as $\alpha_f$ and $\alpha$ for the p-channel and n-channel devices.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Subthresh. slope</th>
<th>$\alpha_f$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-channel</td>
<td>0.35 dec/pH</td>
<td>155mV/dec</td>
<td>0.38</td>
<td>0.92</td>
</tr>
<tr>
<td>n-channel</td>
<td>0.30 dec/pH</td>
<td>170mV/dec</td>
<td>0.31</td>
<td>0.97</td>
</tr>
</tbody>
</table>

These values of $\alpha$ are much higher than expected for a silicon dioxide ISFET. This may be because the thin native oxide layer is somewhat permeable to ions because of its porosity and thus it cannot be represented as a true ISFET gate insulator.

For some of the devices with small widths (100 nm) we observed anomalously high pH sensitivity. For these devices the mean sensitivity over the most sensitive range (pH=7 to pH=8) was 0.64 dec/pH. The subthreshold slope, however, was not significantly higher than for the devices with much higher channel width. This would suggest that the change in surface potential due to the pH change is super-Nernstian, something without a convenient physical explanation.

The majority of the device response to changes in pH occurs within a few seconds. It should be noted that the current levels determined by this measurement when going from high pH to low pH do not exactly match the levels when going from low pH to high pH. This is due to two effects which have been observed and studied in ISFET pH sensors: hysteresis and drift [39, 41, 43-46, 51, 59, 93]. ISFETs were shown to have a biphasic response with an initial response time of less than 1 second followed by
a second phase characterized by a much longer time constant of minutes to hours [41, 43-45, 50, 51, 59, 93]. This slow response leads to a very large hysteresis effect in SiO₂ gate ISFETs which has been reported to be as large as 25 percent of the total response [39, 51]. The second effect is a monotonic drift that is independent of pH. This has also been observed in ISFETs with a variety of gate materials; however, it is most evident in SiO₂ gate devices. The drift is thought to be a result of the combined effects of protonation of buried OH- groups in the bulk of the gate insulator, diffusion of counter ions into the gate oxide as well as physical etching of the surface by OH⁻ ions in the electrolyte. Due to these non-ideal effects, bare SiO₂ was largely abandoned as a gate insulator for ISFETs decades ago as more ideal materials like Al₂O₃ and TaO₅ became available. The device in Figure 48 exhibited a fairly small drift and hysteresis effect; however, these effects are much more evident in some other devices. Figure 50 (a) and (b) show the time recording and current vs. pH plot for a device exhibiting significant hysteresis. Figure 50 (c) shows a device exhibiting current drift. Figure 50 (d) shows a fairly typical measurement when both effects occur.
Figure 50: Nonideal effects in nanowire field-effect pH Sensors  (a) Current sampling vs. time: pH varied from 5 to 9.5 and back in 0.5 pH steps  (b) Current versus pH curve for same device shows clear hysteresis. (c) Time recording of current during repeated changes of 1 pH unit shows monotonic drift effect. (d) Typical devices exhibit both hysteresis and drift.

5.2 LF Noise Spectroscopy as a Diagnostic

Signal-to-noise ratio (SNR) is one the critical factors that may hinder nanowire field-effect sensors from fulfilling their promise of ultimate sensitivity. There can be many sources of noise in a typical sensing measurement but the most fundamental and perhaps most important is the noise intrinsic to the device itself. At low frequencies, this intrinsic noise has been empirically determined to be inversely proportional to
frequency and is thus referred to as 1/f noise [95]. 1/f noise has been studied extensively in MOSFETs and its main source in silicon devices has usually been considered to be conductivity fluctuations caused by charge trapping/detrapping at the silicon-insulator interfaces. The physical cause of these fluctuations may be either mobility fluctuations or carrier population fluctuations and this is still a source of debate [96]. The level of 1/f noise is highly dependent upon device processing conditions as these determine the density of interface states, which are the predominant locations of charge trapping. As such, a quantifiable measure of 1/f noise can provide a useful gauge of device quality. This is especially true for nanowire devices where the surface to volume ratio is very high and thus 1/f noise is likely to be high. 1/f noise has also been studied in planar ISFETs where it has been shown that the low-frequency noise characteristic is dominated by the 1/f noise due to the Si-insulator interface and not the insulator-solution interface as might be expected [48, 49]. Thus, the quality of the FET device is expected to limit the SNR of a field-effect sensor before the solution-induced noise poses a challenge.

The Hooge parameter, $\alpha_H$, is an abstraction of the 1/f noise, which historically has been used as a measure of FET device quality. $\alpha_H$ is defined by the equation:

$$\frac{S_I}{I^2} = \frac{\alpha_H}{fN} \tag{5.2}$$

where $S_I$ is the noise power spectral density (PSD), $I$ is the current through the device, $f$ is the frequency, and $N$ is the density of charge carriers in the device. Recently, a few studies of 1/f noise have been performed on nanowires [97-99], and have yielded Hooge parameters considerably higher than the typical values for commercial MOSFET devices, indicating poor device performance as compared to commercial MOSFETs.
Since interface states play a large role in 1/f noise and are also likely very important in determining the ultimate sensitivity of FET sensors due to their negative effect on subthreshold slope, we have proposed that the Hooge parameter should be a useful diagnostic in predicting the ultimate sensitivity of nanowire sensors. We have measured the Hooge parameter for a number of nanowire FET devices to determine the various processing conditions on 1/f noise.

The experimental setup used for the noise measurements was nearly identical to that which is used for time-dependant sensing measurements. The nanowire FET was contacted using a manual probe-station within a light-tight faraday cage to minimize extrinsic noise due to electrostatic interference or variations in illumination. The source end of the nanowire was grounded while drain was held at a fixed voltage, $V_{ds}$. The gate of the nanowire (either top-gate or back-gate, depending on device geometry) was held at a fixed voltage, $V_{gs}$, and the drain current was measured directly as a function of time. The voltages were sourced and currents measured using a Keithley 2612 dual channel SMU controlled by a custom-written Labview program. The devices were allowed to stabilize for several hours after the voltages were applied before beginning the measurements. The noise power spectral density (PSD) was extracted from the drain current with Matlab's “pwelch” algorithm, which uses an averaged, modified-periodogram method of spectral estimation.

