Integration of Analog VLSI and Thin Films for Chemical Sensing Arrays

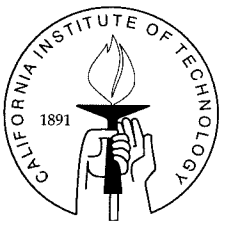
Thesis by

Jeffrey Allen Dickson

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

1999

(Submitted May 21, 1999)

Acknowledgements

I would like to thank my advisor, Rod Goodman, for his support. He gave me an excellent environment in which to investigate many things, beyond that described in this document.

There are many people who are responsible for prolonging my tenure at Caltech. Fortunately, they also enriched it, and for that I owe many thanks. Glen George for many hours of stimulating discussion and his technical assistance. My officemate David Babcock, for his friendship and assistance: when his work too him off-campus his presence was certainly missed. Bhusan Gupta, for his technical expertise and willingness to assist.

I would like to thank Phil Stapleton of Stapleton Technologies, for his assistance in obtaining the plating products critical to this work. Tom Tsao, of the Micromachining Laboratory, assisted in the initial plating attempt and was very helpful answering questions. Alina Moussessian was a helpful courier, taking the die to the Jet Propulsion Laboratory for bonding by Jerry Mulder. Dr. Carver Mead generously allowed me to use his fume hood and equipment to perform the plating.

Abstract

Chemical sensors are typically discrete elements, each sensitive to a particular odor. Current chemical sensing systems are unable to detect a wide range of chemicals. Moreover, the lack of an integrated sensor technology prevents the fabrication large arrays of chemical sensors.

A sensing technology is presented that integrates chemical sensors based on conducting carbon black particles and polymers with standard integrated circuit technology. An integrated circuit for array sensing is presented. The necessary post processing steps, consisting of and electroless gold deposition and subsequent polymer deposition, are discussed. The system is demonstrated to function as an array of individually addressable chemical sensors.

In addition, a novel nonvolatile analog storage element based on floating gates is presented. The application of this circuit and floating gate adaptation to ratiometric sensing is demonstrated.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER

ONE

The Windows of the Soul

Aristotle enumerated five senses – sight, hearing, touch, taste, and smell, which he described as the windows of the soul. Due to the importance of these senses all of them have artificial counterparts: CCDs and CMOS vision sensors for sight, microphones for hearing, pressure sensors for touch, and chemical analysis systems for taste and smell. Although no system comes close to mimicking the brain's sensory processing power, many of these sensors approach or exceed the size efficiency and performance of their biological counterparts. Attempts at mimicking the human olfactary system have fallen short, as large and complicated systems dominate the chemical sensing field. This thesis presents work towards changing that.

Traditional chemical sensing devices are large, expensive, systems such as ion spectroscopy and chromatography. Their size, complexity, and most of all cost have limited the applications of chemical sensing. Small chemical sensors could monitor the integrity of seals or pipes carrying dangerous chemicals, detect overheating transformers, check the quality of fish or beer, detect explosives, even perform medical diagnoses.

Another aspect of traditional chemical sensors is that they are finely tuned, responsive to a specific analyte or ligand. While this approach is acceptable for many applications, specific sensors are available for a small subset of the total sensory space. In addition, this approach has scaling difficulties when a large number of analytes need to be monitored. Situations for multiple analytes are not hard to find: even the simple strawberry has over 35 chemicals that make up its odor, which can vary with time in the same berry.

The most popular individual sensing element are those based on thin film metal oxide technologies. These sensors detect vapors by their interaction with surface states on a semiconductor. The major drawback of these sensors is that they have to be operated at high temperatures, as high as 400°C, in order to have the required sensitivity and response times, and to make the process reversible. This temperature requirement makes integration difficult due to the need for heating elements for the sensors, while being thermally isolated from any associated circuitry. Moreover, the heating can have a large power requirement. In addition, because of the limited number of odors that this method can sense, due to the sensing method, it is not suited for use as an element in a array when a large number of sensors are desired.

A simple but very effective method of chemical sensing uses acoustic sensors as the sensing element. A transducer consists of a bulk acoustic wave

oscillator or surface acoustic wave oscillator, coated with a thin film of a chemically selective material. The odor molecule adsorbs onto the film, or absorbs into its bulk, changing the mass of the oscillator. The change in mass causes a change in frequency of the oscillator. These sensors have been demonstrated to be very sensitive. One drawback in the past has been the requirement that the oscillator be made of quartz, limiting its suitability for integration on a chip. However, recently Sandia Labs reported an array of six sensors fabricated in gallium arsenide, also using GaAs as the oscillator material.

One interesting approach to chemical sensing is the optical sensing developed by Walt et al. In this approach, a fluorescent dye combined with an organic polymer constitutes the chemical sensor. The dye has the property that it is sensitive to the polarity of its surroundings. When the specific odorant adsorbs into the polymer the spectral properties of the dye change, resulting in a change in color. In addition, the polymer undergoes a physical swelling which also affects its response. By using different organic polymers a broad range of chemical sensors is possible. The sensors are fabricated on the end of a fiber optic cable and observed with a CCD. The sensor responses are fast compared to most other technologies. Most importantly, the individual sensors are small, and fabrication technology has been developed to assemble arrays. Drawbacks are that the lifetimes of the sensors is short, and the need for optical processing to read out the sensor responses.

Another array approach based on optical sensing is that of Nelkirk et al., which they have dubbed "The electronic tongue." Like the approach of Walt, color changes in the sensing media indicate the detection of chemicals. Small beads of the sensing material are placed in etched wells of silicon and observed using a conventional camera [6].

For array implementations with sensors that have broad responses, the most popular material has been electrically conducting polymers. These are fabricated by the chemical or electrochemical polymerization of organic monomers. While the initial polymers, such as pyrrole, are not conducting, the crosslinking forms structures where the π-electrons are delocalized, and the material becomes semiconducting. Polar molecules that diffuse into the polymer can alter the distribution of the electron energies, which in turn causes a change in the resistance of the polymer. This change in resistance can then be detected.

There are a number of attractive attributes of conductive polymers, especially for an array implementation. The size can be small, without any need for complicated oscillating masses. They operate at room temperature, so no heater is required. They operate on polar gasses and vapors, which represent a very important class of gases to detect. However, they have their drawbacks. Most serious is the drift of the sensor baseline with time, partly due to their sensitivity to humidity. This drift requires attention in the signal processing or classification stage; however, the rate of drift is much slower than the response to a vapor, making it relatively easy to compensate.

A variation on the conducting polymer approach is that demonstrated by Lewis et al. [7–10, 14, 15], employing carbon black and non-conducting polymers. Carbon black particles mixed with the polymer, as opposed to polymerization, create a conducting polymer. The number and density of the carbon black particles determines the electrical resistance of the polymer. When odor molecules absorb into the polymer the polymer swells, and the distance between the carbon black molecules increases. This in turn increases the electrical resistance of the polymer. Different polymers swell for different odor molecules, enabling the fabrication of a large number of different sensor types;

moreover, the selection of polymers is not limited to those that create semiconducting polymers like pyrrole. However, carbon black based polymers share many of the same drawbacks of the conducting polymers, namely baseline drift. Despite the shortcomings, these sensors make a very attractive candidate for integration with active circuitry, since they are easy to fabricate and present the detected signal in a form readily processed by circuitry.

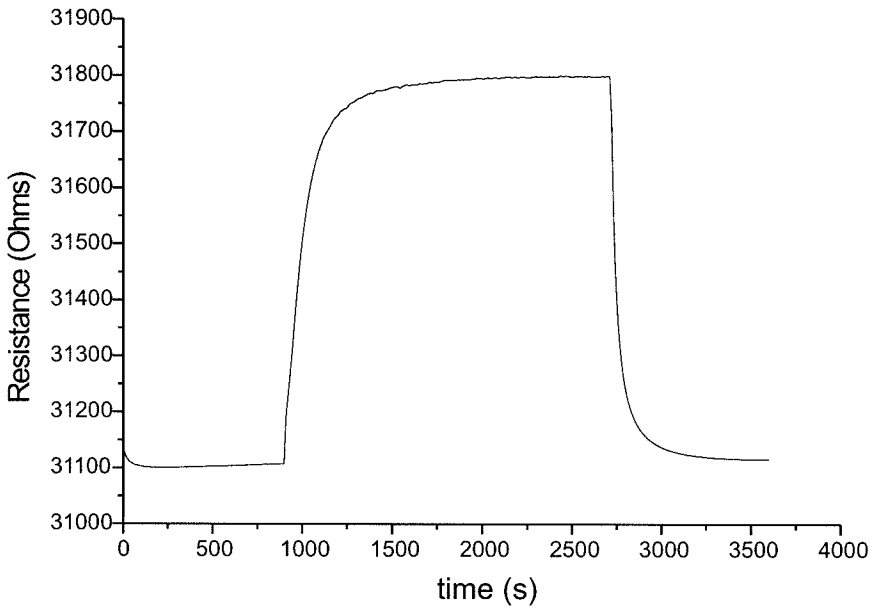


Figure 1.1: Carbon Black sensor Detecting Octanol

1.1 The Biological Nose

Nature evolved a very different sensing system for the chemical sensing needs of organisms. The mammalian nose, for example, has many broadly tuned sensors, not individual sensors for a particular odorant. The biological system is one of the most effective chemical sensing systems. The canine system is partic-

ularly sensitive, and for that reason dogs are the most effective approach for the detection of narcotics, explosives and survivor location in disaster response.

Carver Mead coined and pioneered "Neuromorphic Systems," the investigation of biologically inspired silicon systems. Most of the work concentrates on applications to sensor processing, such as the plexiform modeling in the silicon retina, the basalar membrane model in the silicon cochlea, or the Reichart model for motion sensing. In this work there is no attempt at implementation of biologically inspired sensor processing; however, this does not mean a study of evolution's efforts into chemical sensing should be ignored. In this section I will briefly describe the mammalian olfactory system.