We first demonstrate that the nanowire FET devices follow equation 5.2 by measuring the noise as a function of channel current at a constant carrier density (constant gate voltage). Figure 51 shows that the normalized noise PSD is constant with respect to current suggesting that the Hooge equation accurately describes our devices.
This, however, gives no information regarding the physical mechanism generating the noise.

**Figure 51:** Noise versus drain current  
(a) Normalized noise PSD at varied $V_{ds}$  
(b) Constancy of noise amplitude with respect to channel current

Next, the dependence of the noise on carrier density is extracted by measuring the noise PSD as a function of $V_{gs}$. The dependence of the noise PSD on gate voltage is clearly evident in Figure 52.

**Figure 52:** Normalized noise power spectral density versus frequency at varied $V_g$
Since the device is biased into the linear region of operation and $V_{ds}$ is small compared to $V_{gs}$, we can use the conventional MOSFET equations to calculate the average carrier density in the channel. $N$ is given by equation 5.3.

$$N = \frac{C_{ox}}{q} [V_G - V_T]$$ \hspace{1cm} 5.3

$$\frac{1}{A} = \frac{C_{ox}}{q\alpha_H} [V_G - V_T]$$ \hspace{1cm} 5.4

The gate capacitance $C_{ox}$ used for these calculations was determined from the Silvaco simulations described in section 4.2, setting $V_G = V_T$. From the plot of $1/A$ vs $V_G$, shown for a representative device in Figure 53, we can extract a value for $\alpha_H$.

![Figure 53: Reciprocal noise amplitude versus gate voltage is proportional to $1/\alpha_H$.](image)

Table 7 lists the values of $\alpha_N$ for our top-gated device as a function of gate width. There is no clear trend in $\alpha_N$ as a function of gate width. Because of each measurement takes many hours to complete, we were not able to take enough
measurements to make a statistically significant comparison between the various gate widths. However, the values of $\alpha_N$ that we obtain generally compare very favorably with those reported for other nanowire FETs and commercial MOSFETs [97-100].

**Table 7: Hooge parameter values for various nanowire FETs**

| Device Type | Gate/Channel Width | Length | Etch Type | $|\alpha_H|$ |
|-------------|--------------------|--------|-----------|----------|
| Top-gate    | 45 nm              | 44 µm  | TMAH      | 2.2 x 10^{-3} |
|             | 95 nm              | 44 µm  | TMAH      | 4.8 x 10^{-4} |
|             | 195 nm             | 44 µm  | TMAH      | 2.1 x 10^{-3} |
| Back-gate only | 100 nm         | 3 µm   | TMAH      | 3.0 x 10^{-3} |
|             | 100 nm             | 3 µm   | Cl₂ ICP   | 1.4 x 10^{-2} |
|             | 100 nm             | 3 µm   | CF₄ RIE   | 1.6 x 10^{-2} |

We have also investigated the effects of the etch process used to define the nanowire pattern. We compared three different etch process: an orientation-dependant wet etch using 25% tetramethylammonium hydroxide, a Cl₂ ICP process, and a CH₄ RIE process. The devices used for this comparison were back-gate only devices processed together on a single wafer as described in section 3.2.

At the point where the HSQ pattern was transferred into the active silicon, the three sets of devices were processed in sequence without dividing the wafer. Two-thirds of the devices were protected with photoresist while the native oxide was stripped from the remaining devices, which were to be etched with TMAH. The remaining native oxide was sufficient to prevent the TMAH from attacking the other devices during the 10 minute etch at 50°C. Next, the TMAH-etched devices and half the remaining devices were protected with photoresist while another one third of the devices were etched in CF₄ in the Oxford Plasmalab 80+ RIE (Yale). Finally, the TMAH-
etched devices and CF$_4$-etched devices were protected with photoresist while the final third of the devices were etched in a Cl$_2$ ICP process in the Oxford Plasmalab 100 (Yale). All subsequent processing was done on the entire wafer and followed the steps outlined in section 3.2. The devices were not passivated or annealed.

The $I_d$ vs $V_g$ characteristics of the three sets of devices were measured shortly after processing (1-2 days) and demonstrated significant hysteresis due presumably to the unpassivated and unannealed surfaces which would be expected to have many dangling bonds. This was especially pronounced for the TMAH etched devices. The hysteresis decreased significantly over a period of 2 weeks in air. The characteristics immediately after fabrication and after several weeks are shown in Figure 54.

![Figure 54: Hysteresis in devices used for noise spectroscopy (a) TMAH etched devices (b) Cl$_2$ ICP etched devices](image)

Even though the hysteresis was higher for the TMAH-etched devices, the Hooge parameter was measured to be significantly lower as shown in Table 7, suggesting that the traps which cause hysteresis in these devices do not significantly contribute to $1/f$ noise. The lower $\alpha_N$ confirms our long-held hypothesis that the surfaces produced by
TMAH-etching have lower surface state density than RIE patterned surfaces, leading to lower noise and in-turn better sensor performance.
Chapter 6: Microfluidics for Sensing

6.1 Design Considerations

One of the most challenging practical aspects of biomolecular or chemical sensing is finding an effective method of delivering the fluid that is to be analyzed to the surface of the chip which doesn’t impede the electronic measurements that must be performed. Ideally, these measurements are to be performed during the fluid transfer. The combination of electrical measurements and fluidic transfer has prompted several different approaches in the sensor community; however, none has been completely satisfactory.