The motivation for looking at the biological system is clear - the human sense of smell is extremely capable. The human olfactory system is able to distinguish certain odors in concentrations in the parts-per-trillion range, and the capabilities of other mammals are even greater. The odors we are subjected to are composed of a large number of compounds which contribute to their odor, such as popcorn with one hundred such compounds [11], yet we are able to identify and classify a large number of distinct smells. In addition, we can distinguish subtle differences between odors, such as the smell of strawberries of different stages of ripeness.

The biological implementation of smell is the least understood of our senses. Compared to senses such as Vision and Hearing, where the physical sensing mechanism are well understood and much of the biological processing have been revealed, there is not a well defined model for the olfactory system. For example, we know very little about how a particular chemical structure results in the perception of a particular odor.

The biological sensing system starts with a binding of odor molecules with

receptor proteins situated on the olfactory sensory neurons. These neurons are located in the mucousal layer of the nasal epithelium, or lining, of the nose. In mammals the odor particles must dissolve in the mucous to be sensed. The binding of odorant and the receptor proteins results in the generation of action potentials in the axons of the olfactory sensory neurons. These action potentials, generated by approximately 100 millions olfactory cells, propagate along the axons of the olfactory system to the 5000 glomeruli nodes of the olfactory bulb where the signals are processed. Little is known about how the brain represents or processes odor signals.

The most interesting aspect of the olfactory system for purposes of this research is the mechanism for detecting odorants. Fortunately, there have been some recent insights gained from gene sequencing. Buck has reported that there are approximately one thousand receptor proteins expressed in the human genome [1]. Since a typical human can identify ten times that amount of different odors, this finding indicates that the biological system does not employ specific receptors for specific odorants. Buck produces evidence that there is broad receptivity among the odor receptor proteins [2,3].

Although the human system is remarkably adept, it does have a number of shortcomings. The nose is unable to detect a large number of the chemicals, even deadly gases such as carbon monoxide. The odor can be masked by another, a shortcoming that the entire air freshener industry exploits to great profit. In addition, the brain quickly adapts out responses to a constant odor after a few minutes of exposure. While this is advantageous from the standpoints of bandwidth and energy, we do not generally have a method of distinguishing between adaptation of the odor and cessation of its presence.

1.2 Applications of Chemical Sensing

Funding for this work was in part provided by DARPA for the investigation of land mine detection. There are an estimated 100 million landmines buried on the planet, one of which explodes every 20 minutes. Many of these are legacies of past wars, and now detonate not in fields of war, but in school playgrounds. Most of the injuries, many times fatal, are to civilians. Detection of these land mines is very difficult. Development of a portable and cheap sensor to detect the explosives in land mines, typically RDX and TNT, would be a great tool in effective mine elimination.

More applications and and other technologies will be added by handin.

1.3 Thesis Organization

In chapter 2 I will discuss the development of a nonvolatile memory for analog VLSI. While not directly related to chemical sensing, I will show in later chapters its application to the integration of chemical sensors.

Chapter 3 discusses adaptive analog circuitry for integrated chemical sensing. I present a discussion of the basics of ratio metric sensing, the Wheatstone bridge, and discuss its active implementation in an integrated technology. A novel circuit based on gate oxide tunneling currents is presented that biases a resistive sensor and adapts out baseline drift.

In Chapter 4, I discuss the post processing steps for a standard VLSI fabrication process that allows the integration of chemical sensors with integrated circuitry. The problem of making contact to the exposed aluminum of a standard integrated circuit is solved using an electroless gold plating process.

Chapter 5 presents an integrated array of chemical sensors. While simple in scope, the circuits allow the creation of an array of individually addressed sensors. The results of applying the post processing results of chapter 4 are presented.

In Chapter 6 I summarize the thesis and discuss future directions.

CHAPTER

TWO

Persistence of Memory

Analog VLSI is a compelling paradigm for the implementation of sophisticated computation from both an area and energy perspective. One of the major impediments to the practical application of analog computation to large scale systems, such as neural networks, is the lack of a non-volatile storage element. This is unfortunate since analog VLSI implementations of these systems have many otherwise attractive elements. Many of the computations involved in adaptive learning systems have elegant implementations in analog circuitry. Moreover, the underlying physics of the devices can be exploited for incredible density and complexity. In this chapter I describe a non-volatile analog memory cell for use in analog learning systems. While not directly related to the problem of chemical sensing, there are applications to this thesis, as will be shown in the next chapter. In addition, the ultimate goal of on-chip classification can benefit from the work in this chapter.

2.1 Analog Memories

There are a number of methods to circumvent the lack of analog storage for integrated circuit implementations of learning systems. This section presents a brief overview of these approaches.

The natural storage element in VLSI is the capacitor, which can be charged or discharged to the desired voltage. The capacitor is a natural element of VLSI technologies - all MOS gates are effectively a capacitor. A MOS transistor, with its insulated gates, makes an ideal readout device as it does not remove or add charge to the capacitor. However, it is impossible to make an ideal switch in MOS technology. Any switch used to add or remove charge to the capacitor, in order to change the value of the memory, forms a leakage path that will quickly remove the charge. This prevents the storage of parameters.

Similar to the refreshing of digital bits in a standard digital DRAM, researchers have applied a refreshing scheme to analog implementations. By reading the value of the capacitor and converting to digital form, the leakage can be compensated by rounding up to the next higher quantization. For example, in a system with the levels (0,1,2,3,4) volts, the refreshing system would refresh a voltage of 3.9 volts to 4 volts. Typically, the refreshing circuitry is implemented off chip, resulting in complicated systems. However, this allows modification of the values in real time by the external control system. This could be beneficial in a learning system where the learning is performed externally, allowing investigation of a variety of algorithms. Moreover, the long term storage of the values is not a problem, as they can be stored in the external memory.

One drawback with a refreshing system is that as the number of the ele-

ments increases, the demands on the refreshing system increase. As the number of nodes increases, the rate of the refreshing per cell drops, limiting the precision of the refresh system.

There have been systems that integrate the refreshing on-chip. Caughenburgs developed an interesting system with on-chip algorithmic analog to digital converters that eliminates the off-chip complexity [4]. Multiple units of the onboard units can increase the frequency of updates by dividing the workload. Hochet reported another refreshing scheme that has the refreshing circuitry local to each cell, eliminating the need for multiplexing [12]. The circuit simply requires a square wave and triangle input that is global to all cells. While both of these systems do not have the requirement of external circuitry to store and refresh the analog values, this functionality is not without its cost. Neither have a capability for nonvolatile storage, so when the power is removed the values are lost. This could be a serious limitation in systems where the values are weights of a neural network. Without an ability to store the weights, they have to be relearned each time the power is restored. There has to be some external mechanism to store and apply the training signals, which reduces some of the advantages inherent in this approach.

We have developed a seven-transistor circuit that is compact; is easy to use; possesses a 14-bit accurate, nonvolatile analog memory; and has a rail-to-rail buffered voltage output. We believe that this memory cell is well suited for use in silicon neural networks.

2.2 The PWM Memory Cell

The PWM memory cell is shown in Figure 2.1. The memory is stored as charge on the floating-gate. Because the floating gate is completely insulated by S_iO_2 , this stored charge is retained in a nonvolatile state. The feedback capacitor, C_{int}, integrates this charge to set the memory-cell output voltage V_{out}. We add electrons to the floating gate using hot-electron injection, and remove them using Fowler-Nordheim tunneling. M1 and M2 are the injection and tunneling transistors; they are described in detail in Sections 2.3 and 2.4, respectively.

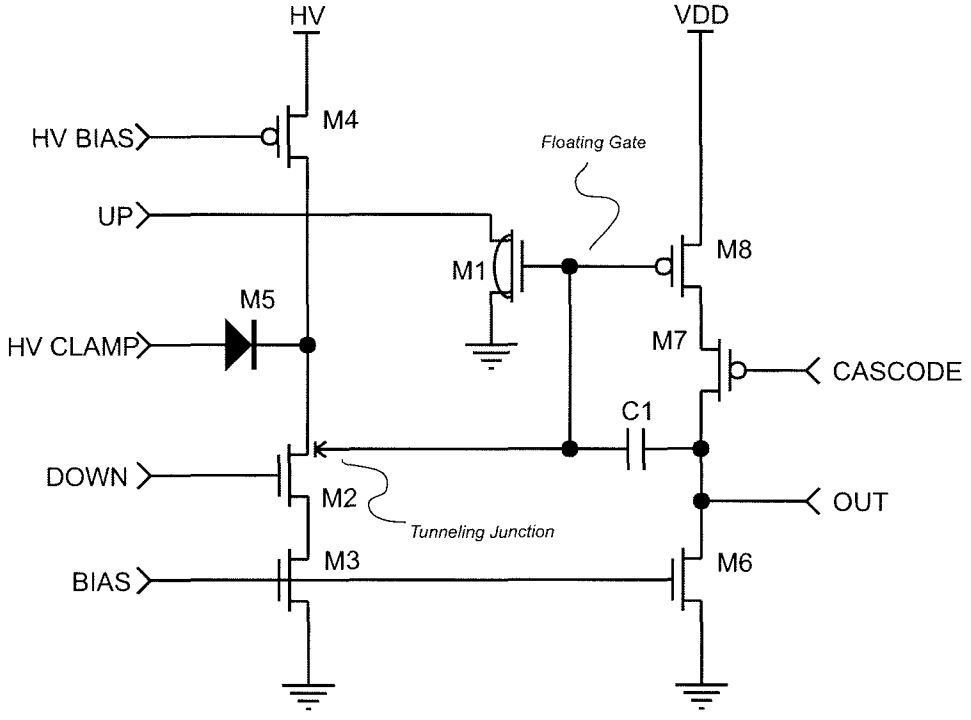


Figure 2.1: Schematic of the PWM memory cell.

We typically set V_{bias} such that transistor M3 operates in the subthreshold regime. Subthreshold operation maximizes the gain of the M3-M4 inverting amplifier, permits rail-to-rail output voltages, and conserves power. When we

apply a logic high to the up input, electrons inject onto the floating gate. Because the M3-M4-C_{int} feedback loop pins the floating-gate voltage, electron injection causes V_{out} to rise.

M2 and M6 form a high-voltage, pseudo-nMOS inverter. When we apply a logic low to the active low $\overline{down}$ input, M2 turns off, M6 pulls V_{tun} up to HV, electrons tunnel off the floating gate, and V_{out} falls. When we apply a logic high to the $\overline{down}$ input, M2 turns on, V_{tun} is pulled down until clamped by diode-connected transistor D1, and the tunneling process turns off. D1 prevents voltage-induced breakdown of M6's drain terminal. We use transistor M5 to limit M2's power dissipation, by restricting the source current.

We typically generate V_{hvbias} locally, using a high-voltage nFET with its source grounded, its gate driven by V_{bias}, and its drain tied to an HV-referenced diode-connected pFET. We can also generate V_{clamp} locally, by means of either a voltage regulator or a simple diode stack, referenced to HV.

Because we intend to fabricate silicon learning systems that employ a large number of memory elements, we designed our memory cell to be simple and compact. It therefore has neither power-supply rejection, nor temperature compensation. We intend instead to fabricate, on each chip, a circuit that globally adjusts the bias voltage, V_{bias}, to compensate changes in V_{dd} and in temperature. Although we do not believe that this approach can preserve 14-bit accuracy across an entire chip, we expect that the residual error will comprise principally a long-term drift, which we intend to compensate using the network learning algorithm itself.

2.3 The Injection Transistor

We add electrons to the floating gate using injection transistor M1, shown in detail in Figure 2.2 (A) and (B). M1 is a conventional nFET, to which we add a p-type (1E17cm^3) bulk channel implant. We employ a bipolar-transistor base layer, available in the 2.0 micron MOSIS (Orbit) BiCMOS process, as this channel implant.

The hot-electron injection process is shown in the energy-band diagram of Figure 2.2 (C). We accelerate channel electrons in the transistor's channel-to-drain E-field. These electrons inject from the channel-to-drain space-charge layer, over the 3.2V Si-S_iO_2 work-function barrier, into the oxide conduction band. They then are swept by the oxide E-field over to the floating gate. Successful injection requires that three conditions be satisfied:

1. the electrons must possess the 3.2eV required to surmount the Si-S_iO_2 barrier,

2. the electrons must scatter upward into the gate oxide, and

3. the oxide E-field must be oriented in the proper direction to transport the electrons over to the floating gate.

Using a conventional n-type MOSFET, we can easily satisfy the first two requirements. We merely operate the transistor in its subthreshold regime, with a drain-to-source voltage greater than about 3V. Because the subthreshold channel-conduction band is flat, the channel-to-drain transition is steep, and the E-field is high. Channel electrons are accelerated rapidly in this field; a fraction of them acquire the 3.2eV required for injection. A fraction of these

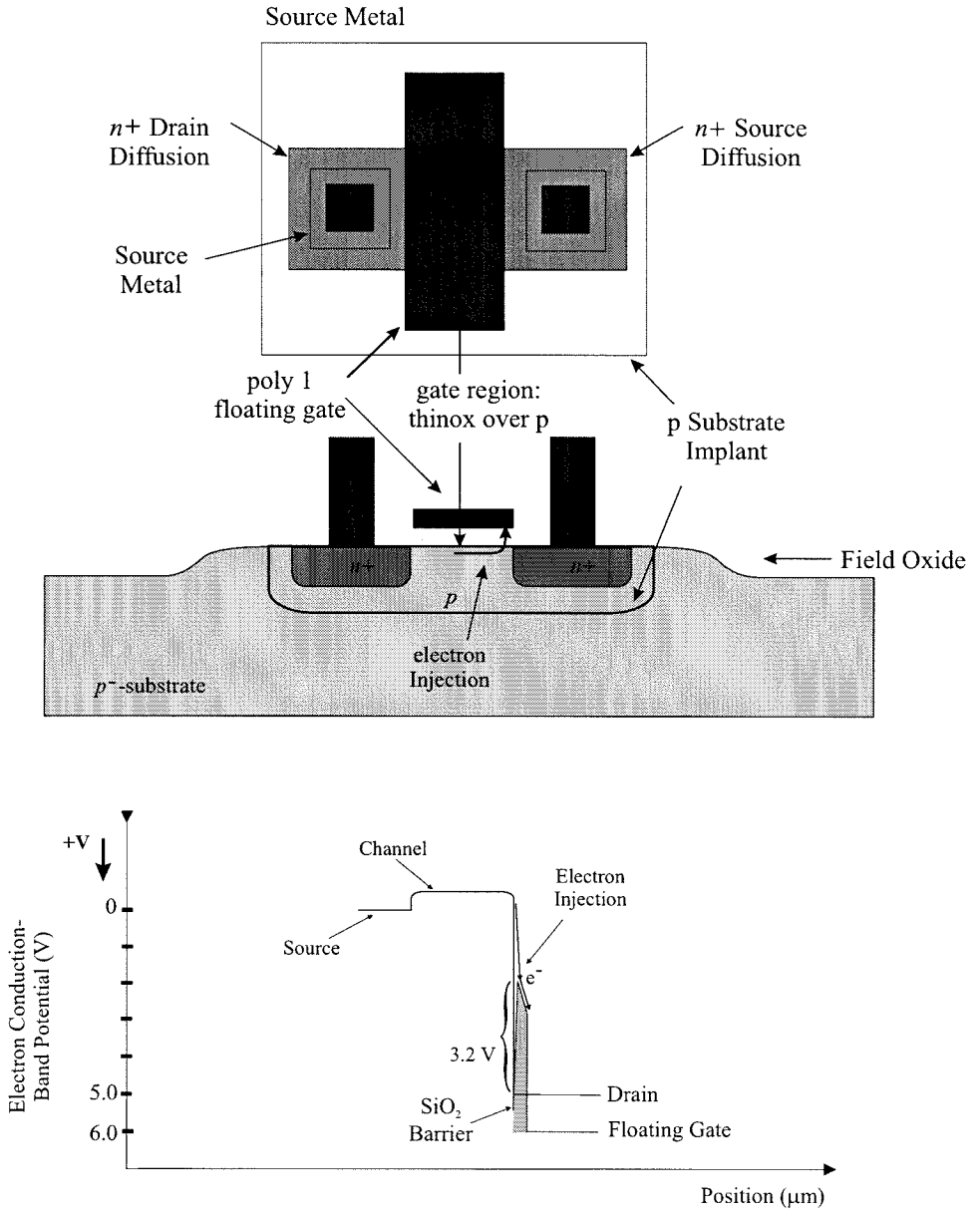


Figure 2.2: Injection in the pbase transistor.

3.2eV electrons naturally scatter, by means of collisions with the semiconductor lattice, upward into the gate oxide. Unfortunately, the third requirement prevents a hot-electron gate current. Because subthreshold operation typically requires $V_{gs} < 0.8$V, when V_{ds} = 3V, the drain-to-gate E-field opposes transport of the injected electrons to the floating gate. The electrons return instead to the drain.

If we instead operate a conventional transistor above threshold, with its drain at 3V and its gate at 4V, we satisfy requirement three. Unfortunately, with these terminal voltages there is a potential drop of about 3V along the channel. The channel-to-drain transition becomes shallow, the E-field drops, and channel electrons are no longer accelerated to 3.2eV. We satisfy requirement three, but fail to satisfy the first requirement. It is not until we apply 6V to the drain, and 5V to the gate, that the hot-electron gate current becomes large. Unfortunately, with these terminal voltages the channel current exceeds 1mA. We believe that the power consumption of an above-threshold injection transistor is too high for use in a silicon learning system.

We instead apply a bulk p-type channel implant to the injection transistor's channel region, thereby raising the threshold voltage from 0.8V to 6.1V. When V_{gs} = 6V, and V_{ds} = 3V, the channel current still is subthreshold, but now the oxide E-field sweeps injected electrons over to the floating gate, rather than returning them to the drain. With a channel current of 100nA, the injection transistor's gate current can exceed 10pA.

In Figure 2.3 we plot hot-electron injection efficiency (gate current divided by source current), versus both the gate-to-channel and the drain-to-channel potentials. We plot the data as efficiency because the gate current increases linearly with the source current over the entire subthreshold range. We refer-

ence all the terminal voltages to the channel because the injection probability varies with the channel potential. We can re-reference our results to the source terminal using the relationship between source and channel potential in a subthreshold MOSFET.

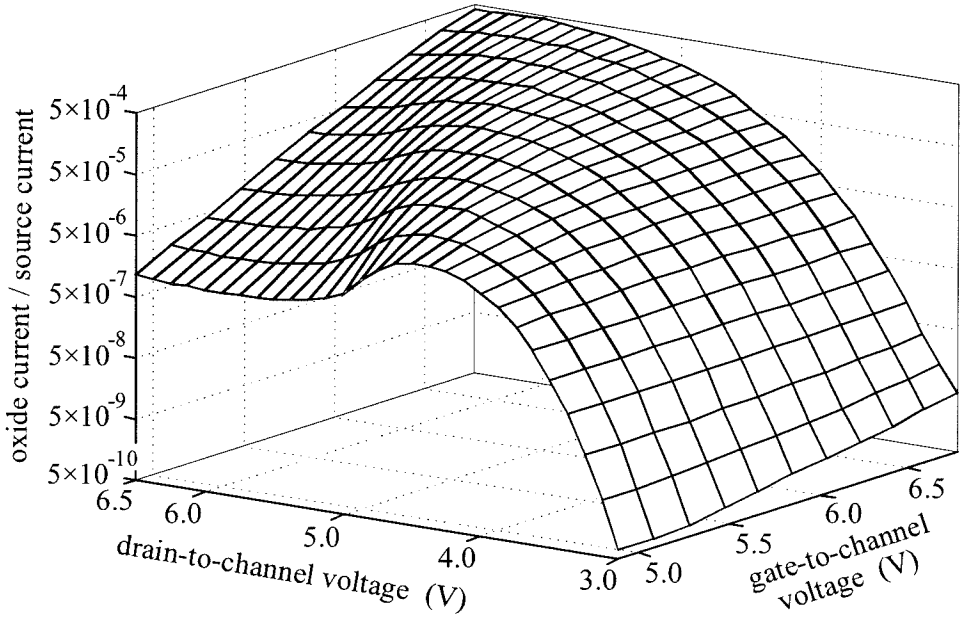


Figure 2.3: Injection current in a pbase transistor.