There are several characteristics to be considered when designing a fluid delivery system and many of these are seemingly mutually exclusive. Some of these characteristics are mandatory, while others are optional. Free access to electronic contacts is a non-negotiable characteristic of a useable fluid transfer system. If the electronic probing or bonding pads cannot be accessed, the system is useless. Further, the pads and wires or fan-in must be protected from the fluid to prevent the formation of a short circuit through the fluid. Another property that would be very useful is a high fluid mixing rate that is caused by turbulent flow. As discussed in chapter 2, microfluidic channels that demonstrate laminar flow should be diffusion limited causing low sensitivity and very slow response at low analyte concentration. In practice, this means that if the fluidic transfer system consists of a microfluidic channel formed at the surface of the chip, the channel width and height must be around 500 um or greater [84,
In the past, our lab has avoided this problem by using an open-topped reservoir permanently attached to the surface of the chip.

The open reservoir has been used successfully but is very tedious and presents a number of problems of its own. Before the reservoir is attached, the surface of the chip is protected using a lithographically patterned polymer film such as SU-8, polyimide or even photoresist. Each chip is then prepared separately by the following process: A 2-3 mm long piece of PTFE or Tygon tubing is cut with a razor blade. The end of the tubing is dipped in epoxy or cyanoacrylate glue and under a dissecting stereoscope the tubing is manually placed onto the surface of the chip. Great care must be taken to ensure that the tubing sits within the pad frame and that none of the pads or devices is covered by the glue. Figure 55 shows a micrograph of a chip with a tygon tube glued in place.

Figure 55: Nanowire FET chip prepared for sensing with Tygon reservoir

The main drawbacks of this method are the time involved in preparing chips, the permanent and destructive nature of the method, the yield and the issue of
chemical/bio-compatibility of the glue. Each chip takes a few minutes to prepare. Once the tubing is glued to even a single chip on a wafer, the wafer can no longer be subjected to additional processing steps in the cleanroom because complete removal of the glue is nearly impossible. When positioning the tubing on the chips, an extremely delicate procedure, it is not unusual to contact the devices with glue which renders them useless. With experience, the yield for this process is about 80-90%. Neither epoxy nor cyanoacrylate glue is completely biocompatible [102] so this may interfere with some types of cellular measurements, though both are fairly chemically inert in aqueous solutions so macromolecular sensing may not be affected.

Another drawback of the open reservoir method is encountered when one tries to actually use this type of approach. The sensing experiments are typically done on an electrical probe station under a stereomicroscope. The chip is placed on the platen of the probe station and the pads are contacted via micromanipulator probes. Initially, the solution is introduced by manually pipetting into the open top of the reservoir. This must be done rather forcefully to ensure displacement of air bubbles at the solution/chip interface. Care must be taken to avoid overflowing the reservoir, which only holds a few microliters, as this will cause a short circuit between the pads or probe tips. When a new solution is to be introduced, the old solution must be aspirated so as not to overflow the reservoir. This has been accomplished by using an absorptive material to draw off the old solution as the new solution is introduced by pipet. If the solution is aspirated too fast, the surface will become dry, ruining the experiment. Because of the delicate nature of this operation and the variation from chip to chip, it would be prohibitively difficult to automate this process using the open reservoir method.
We have developed a new tool for fluidic transfer that addresses many of the shortcomings of the previously employed methods. This tool, which we call the *microfluidic probe* is analogous to an electrical needle-probe in that it is attached to a micromanipulator for use on a wafer-probing station. The *microfluidic probe*, shown in use in Figure 56 allows us to form a non-permanent, sealed, microfluidic channel at the surface of the chip.

![Image of microfluidic probe](image.png)

**Figure 56:** *Microfluidic probe* allows non-destructive microfluidic and electronic probing of individual dice on a wafer.

The channel is large enough to ensure turbulent conditions and thus a fast device response; however, it does not interfere with the electrical probe pads like a typical PDMS “peel-and-stick” approach, and can be employed on chips as small as 2mm x 2mm. This represents a decrease in die area of at least a factor of 25 and a corresponding increase in number of sensing experiments per wafer. Considering that
the cost of a 4” wafer of nanowire devices may be several thousand dollars, this represents a significant savings. The probe is also completely non-destructive. It requires no permanent modification of the wafer and wafers screened using it may undergo further cleanroom processing after fluidic probing.

The fluidic probe tip is entirely made of the biocompatible materials polystyrene, PDMS, and PTFE. It may be reused many times and can be set up on a chip just as quickly as an electronic probe tip. We have also incorporated the fluidic probe tip into an automated wafer-prober which uses a probe-card to all 32 pads on the chip simultaneously. This gives us the option of high-throughput electrical and fluidic screening of devices. This device is shown in Figure 57.

Figure 57: Microfluidic probe for automatic probe station allow rapid fluidic screening at the wafer scale.
6.2 Microfluidic probe designs

Three different designs for the microfluidic probe were developed, each with a slightly different capability. The first of these, as described above, requires no modification of the test sample; the flow channel is formed temporarily at the interface between the probe tip and the sample surface. The second and third methods require extra lithographic patterning of the surface. A schematic comparison of the three methods is shown in Figure 58.

Figure 58: Three approaches to microfluidic probing: (a) PDMS gasket surrounding recessed channel. (b) PDMS gaskets interface with ports on SU-8 buried channel. (c) flat bottomed PDMS seals against top of SU-8 walls.
Method 1 consists of a piece of PDMS elastomer with a molded impression in the bottom surface. This impression is surrounded by a raised ridge which acts as a gasket and forms a water-tight seal around the channel. This gasket is easily deformed with minimal pressure and can thus form a seal even on a chip with moderate surface topography. Two holes in the PDMS serve as an inlet and outlet to the channel. This method is most desirable when the sample cannot be modified.

In the second method, the channel is not formed at the interface but is instead formed on the test wafer itself as a self-contained buried channel structure. This buried channel has an inlet and outlet that mate with inlet and outlet ports on the fluid probe. Because the channel is a rigid permanent structure, it is able to withstand higher pressure than the other methods. Thus this method is most useful for experiments where high pressures are required to achieve extremely high flow rates.