We typically operate the injection transistor with its floating gate voltage V_{fg} = 6V, and with its source at ground. We chose V_{fg} = 6V for three reasons. First, with V_{fg} = 6V, the source current still is subthreshold, but is quite large (about 100nA). Therefore the gate current will also be large. Second, because we inject with V_{ds} = 5V, when V_{fg} = 6V, the oxide E-field is in the proper direction to transport the injected electrons to the floating gate. Third, when V_{out} is near ground, pFET M4 can itself inject electrons onto the floating gate. The lower the gate voltage, the smaller this parasitic gate current. When V_{fg} = 6V and V_{out} = 0.5V, the pFET gate current is less than 1E-19A; when V_{out} = 1V, it is less

than 1E-21A.

To turn on the injection, we raise the drain to 5V; to turn it off, we lower the drain to ground. We chose V_{ds} = 5V because, as shown in Figure 2.3, at this value the oxide current is relatively insensitive to small variations in drain voltage. Because the M3-M4-C_{int} feedback loop pins the floating-gate voltage, M1's channel current is constant, and therefore its gate current also is constant. By controlling how long the drain is at 5V, we control how much charge is injected off the floating gate; the memory cell output voltage therefore increases linearly with up pulse width. If we apply $V_{dd} \approx$ 7V and $V_{bias} \approx$ 0.8V, the feedback loop will maintain $V_{fg} \approx$ 6V so long as M3 and M4 remain saturated (0.1V $< V_{out} <$ 6.9V).

2.4 The Tunneling Transistor

We remove electrons from the floating gate using tunneling-transistor M2, shown in detail in Figure 2.4(A) and (B). Electrons tunnel from the floating gate, through the S_iO_2 , to M2's drain. In order to minimize both the initial tunneling-rate mismatch between memory cells, and the decrease in tunneling rates due to oxide trapping, we tunnel through the thermally grown gate oxide. Because this oxide is 350-400Å thick, we must apply 25-30V across it to obtain reasonable tunneling currents. We therefore designed the tunneling transistor for 50V operation.

In a conventional nFET, the drain implant is heavily doped n+ (1E20 cm^3). The reverse-biased drain-to-bulk depletion region approximates a one-sided step junction, the E-field at this step boundary is high, and the pn-junction breakdown is approximately 20V. In the tunneling transistor, we replace the

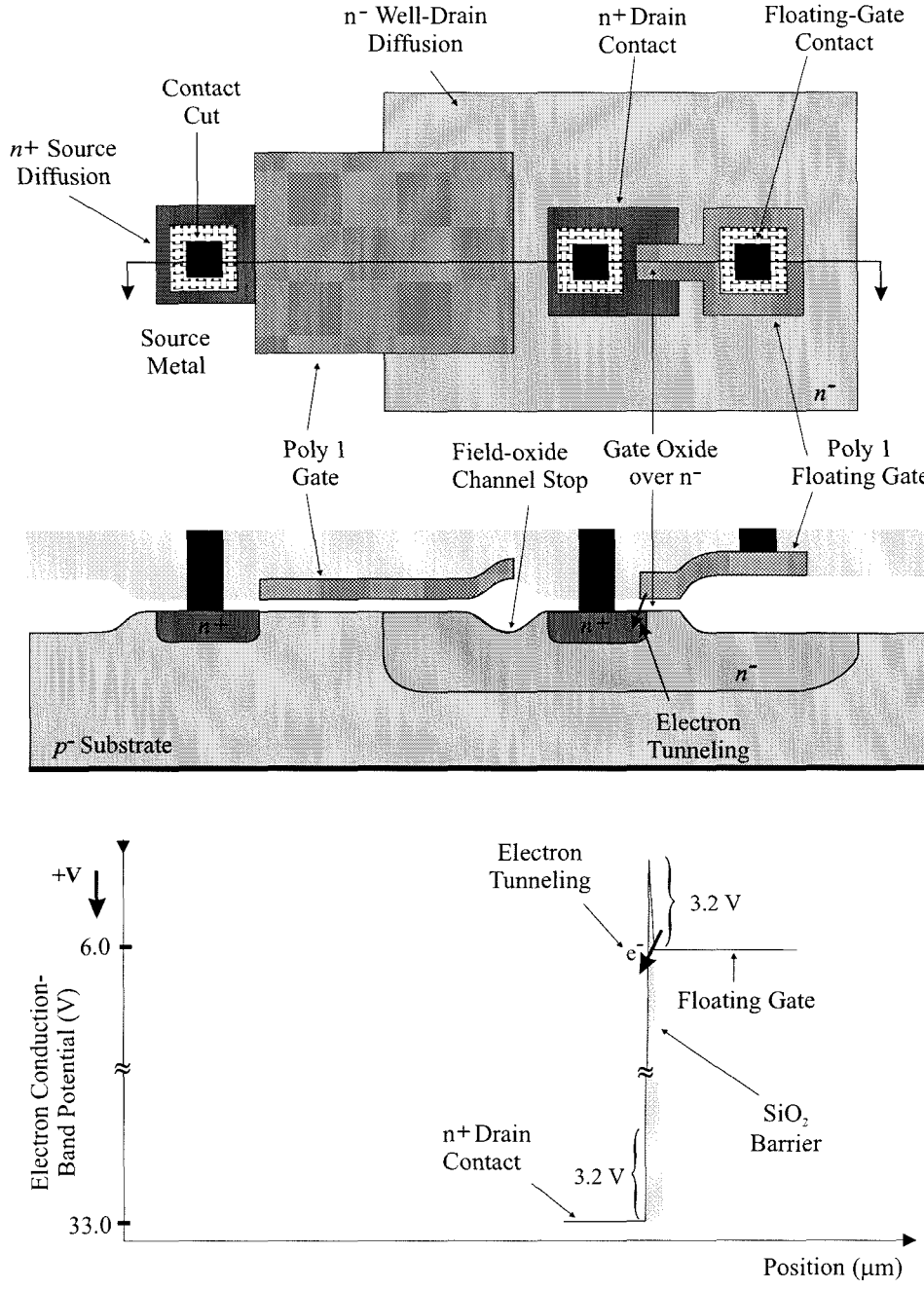


Figure 2.4: The tunneling junction.

n+ drain with a lightly doped (5E15 cm^3) n-well-drain. With this change, the tunneling transistor's drain-to-bulk depletion region becomes two-sided, the peak E-field is reduced, and the breakdown voltage is extended to beyond 50V.

The need for a 50V tunneling transistor is a consequence of the 350-400 Å gate oxides found in the 2.0 micron Orbit process, rather than a limitation in the memory cell itself. If the memory cell was fabricated in a modern process with 80 Å gate oxides, it would operate from a 12V supply. Unfortunately, in a more modern process the substrate impurity concentration is likely to be higher, and the n+-to-bulk breakdown voltage will therefore be lower. Thus, although a modern process will afford lower tunneling voltages, it is unlikely that the tunneling transistor will become a conventional nFET; it is instead likely to remain a high-voltage device.

As shown in Figure 2.4 (A) and (B), we form the gate-oxide tunneling junction by extending the floating gate over the n+ well-drain contact. The FN-tunneling process is illustrated in Figure 2.4(C). The high drain-to-floating-gate voltage reduces the effective oxide thickness, facilitating electron tunneling from the floating gate, through the S_iO_2 barrier, into the oxide conduction band. The electrons are then swept by the oxide E-field over to the tunneling transistor's drain.

Figure 2.5 shows the tunneling current versus the oxide voltage. The data are normalized to the gate-to-n+ edge length, in lineal microns. The reason is that the tunneling is enhanced where the self-aligned floating gate overlaps the heavily doped n+ well-drain contact, for two reasons. First, the gate cannot deplete the n+ drain contact, whereas it does deplete the n- well-drain. Thus, the oxide E-field is higher over the n+. Second, enhancement at the gate edge further augments the oxide field. Because the tunneling current is exponential

in the oxide voltage, gate-oxide tunneling is primarily an edge phenomenon.

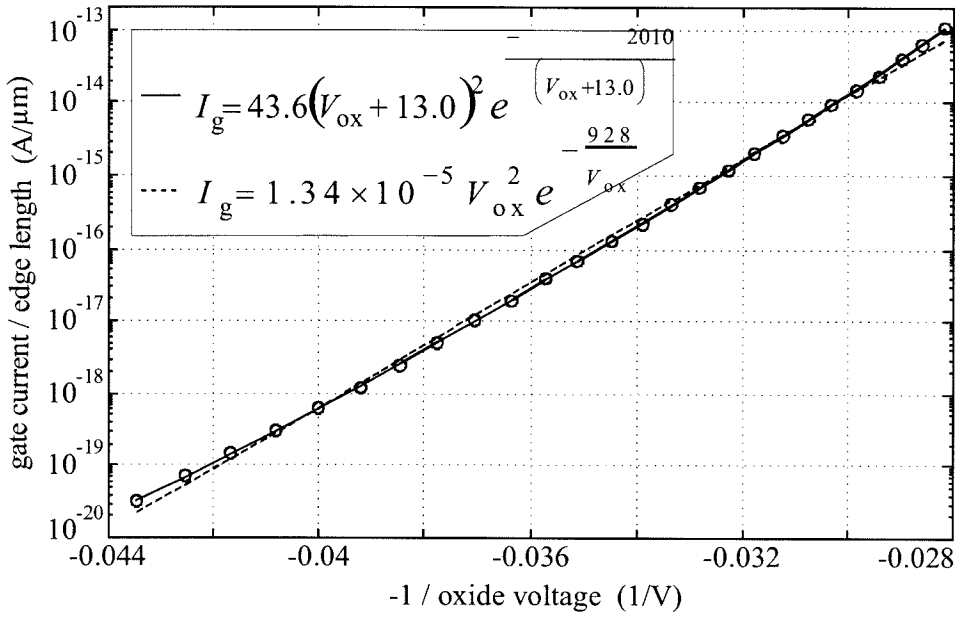


Figure 2.5: Plot of tunneling current for varying junction size.

As shown in Figure 2.4, where the floating gate exits the n- well-drain we route it in metal, rather than in polysilicon. At first glance, this may seem to be an undesirable thing to do, because (1) charge leaks through a deposited (metal-metal) oxide at a much faster rate than it does through the thermally-grown gate oxide, and (2) routing polysilicon to a metal layer consumes layout area. Unfortunately, in the Orbit process, the threshold voltage for a field-oxide transistor is approximately 20V. If the polysilicon exits directly the n- well-drain, when V_{dg} exceeds 20V a parasitic p-type MOS channel forms in the n-. When V_{dg} exceeds 35V, junction breakdown occurs where the field-oxide-induced, p-type channel met the n+ well contact. Although we hope that a more modern process, with its lower tunneling voltages, will permit us to route the floating gate entirely in polysilicon, in the present process the floating gate must exit

the well-drain on a metal layer.

To turn on the tunneling, we lower $\overline{down}$ to ground; to turn it off, we raise $\overline{down}$ to 5V. By controlling how long $\overline{down}$ is at ground, we control how much charge is tunneled onto the floating gate. Because the M3-M4-C_{int} feedback loop pins the floating-gate voltage, the voltage across the tunneling oxide is constant while tunneling; the oxide current therefore also is constant. As a result, the memory cell output voltage decreases linearly with $\overline{down}$ pulse width.

2.5 Pfet Memory

During initial testing of the memory cell, long-term storage of output voltages close to ground was found to be impossible. The output voltage drifted upwards, and appreciable error would manifest over a few hours. This source of this drift was found to be injection in the pfet of the output amplifier. The pfet injection was not an expected effect, and was observed because of two unusual aspects of the pfet's operation:

- Operating a pfet with a high V_{ds} due to the higher than normal V_{dd} mandated by the higher threshold voltage of the pbase transistor, and
- A floating gate on the pfet.

The gain of the system also inherently amplified this effect. Colleagues in the Physics of Computation Laboratory at Caltech have used this effect in a variety of novel circuitry.

The addition of the cascode transistor eliminates the pfet injection by pinning the V_{ds} of transistor with the floating gate at a constant voltage.

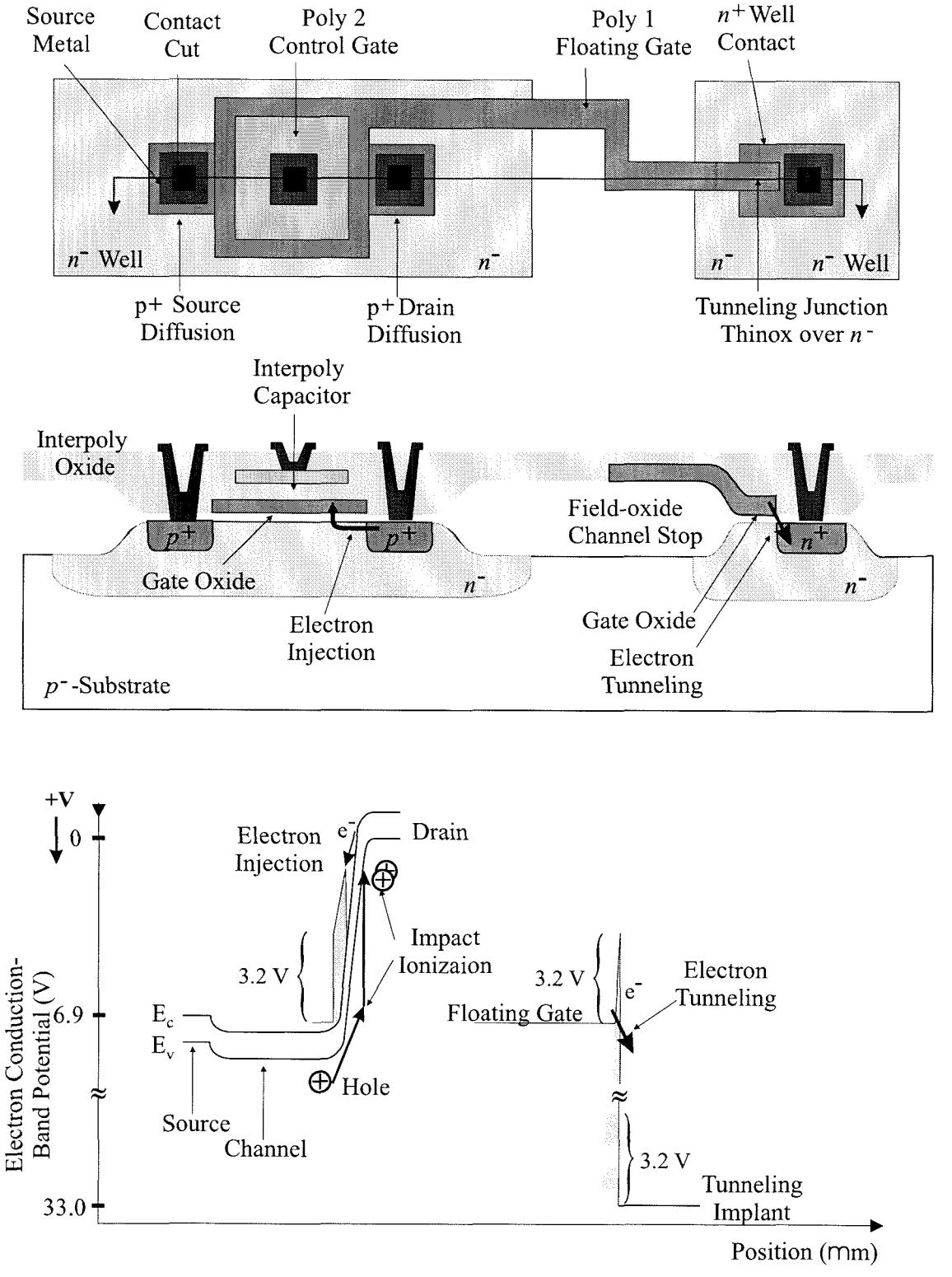


Figure 2.6: Pfet injection band diagram

Injection in the pfets allows the replacement of the pbase/nfet injection transistor with a pfet. This is advantageous as it removes the necessity of the special pbase layer, and makes the memory cell compatible with standard CMOS fabrication processes. Figure 2.7 shows a pfet based memory cell. The operation of the cell is similar to the pbase version. The injection is controlled by pulling the drain of the injection pfet to ground. The V_{ds} of the transistor controls the injection current.

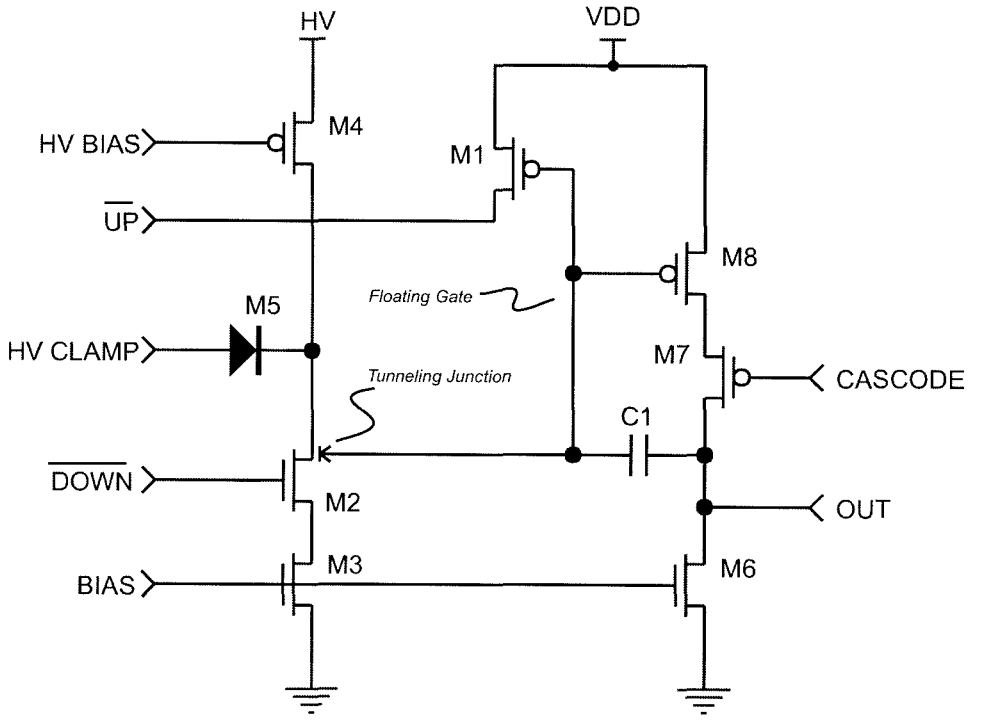


Figure 2.7: Schematic of the memory cell using a pfet for injection.

2.6 Results

I fabricated the memory cell in a wide variety of technologies. The nfet memory cell was limited to the 2.0 micron technology because of the availability of the pbase implant. The pfet memory cell was fabricated in the 2.0 micron Orbit, the 1.2 micron Orbit, the 1.2 micron AMI, the .8 micron HP, and the .35 micron TSMC process. In this section I discuss the results of testing the nfet memory cell.