The third method forms the flow channel at the interface between the probe and the sample; however, it also requires patterning of the test sample. The walls of the channel are formed as rigid, permanent structures on the sample surface and the ceiling of the channel is formed by the bottom surface of the fluidic probe. Because of the simplicity of the probe in this case, it can allow for smaller chip size than the other methods.

In each case, the pyramid shaped PDMS probe tip is bonded to a rigid polystyrene or glass support into which tubing is inserted to deliver and remove the fluid. These materials are chosen in part because they are optically transparent, allowing visual inspection and alignment of the device through the probe.
6.3 Microfluidic probe fabrication

The fabrication of each of the three designs is slightly different but the steps can generally be grouped into overall processes: sample patterning, PDMS Molding, and probe assembly.

Method 1:

In this case, no additional patterning of the sample is required. The process to fabricate the microfluidic probe is shown in Figure 59.

![Diagram showing fabrication process](image)

**Figure 59: Outline of fabrication process for method 1 microfluidic probe**
A mold for the PDMS is fabricated on a 4-inch silicon wafer using SU-8 2150 (Microchem) an epoxy-based negative UV photoresist that can produce features up to 1 mm high with a single coating. Photomasks for all three methods were written in the Yale Cleanroom on a TRE Electromask CC-251 optical pattern generator and all exposures were performed on an EVG-620 contact aligner in soft contact mode with a 350 nm low-pass filter to remove deep UV and reduce T-topping. To yield the two different thicknesses of SU-8 required to mold both the channel and the surrounding gasket, a two step lithographic process was required. The first SU-8 layer which forms the pattern for the gasket was spin-coated and soft-baked to produce a 100 um thick layer. The pattern for the first layer is shown in Figure 60 (a).

![Image of mask patterns](image)

**Figure 60:** Mask patterns for microfluidic probe method 1 (a) First mask patterns gasket. (b) Second mask patterns channel.

The SU-8 was exposed and post-exposure baked to crosslink the resist. Then without developing, a second layer of SU-8 2150 was spun and soft-baked to yield 300-500 um of resist. This layer was exposed with the mask for the channel pattern, shown in Figure 60 (b), using the latent image from the first layer to align the exposure. After baking to crosslink the second layer, the entire structure is developed to yield the mold structure shown in Figure 61. We found it necessary to increase the bake times given
on the photoresist datasheet by a factor of 2 to 4 to produce repeatable results. The SU-8 is annealed at 150°C for 10 minutes to remove micro-cracks and to smooth out rough edges.

![SU-8 mold](image)

**Figure 61: Two layer SU-8 mold for microfluidic probe method 1**

Prior to using the mold, it is vapor primed with chlorotrimethylsilane, which prevents the PDMS from sticking to the SU-8 on the mold and tearing when it is removed. PDMS elastomer (Silgard 184, Dow Corning) is mixed in a ratio of 10:1 elastomer base to curing agent by weight. The mixture is stirred thoroughly and degassed in a vacuum desiccator for 30 minutes to remove bubbles. The PDMS is then poured over the mold to a thickness of 5-7 mm and is again degassed to remove bubbles trapped between the PDMS and mold. Finally, the elastomer is cured for 30 minutes at 80°C in a vacuum oven. The cured PDMS casting is peeled off of the mold which may be reused. The PDMS is cut with a razor blade into the pyramid shapes shown in Figure 62.
The rigid supports for the PDMS may be formed from either glass or polystyrene. Glass is desirable because it allows the probes to be autoclavable; however, polystyrene is much easier to machine so after trying both we settled on polystyrene. The polystyrene is cut from standard Falcon petri dishes into 1 cm squares using a hot razor blade. Two holes are drilled in each to accept inlet and outlet tubing. The PDMS is then bonded to the polystyrene using an oxygen plasma bonding technique. Both the PDMS and polystyrene are exposed briefly to a low-power oxygen plasma (MCS HT-6, Yale: 75W, 12 seconds). The exposed sides are immediately pressed together and placed into a 70°C oven for 1 hour. This forms an irreversible covalent bond between the PDMS and polystyrene. Using a modified syringe needle, holes are cored in the PDMS to connect the channel with the inlet and outlet holes in the polystyrene support. Teflon tubing is then inserted into the holes in the PS and secured with epoxy. The *microfluidic probe* tip is mounted on a machined tip holder which connects it to the micromanipulator as was shown in Figure 56. Figure 62 shows the finished PDMS probe.

![Figure 62](image)

*Figure 62: Method 1 microfluidic probe ready for use*
Method 2:

In this case we fabricated buried channels on the surface of the wafer as well as a matching *microfluidic probe*. The overview of the process is shown in Figure 63.

**Figure 63:** Overview of fabrication process for method 2 *microfluidic probe*
Again we used SU-8 resist, this time as the structural material for the channel walls and ceiling. SU-8 2150 was spun and baked to yield a layer that is 350-600 microns thick. This layer was exposed using the pattern shown in Figure 64 (a) to produce the sidewalls of the channels. The resist was post-exposure-baked just long enough for the latent image to appear. A second exposure was performed to define the ceilings of the channels using the pattern shown in Figure 64 (b). In this case the dose of the exposure was drastically reduced and the Low-pass filter was removed from the mask aligner. Because of the high absorption of deep UV light by SU-8, we found it possible to expose and crosslink to top 50 um of SU-8 and leave the remaining 300 um unexposed. After baking to fully crosslink all of the exposed SU-8, we develop the unexposed resist leaving the buried channel with inlet and outlets as shown in Figure 65 (a).

![Figure 64: Mask patterns for buried microfluidic channels](image)

The PDMS molding process is the same as method 1 except that the pattern is simpler. Instead of the channel surrounded by a gasket, we simply need two round gaskets, which will surround the inlet and outlet ports on the surface of the SU-8 coated
sample. This requires only a single 100 um thick SU-8 pattern on the mold. Again, the PDMS is cast against the mold, cured, and cut with a razor blade to produce the structure shown in Figure 65 (b). The assembly of the PDMS/PS/tubing structure is exactly the same as in *method 1*.