The fabricated memory cell is shown in Figure 2.8. The memory cell pictured is the pfet memory cell in the AMI 1.2 micron process. The area required by the nfet and pfet memory cell are roughly the same, as they only differ by the single injection transistor. The results of testing the memory cell as an ana-

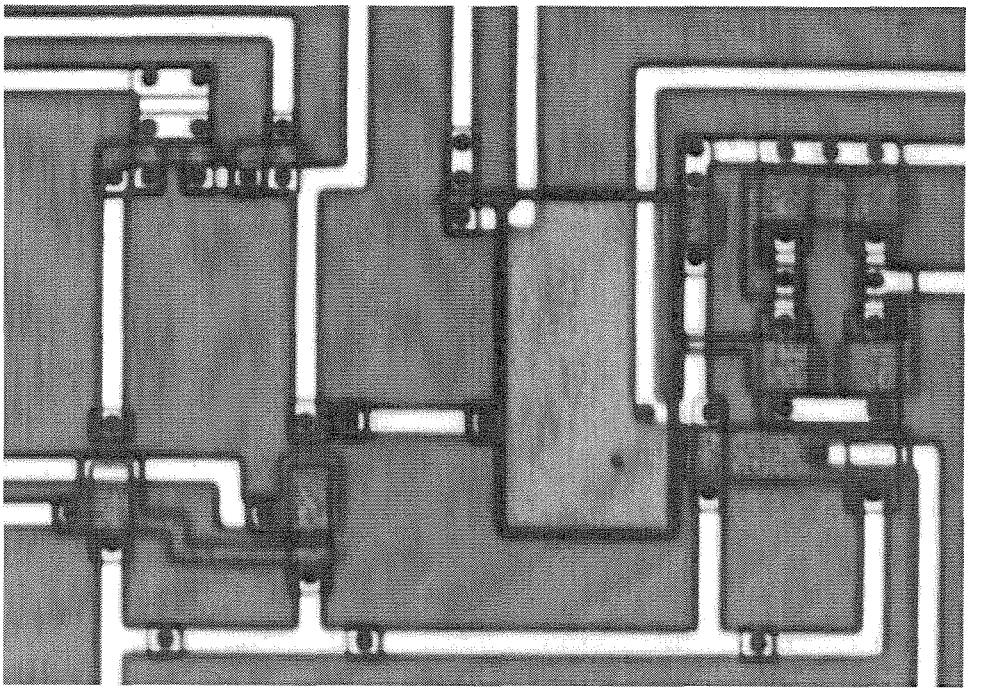


Figure 2.8: Picture of the fabrication memory cell

log counter are shown in Figure 2.9. I set the tunneling voltage at 40V and the VDD to 6.5V. The pulse width was selected to produce approximately 10 bits worth of count data. The output voltage of the count is plotted against the iteration count. The count begins after resetting the output of the memory cell to its lowest output, counting up (via injection) until the maximum VDD is reached, and then counting down (via tunneling).

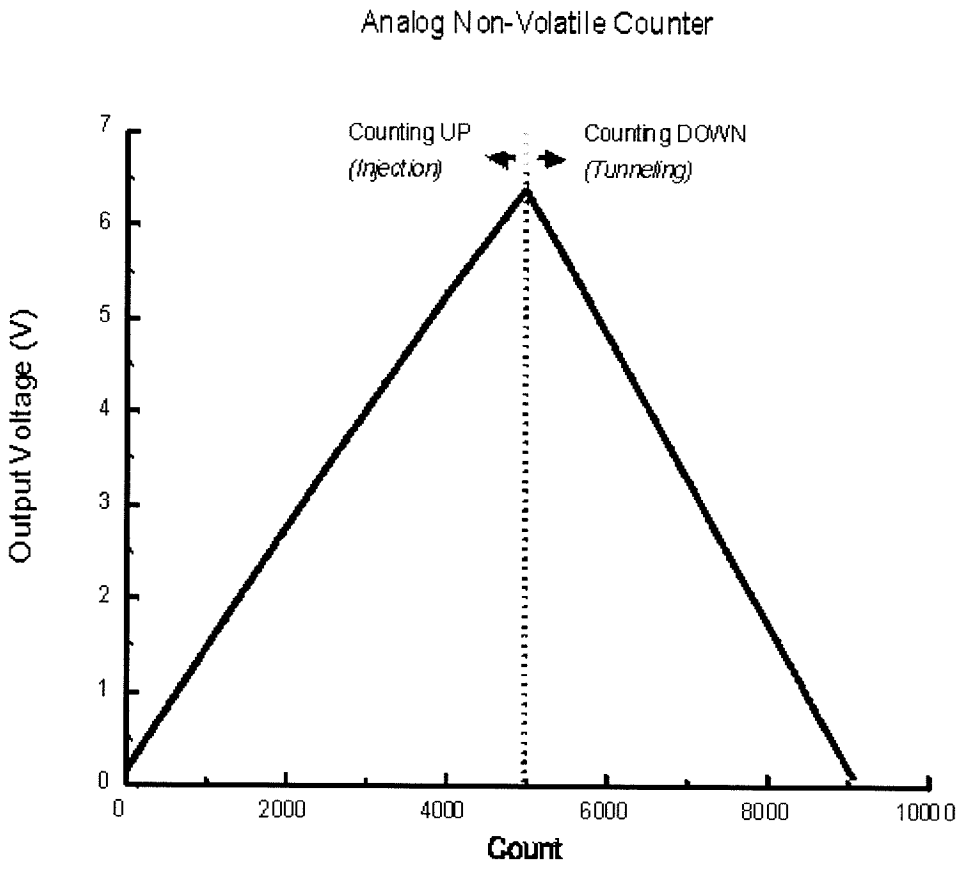


Figure 2.9: The operation of the memory cell as an analog counter.

The use of the memory cell as a pulse width modulated storage device is demonstrated in Figure 2.10. For this data I set the tunneling voltage at 40V, the injection VDD at 6.5 V, and varied the pulse width of the programming pulse.

The data are plotted on a log-log scale, showing the linearity of the change in memory voltage vs. pulse width over a large range of inputs.

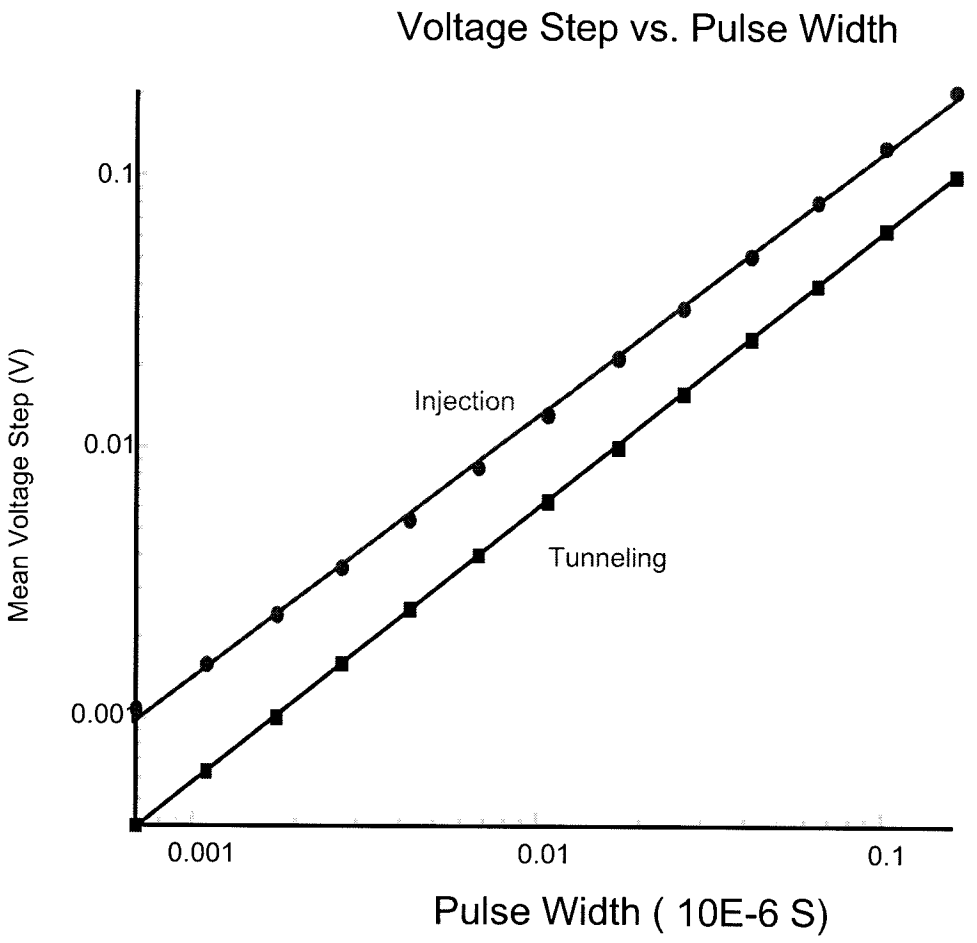


Figure 2.10: Change in output voltage vs. pulse width.

Figure 2.11 presents the delta in output voltage vs. the output voltage, to examine any dependencies between operating voltage and the injection rate. The data shown are for a VDD of 6.5V, a bias voltage of 0.75 V, and a pulse width of 50ms. The distribution histogram of the delta data is shown in Figure 2.12.

The same analysis is presented in Figure 2.13 and Figure 2.14 The data

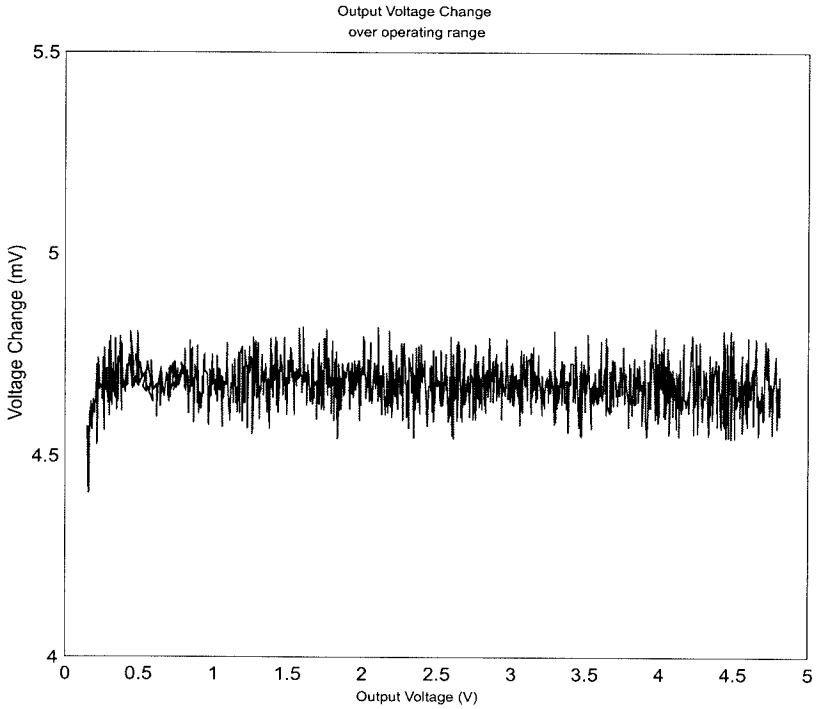


Figure 2.11: Change in memory voltage vs output for injection

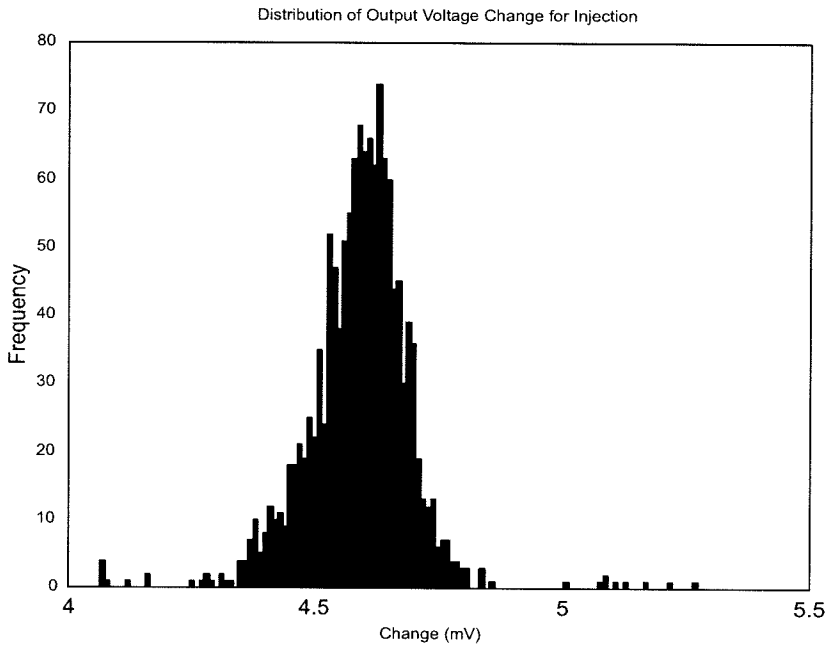


Figure 2.12: Distribution of injection deltas for a pulse width of 50ms

shown are for a VDD of 6.5V, a bias voltage of 0.75 V, and a pulse width of 50ms, and a tunneling voltage of 35V.

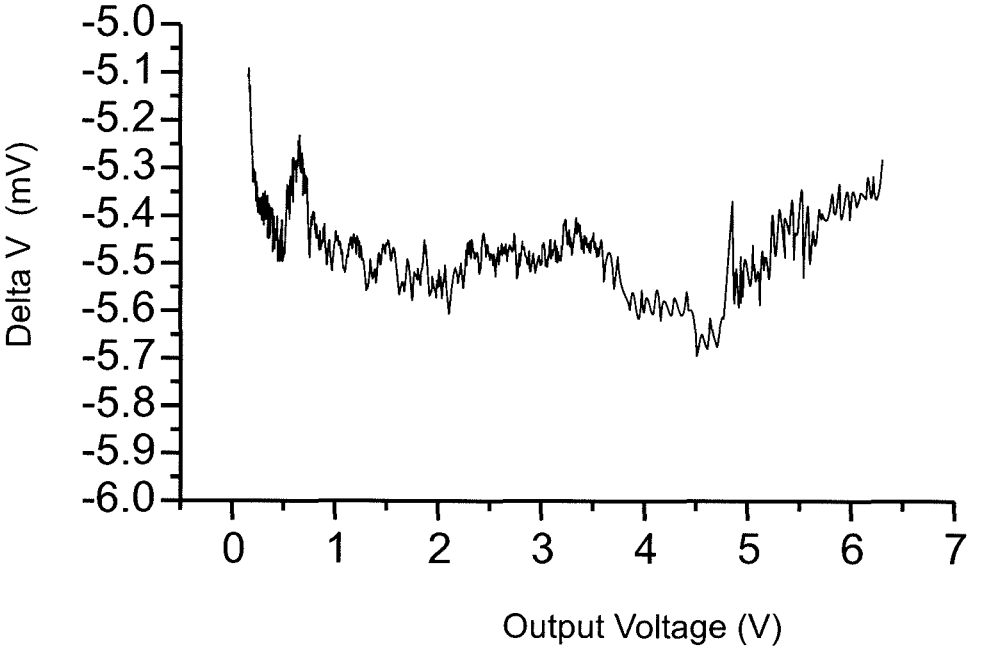


Figure 2.13: Change in memory voltage vs output for tunneling

2.7 Conclusion

In this chapter I have presented a nonvolatile memory element. This device has a number of applications, from parameter storage to an adjustable weight storage device for neural networks. It is different from other reported devices in that it possesses local control of the tunneling voltage in a standard CMOS process. This makes it suitable for implementation in arrays.

One drawback of its use in the older 2.0 and 1.2 micron processes is the high tunneling voltage required for tunneling across the gate oxide. The smaller fea-

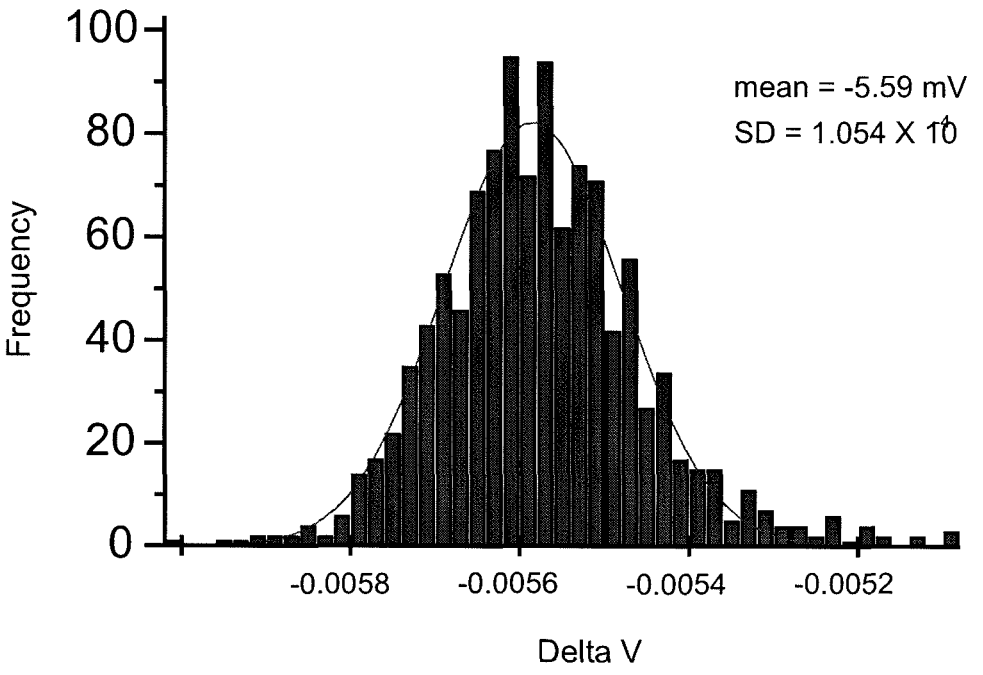


Figure 2.14: Distribution of tunneling deltas for a pulse width of 50ms

ture size processes have much thinner oxides, and thus the situation becomes more favorable as the feature size is reduced further.

Another drawback is that the programming of the memory is its speed. In a neural network application the rate of weight updates is the performance limiting factor during the learning process. For some applications the speed of the memory cell may be too slow to support high speed learning. However, during high speed learning there is no need to store the weights in a nonvolatile manner, in fact it is inefficient to do so. A scheme whereby a short term memory consisting of a capacitor and a long term memory using this cell is a possible solution.

CHAPTER

THREE

Ratiometric Sensing Building Blocks

Sensor transduction is a very important aspect of any sensor design. An integrated sensor design faces the additional challenges that it must be compatible with the available technology. The flexibility of choosing from a variety of technologies and devices is not available. The carbon black based conductive sensor used in this work produce a ratiometric output: while nominal resistances may differ, the percentage change is consistent for like sensors. This property can overcome the limited control over the sensor resistances. This chapter explores a sensing approach that outputs the ratiometric output of the sensor, while compensating for baseline drift.

3.1 Ratiometric Sensing: The Wheatstone Bridge

Figure 3.1 shows a typical Wheatstone bridge configuration, consisting of three biasing resistors, the sensor resistor, and a differential voltage measurement. The circuit outputs a voltage proportional to the percentage change of the sensing resistance, not by some magical computation, but by judicious selection of the bias point. Other configurations of the bridge are possible, but for the purposes of this discussion, I will focus on the most simplistic.