![Figure 65: Method 2 microfluidic probe](image)

(a) SU-8 buried channel  (b) Gaskets that interface with buried channel

**Method 3.**

This method is something of a combination of the previous two, although it is simpler than either. The overview is shown in Figure 66.
The sample wafer is prepared by fabricating a 300-500 µm high ridge in SU-8 that will serve as the walls of the flow channel. As with the molds for the other 2 methods, the SU-8 is spun and soft-baked to yield a 300-500 µm coating. It is exposed using a mask with the pattern shown in Figure 67 and baked to crosslink the resist. The
structure is developed and annealed to remove cracks. The SU-8 ridge is shown in Figure 68 (a)

![Figure 67: Mask pattern for SU-8 ridge](image)

The PDMS is not patterned at all in this case. A flat sheet of PDMS is poured out on a wafer, cured and cut with a razor. The assembly proceeds as before except that the two holes cored in the PDMS are now not surrounded by a gasket or channel. They are simply punched close together on the flat bottom surface of the PDMS as shown in Figure 68 (b).

![Figure 68: Method 3 microfluidic probe](image)

(a) SU-8 ridge surrounds devices. (b) Flat bottomed probe forms ceiling over ridge.
6.4 Microfluidic probe testing

A typical use of the microfluidic probe on a manual probe station involves positioning several electrical probe tips on the wafer and then bringing the microfluidic probe into contact with the surface until the PDMS is slightly deformed. This ensures that a seal is formed at the PDMS/sample interface. To determine the quality of this seal and compare the performance of the three designs, we pressure-tested each design. In each case, the probe tip was positioned on a bare silicon wafer placed on a digital balance which was used to measure the initial force exerted on the wafer by the deformed PDMS of the probe tip. The outlet tube was then clamped and the inlet tube connected to a syringe. The force on this syringe was measured and divided by the barrel cross sectional area to determine the pressure in the fluid line. The syringe was gradually compressed, increasing the internal pressure until leakage occurred at the wafer/PDMS interface. In many cases, the PDMS/PS bond failed before leakage occurred at the surface. The internal pressure required to induce leakage was measured for a range of forces applied to the surface. Figure 69 shows the pressure versus force for each design. As expected, the buried channel is able to withstand the highest pressure before leaking. This pressure is equal to about 1 atmosphere at a downward force of 12.5 grams.

In the automated probe station we set the height of the fluidic probe tip slightly below the height of the electrical probe tips. When the platen is raised to bring the sample into contact with the pins, the PDMS contacts the wafer first and is slightly deformed. A 50 um height difference was sufficient to produce repeatable non-leaking seals.
Figure 69: Comparison of leakage in fluidic probes. Internal pressure required to cause leak is plotted versus downward force applied to surface.

We used the microfluidic probe tip in a number of the pH tests described in chapter Chapter 5:. For these tests the syringes were switched and pushed manually, however it would be trivial to connect the syringes to a system of electronically controlled valves and syringe pumps to automate this process. Figure 70 shows results of a typical pH sensing test performed using the microfluidic probe.

Figure 70: Sampling measurement of drain current at various pH using microfluid probe method 1
For each step, we change syringes and flow a total of 100 microliters of solution through the channel. The fluid exchange takes about 5 seconds and the sensor responds rapidly. We see similar response whether we use the microfluidic probe or the open-reservoir method.

We envision the microfluidic probe become a common tool in the electronics laboratory as it provides an easy way to unify microfluidic and microelectronic capabilities. In the automatic prober it could enable high-throughput fluidic screening of devices for their sensing properties in the same way that we currently screen devices for their electrical characteristics. Other labs employ technicians to do this work.

6.5 Microfluidic surface patterning

Another use that we have envisioned for the microfluidic probe is chemically patterning or selectively functionalizing surfaces. While qualitatively similar to microarray printing, this technique would allow arbitrarily shaped patterns such as spots, rings, lines, and serpentines. It would also allow a variety of chemical surface modifications and allow for rinsing off non-bonded reagents to produce clean monolayer coatings, unlike fast-printing techniques that deposit large quantities of material. To validate this patterning technique we used a well-known conjugation method to produce fluorescently labeled patterns on a glass slide.

Patterns with minimum dimensions larger than a few hundred microns could be prepared in exactly the same way as method 1 for the fluidic probe tips. The limitation to further scaling of that technique arises from the difficulty in coring the inlet and outlet holes to access the channel. Because of this, a section at each end of the channel was required to be larger than the diameter of the needle used to core the holes.
Devices with much smaller channels were produced by lithographically patterning the inlet an outlet ports. We were able to produce devices with 75 micron wide inlet and outlet ports and 75 micron wide channels using the process shown below.

A *microfluidic probe* was prepared with a 3 separately-addressable arbitrarily shaped channels by the microfabrication process shown in Figure 71.

![Fabrication process for microfluidic surface patterning probe tip](image)

**Figure 71: Fabrication process for microfluidic surface patterning probe tip**

This process produces two separate layers of PDMS which are then aligned and bonded, before bonding the entire structure to a polystyrene support layer as in the fluidic probe tip fabrication.
The first PDMS layer contains the channels/chambers which define the shape of the pattern to be printed. This layer also has inlet and outlet ports. A mold is formed by a two layer SU-8 process similar to the one used in method 1 above. The first layer, which is formed in 100 micron thick SU-8 2035, is patterned using a mask corresponding to the pattern of the channels, shown in Figure 72 (a).

![Figure 72: Mask patterns for patterning probe device](image)

(a) Channel patterns with ports (shaded regions)  
(b) Fan-in allows access to closely spaced channels.

This pattern is exposed, fully post-baked and developed. A second, 1 mm thick layer of SU-8 2150 is prepared on the wafer and exposed with a mask that defines the inlet and outlet ports, corresponding to the shaded regions in Figure 72 (a). This layer is baked and developed to yield 1 mm high pillars corresponding to the ports. When a layer of PDMS thinner than the height of these pillars is cast against this mold, the pillars extend all the way through the PDMS and form through-holes. The channels and through-holes are visible in an optical micrograph of the PDMS casting is shown in Figure 73.
Because the ports are too close together to connect to separate holes in the PS support layer, a separate fluidic fan-out layer is required. This layer is cast in PDMS by using a single layer SU-8 2150 mold, 200 um thick, to produce a network of channels leading outwards from the ports corresponding to the pattern shown in Figure 72 (b). The PDMS is cut into individual devices using a razor blade. The fan-out layer is plasma-bonded to a polystyrene support layer that has holes drilled corresponding to the ends of the fan-out pattern. These holes are extended through the PDMS layer into the fan-out channels by coring with a syringe needle.

The last step is to align and bond the two layers of PDMS together. The two layers are exposed briefly to a low-power oxygen plasma (MCS, Yale), then a small amount of methanol is applied to the surface of the fan-out/PS assembly. The first PDMS layer is then placed on this surface and under a microscope the two layers are brought into alignment. The methanol prevents immediate binding of the two layers and allows them to be repositioned until satisfactory alignment is achieved. The assembly is then baked at 70°C for an hour to promote binding of the PDMS layers. The two layers are shown after alignment and bonding in Figure 74.
Figure 74: Channel layer has been plasma bonded to fan-in layer to form completed flow-channel network.

The PDMS/PDMS/PS layer is fitted with tubing as in the regular *microfluidic probes* and is mounted on a tip holder for attachment to a micropositioner or the automated wafer prober. The final assembly is shown in Figure 75.

Figure 75: Completed patterning device  (a) Fan-in connects cored holes to channels.  (b) Final assembly with tubing and bracket for mounting on micropositioner
This probe tip was used to selectively functionalize the surface of a glass slide. The slide was cleaned with piranha solution (3:1 H$_2$SO$_4$ : H$_2$O$_2$) and dehydrated in a vacuum desiccator for 1 hour. The entire slide was immersed in 1% 3-aminopropyltriethoxysilane (APTES) in anhydrous toluene in a nitrogen glove-box for 1 hour. The APTES binds covalently to the glass and produces a surface with primary amine functionality. The slide was rinsed in fresh toluene and dried in an oven at 70°C for 1 hour. The effectiveness of the functionalization could be observed from the dramatic change of the surface from hydrophilic to hydrophobic.

The microfluidic patterning probe was positioned over the surface of the amine-functionalized slide and brought into contact using the micropositioner. Fluorescein isothiocyanate (FITC), a green fluorescent label that binds to amine groups was mixed in sodium bicarbonate buffer at 8.5 and briefly flowed through each of the three channels using a syringe. The channels were then flushed with fresh buffer to remove any unbound label and cleared with air before the probe tip was removed from the surface. The patterned surface was imaged with a Nikon epifluorescence microscope using the GFP filters. Figure 76 shows the patterned surface.

![Figure 76: Glass slide selectively labeled with fluorescein isothiocyanate to demonstrate microfluidic selective surface patterning](image)
Since the three channels are independent, we could have used different solutions in each channel to produce closely spaced regions of different functionality. When we tried this, some leakage between channels occurred, as a result of a poor seal. This could probably be improved by redesigning the mold to produce a narrow raised gasket surrounding the channels as in the microfluidic probe method 1.

Here we have demonstrated that we can produce arbitrarily patterned regions of chemical functionality down to a few tens of microns on a substrate. By incorporating this into the automatic wafer prober, we can reproduce the pattern thousands of times on a substrate giving us great flexibility.
Chapter 7: Conclusions

This thesis builds upon a foundation laid by the very successful work of a previous graduate student. In seeking to carry the ball forward, we have tried to improve upon each piece of the biosensing system. We started with the electronic devices, improving their performance, reliability and repeatability by redesigning the layout and fabrication methods. We continued by improving the characterization methods and developing new methods where necessary. Finally, we revisited the question of addressing the sensors with a fluid, developing techniques and tools to maximize throughput, yield and sensitivity.

In our efforts to improve upon the electronic devices, we developed several successful processes for fabricating nanowire FETs. We developed a simple, 6-inch wafer-scale, CMOS compatible process for fabricating back-gated nanowire devices. This approach preserved the critical TMAH wet-etching step of the previous generation that allowed us to retain excellent carrier mobility and noise characteristics; however by revamping the fabrication process, we were able to push the device size down by a factor of 3 to channel widths of 30 nm and achieve much higher yield and device-to-device repeatability. Further, we were able to fabricate suspended-channel devices, which may increase sensor sensitivity in both the diffusive and convective limits. Perhaps more importantly were able to produce high-quality complementary devices. Since nanowire sensors are most sensitive when sensing in a depletion modality, it is important to pick the correct type of device for sign of the charge on the target analyte.
thus complementary sensors are absolutely necessary for achieving maximum sensitivity.

We also developed fabrication processes for locally-gated nanowire FETs. Using a self-aligned process that once again took advantage of the favorable results of TMAH etching we fabricated nanowire devices with individually-addressable polysilicon top-gates. Individual gating is critical for operating complementary devices on the same chip, something that will become increasingly important as sensor devices are integrated into massively parallel systems for detecting multiple analytes. We were able to produce suspended-channel top-gate nanowire FETs as well, which replaces the sensor surface area lost by the addition of the top-gate. To our knowledge these are the first demonstration of suspended top-gate nanowire FETs. The top-gated nanowire FETs, both suspended and unsuspended had excellent transistor characteristics, comparable to commercial FETs.

In our efforts to better understand the performance of nanowire FET sensors, we used two new methods for characterizing their quality. The first of these methods used the measured sensitivity of a device to pH and to a reference electrode to separate the electronic and surface-chemistry components of the deviation from an ideal sensor. In our case, where we used thin silicon dioxide as the insulator, we observed that most of the degradation was electronic suggesting mobile ions from the buffer as a source of interface charge. This leads to the conclusion that a different gate insulator, such as silicon nitride, aluminum oxide, or tantalum pentoxide should be employed in the future.
The second method developed was a low-frequency noise spectroscopy technique that allowed us to quantitatively compare nanowire devices of different geometry and processing conditions. This technique allowed us to extract a single parameter, the Hooge parameter, which represented the level of 1/f noise intrinsic to the device. We saw that our devices compare very favorably to other nanowire FETs, as well as to commercial MOSFETs of similar dimensions. We also used this technique to compare the effects of various processing conditions. Comparing nominally identical devices prepared by our TMAH etching technique and by two commonly used plasma-etching techniques showed that the TMAH-etching yielded a five-fold improvement in noise performance over the plasma techniques. The TMAH etched devices also showed significantly higher mobility and lower subthreshold slope, suggesting that indeed they had a lower density of surface states created by process-induced damage.

Finally, we developed a new tool to address the problem of transporting the analyte-bearing fluid to and from the surface of the sensor. This new technique which we call the microfluidic probe, is a fluidic analogue to the common electronic probes used in most electronics labs. It allows one to quickly and non-destructively contact a wafer to form a sealed continuous-flow fluidic system without disturbing the capacity for electronic wafer probing. We have successfully integrated this system into both a manual probe-station as well as an automated probe-station with a probe-card capable of simultaneously measuring all the devices on a chip. This device has the capacity to bring fluidics into the electronics lab without the need for major changes in equipment or device layout. We also used a variation of the fluidic probe to demonstrate arbitrary micro-scale substrate patterning by chemical functionalization.
There are several issues that should be addressed in future to advance this work and maximize its utility. The easiest and most effective change would be to adopt a new gate material as opposed to the silicon dioxide that we used. Several materials are available which should drastically improve stability and leakage performance of devices in solution based on their resistance to metal-ion diffusion, such as silicon nitride. However, great care must be taken to avoid high interface state densities that would rob sensors of their sensitivity. Our initial plans employed an oxide/nitride layer but we fell away from this in later fabrication runs, a mistake that proved costly.

The suspended top-gate nanowires open up several new avenues of research which should be pursued. The most obvious is integration. Locally gated nanowires can be integrated to form differential amplifiers that would improve sensitivity. Once functionalization techniques become more reliable and selective, it will be important to tune each wire in an array individually to maximize sensitivity as the different functional groups on each wire will produce different threshold or flat-band voltages for each device. This can only be achieved using locally gated devices. The top-gated suspended wires should also allow the employment of AC analyte concentration techniques such as electro-diffusion flow [103]. Further, there are many interesting electromechanical and low-dimensional physics experiments that are enabled by these devices and it is our hope that some of these areas will be investigated by the next generation of graduate students. To them, we say “good luck and have fun”.
Appendix I: HgFET characterization of SOI Wafers

In order to determine whether process-induced damage is degrading the quality of our devices, it is important to have a method to determine the quality of the starting material. In our case, we would like to know the carrier mobility and threshold voltage of the SOI wafers from which we fabricate our nanowire FETs prior to any processing. A pseudo-MOS technique using mercury contacts was reported by Hovel [92] and used to characterize SOI films non-destructively. The HgFET is shown schematically in Figure 77.

![HgFET schematic view](image)

**Figure 77: Schematic view of HgFET (a) Cross-section (b) Plan-view**

Building upon this approach, we developed a probing device for performing HgFET measurements. We molded a device out of PDMS with two concentric channels that could be placed on the surface of a wafer and filled with mercury to form
concentric source and drain contacts. The two mercury electrodes are contacted through the syringe needles used to fill the tubing that leads to the channels. This device is shown in Figure 78.

![Image](image.png)

**Figure 78: PDMS probe for HgFET measurement**

For mercury contacts on silicon with no surface states, such as is the case when the native oxide has recently been removed with hydrofluoric acid, the barrier heights for n-type and p-type silicon are given by:

\[
\phi_{Bn} = \phi_M - \chi \approx 0.45eV
\]

\[
\phi_{Bp} = E_G - \phi_{Bn} \approx 0.67eV
\]

Thus, electron injection is relatively easy but hole injection is more difficult. Once a native oxide regrows, the density of surface states increases. As a result, \( \phi_{Bn} \) increases to about 0.7 and \( \phi_{Bp} \) decreases to about 0.42 \( [104, 105] \) thus hole injection becomes easier and electron injection becomes more difficult.

We measured the \( I_d vs. V_g \) characteristics of the HgFETs formed on the SOITEC ultra-thin SOI wafers used to fabricate the nanowire FETs. The accumulation-mode (p-
type) operation was measured prior to any surface treatment because of the barrier height consideration. The $I_d$ vs. $V_g$ characteristic is shown in Figure 79 (a). The mobility is calculated as in chapter 4 from the equation

$$
\mu = f_g \frac{g_m}{C_{ox} V_{ds}}
$$

where $f_g$ is a geometric factor that replaces $W/L$ in the conventional MOSFET equation and is given by the expression $f_g = 2\pi/\ln(R_2/R_1)$ [106]. The transconductance $g_m$ is calculated by differentiating the measured $I_d$ vs. $V_g$ characteristic with respect to $V_g$.

![Figure 79](image-url)

**Figure 79:** Accumulation mode characteristics of SOI HgFET (a) Drain current versus backgate voltage (b) Field-effect mobility of holes

The inversion mode characteristics were measured after treating the surface with dilute hydrofluoric acid to remove the surface oxide. This yields a surface that is nearly completely terminated with silicon-hydrogen bonds and thus has a very low density of interface traps and consequently a lowered barrier for electron injection. The characteristics gradually returned to the untreated state over a period of 2-3 days.
as the native oxide was re-grown. The inversion mode $I_d$ vs. $V_g$ characteristics and calculated electron field-effect mobility are shown in Figure 80.

![Figure 80: Inversion mode characteristics of SOI HgFET](image)

(a) Drain current versus backgate voltage (b) Field-effect mobility of electrons

From the $I_d$ vs. $V_g$ characteristics, we can also extract the flat-band and threshold voltages, as well as the subthreshold slope. With these values, we can calculate the density of interface states at the buried interface as well as the doping of the film according to the following equations [92, 106].

\[
V_t = V_{fb} + \frac{2\phi_s}{C_{ox}} (C_{ox} + C_{it} + C_d) - \frac{q(N_D - N_A)t_{Si}}{2C_{ox}}
\]

\[
S = 2.3 \frac{kT}{q} \frac{(C_{ox} + C_{it} + C_d)}{C_{ox}}
\]

where $\phi_s$ is the surface potential, $C_{it}$ is the interface state capacitance given by $C_{it} = qD_{it}$, and $C_d$ is the depletion layer capacitance given by the permittivity of silicon divided by the depletion width, which for thin films can be approximated as the full thickness of the silicon film, $t_{Si}$. The extracted values of $V_t$, $V_{fb}$, $C_{it}$, and doping are given...
in Table 8. The calculated doping level agrees with the value specified by the manufacturer and the interface state density is acceptably low.

### Table 8: SOI parameters extracted from HgFET measurements

<table>
<thead>
<tr>
<th></th>
<th>$V_{fb}$ or $V_t$</th>
<th>$D_{it}$</th>
<th>$N_A$</th>
<th>$\mu_{FE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accumulation</td>
<td>-13.8 V</td>
<td>&lt; $1 \times 10^{11}$ cm$^{-2}$</td>
<td>$5 \times 10^{14}$</td>
<td>775 cm$^2$/V·s</td>
</tr>
<tr>
<td>Inversion</td>
<td>-7.9 V</td>
<td>$3 \times 10^{11}$ cm$^{-2}$</td>
<td>x</td>
<td>250 cm$^2$/V·s</td>
</tr>
</tbody>
</table>
Appendix II: Low temperature measurements

Cryogenic characterization of the top-gated nanowire FET devices was intended to be a considerable portion of this thesis. The top-gated nanowire FET devices are a promising system for investigating low-dimensional physics in silicon for a few reasons. They have long narrow channels, high room-temperature mobility, and a degenerately doped gate, something that the back-gate devices lack. This is important as it allows the devices to operate at temperatures below 4.2 K. In contrast, the back-gated devices we employed previously could not be operated below 30 K as the resistance of the back-gate became prohibitively high due to carrier freeze-out or to Schottky barrier contacts. These devices should be ideally suited for observing the population of individual 1-dimensional sub-bands that has been predicted to occur in nanowire type devices as a result of physical confinement. An increase in mobility has also been predicted to occur as carriers in the first sub-band are concentrated in the center of the silicon channel, minimizing interface scattering.

Devices were diced and packaged in 16-pin dual-inline packages as described in chapter Chapter 3:. The cryogenic measurements were carried out in 2 different cryogenic systems. The first was a Janis variable temperature liquid helium (LHe) cryostat capable of reaching temperatures below 1.4K. This system has a superconducting magnet capable of producing a magnetic field of 9T. The second system was a Janis LHe continuous flow cryogenic probe-station which could only reach 5K. The electrical measurements were taken with an Agilent 4156B semiconductor parameter analyzer except for the Hall-effect measurements, which were taken using a
Yokogawa 7651 DC voltage source and a Keithley 6514 system electrometer to measure current.

A2.1 Hall effect and field-effect mobility versus temperature

Hall effect and field-effect mobility versus temperature measurements were taken on back-gate-only nanowire FETs fabricated by Eric Stern. A representative device is shown in Figure 81.

![Nanowire FET with Hall-bar geometry](image)

**Figure 81: Nanowire FET with Hall-bar geometry**

The field-effect mobility was determined by measuring $I_d$ vs $V_{gs}$ and calculating the peak transconductance from the numerical derivative as discussed in chapter Chapter 4:. Hall-effect mobility was calculated using the equation

$$\mu_H = \frac{|R_H|}{|\rho|}$$

where $|R_H|$ is the mean value of the hall coefficient, given by
and $|\rho|$ is the mean value of the resistivity field. The measurements are performed in the LHe cryostat. A known current is sourced through the axial terminals of the Hall bar in zero magnetic field and the voltage across each pair of adjacent terminals is measured to determine each value of $\rho$. The Hall voltage was measured across each pair of transverse electrodes in the presence of a strong magnetic field. The field-effect mobility and Hall coefficients were measured from room temperature down to the point where the back-gate conduction froze-out, around 30K. We plot the Hall-effect mobility versus temperature in figure. To our knowledge this is the first measurement of Hall mobility in a silicon nanowire FET.

Figure 82: Hall-effect and field-effect hole mobility versus temperature
A2.2 One-Dimensional Subbands

We were unable to perform extensive low-temperature characterization of the devices due to numerous equipment malfunctions with our cryostats. We were; however, able to observe the presence of sub-bands in the devices using a continuous-flow cryogenic probe station in the brief time we were allowed to use it. This tantalizing yet unsatisfying result suggests that indeed the top-gated nanowire FETs should be a good model system.

We began the cryogenic characterization with a survey of carrier mobility as a function of temperature. This can tell us a number of things about the devices. Particularly, we can extract information about the nature of carrier scattering in the devices. Figure 83 shows a typical mobility versus temperature curve.

![Graph](image)

**Figure 83:** Field-effect mobility versus temperature  (a) Accumulation mode device  (b) Inversion mode device

The 1-dimensional density of states was calculated for a device of similar dimensions to ours by Colinge [107] and used to predict that the energy spacing between the first few sub-bands would be around 150 µeV. At 5K, the minimum
temperature we could reach in the cryo-probe station, the thermal energy is only a few times larger than the predicted sub-band spacing and thus we should be able to observe the population of the individual sub-bands. We measured $I_d$ vs $V_{gs}$ at very low source drain current and indeed we observed the presence of oscillations in the transconductance characteristic of 1 dimensional subbands, as shown in Figure 84.

**Figure 84:** Oscillations in transconductance due to 1-dimensional sub-band filling


144