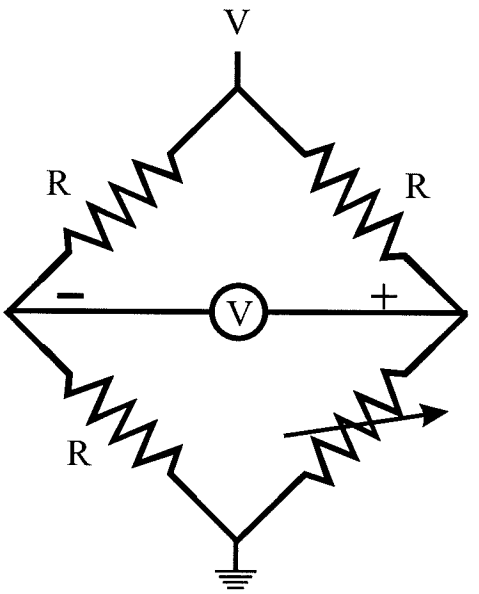


Figure 3.1: The Wheatstone Bridge.

The two resistors R_1 and R_2 on the left side of the bridge simply provide a reference voltage for the calculation. They could simply be replaced by a voltage source equal to

$$V_{dd}\left(\frac{R_1}{R_1 + R_2}\right)$$

for this discussion. However, the resistors are an easy and efficient means of

generating the necessary voltage.

The right side of the bridge formed by R_3 and R_{sensor} perform the sensor transduction. The bridge is balanced in the steady state by adjusting the value of R_3 such that the differential voltage across the bridge is zero. If R_1 and R_2 are equal, then setting R_3 equal to R_{sensor} will satisfy this requirement. In a typical system this is generally accomplished by fabricating R_3 and R_{sensor} in close proximity with the same physical parameters. For example, in a strain gauge application R_3 and R_{sensor} are fabricated near each other, the only difference is that R_{sensor} is the resistor that is subjected to the applied force. R_3 can also be an adjustable resistor, such as a potentiometer, adjusted until R_3 equals R_{sensor}.

The simple voltage divider ratio determines the steady state voltage of the sensor node as

$$V_{sensor} = V_{dd}\left(\frac{R_{sensor}}{R_3 + R_{sensor}}\right).$$

If now the resistance changes by $\triangle R$ in response to chemical exposure the voltage will be

$$V'_{sensor} = V_{dd}\left(\frac{R_{sensor} + \triangle R}{R_3 + R_{sensor} + \triangle R}\right).$$

If we assume that $\triangle R << (R_3 + R_{sensor})$ we can eliminate $\triangle R$ from the denominator, giving

$$V'_{sensor} = V_{dd}\left(\frac{R_{sensor} + \triangle R}{R_3 + R_{sensor}}\right).$$

Computing the change in voltage we have

$$\begin{aligned}\triangle V &= V'_{sensor} - V_{sensor} \\ &= V_{dd}\left(\frac{R_{sensor} + \triangle R}{R_3 + R_{sensor}}\right) - V_{dd}\left(\frac{R_{sensor}}{R_3 + R_{sensor}}\right)\end{aligned}$$

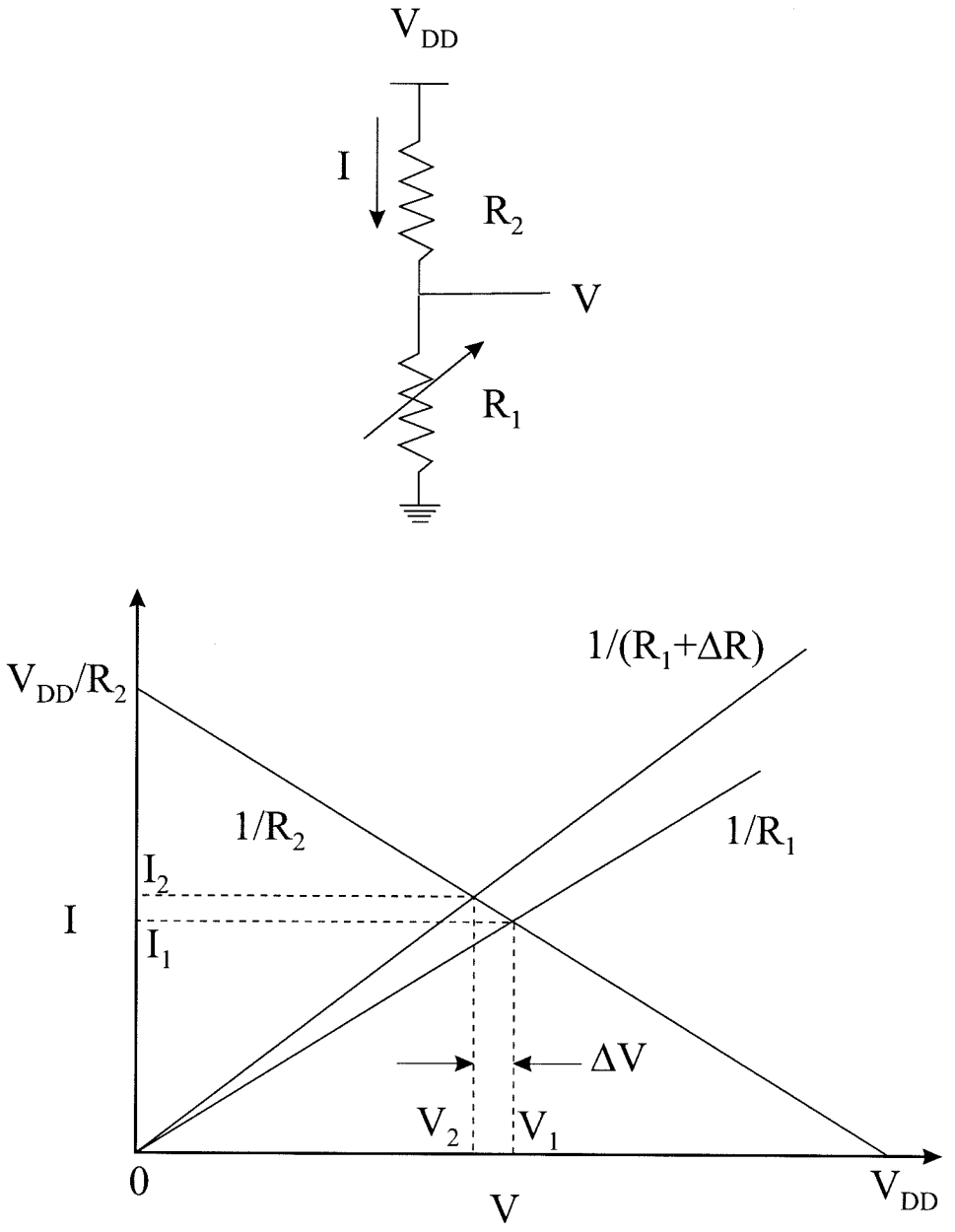


Figure 3.2: Resistance load lines for a Wheatstone bridge.

$$= V_{dd}\frac{R_{sensor} + \triangle R - R_{sensor}}{R_3 + R_{sensor}}$$
$$= V_{dd}\frac{\triangle R}{R_3 + R_{sensor}}$$

Recall that the left side of the bridge was also V_{sensor}, so that by taking the differential voltage between the left and right sides of the bridge we don't need to remember the initial voltage across the sensor. Therefore, the circuit computes an output voltage proportional to the ratiometric change in sensor resistance, neglecting the error caused by our simplification of $\triangle R << R_3 + R_{sensor}$.

We can minimize this error by making R_3 larger. However, this will force the size of $\triangle V$ smaller, and make detection difficult. Moreover, a larger V_{dd} voltage may be required to keep the sensor voltage within acceptable ranges for measurement apparatus.

3.2 An Active Bridge Circuit

For our application of chemical sensing, the fabrication of a balanced bridge is not as straightforward as making two resistors. We have not yet developed deposition methods to the point where we can guarantee the fabrication of matched resistances. In addition, even if such methods were available, the matching resistance would have to be shielded in some manner from exposure to the chemical vapors, complicating fabrication. One alternative would be to fabricate the matching resistance in another material. Again, because we cannot know in advance what the value of the sensor resistance will be, we must either form the matching resistance after the sensor resistances, or modify the resistances after fabrication. Precision analog fabrication, for example,

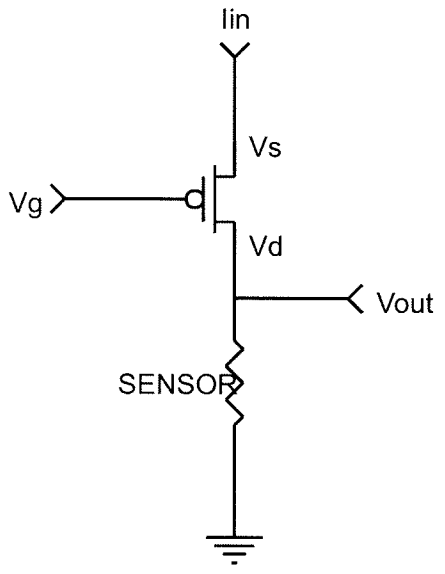
employs laser trimming of resistances to compensate for mismatch and offsets. However, laser trimming is expensive, time consuming, and not reversible and therefore is impractical for our goal of developing on chip arrays. The arrays require a resistance that is adjustable, over a large range of target resistances. It is desirable that we have the ability to adjust the resistance after the initial calibration to compensate for sensor drift or temperature changes.

Fortunately, such a resistance is obtained with a transistor, with some advantages and some caveats. In this section I discuss the use of a transistor as an element to balance bridges and discuss two configurations to achieve it in practice.

Figure 3.3 shows the balancing resistor replaced with a p-type MOS transistor. The transistor can act as a voltage-controlled resistor, with the gate voltage Vg controlling the resistance. The load lines of the transistor/resistor circuit are shown in 3.3, using transistor data from an AMI 1.2 micron run available through MOSIS.

The current in the transistor is dependent mainly on the voltage between the gate and the source. The voltage on the drain has a mild impact on the current, illustrated by the slight slope of the current at the region of intersection. This means that only the gate voltage on the transistor will set the current through the sensor, and the current will be independent of the voltage across the sensor. Since the voltage across the sensor is proportional to its resistance, this has a beneficial impact on the operation of this circuit for ratiometric sensing.

There are many approaches to improving the transistor current source, such as cascoding or more sophisticated feedback schemes.



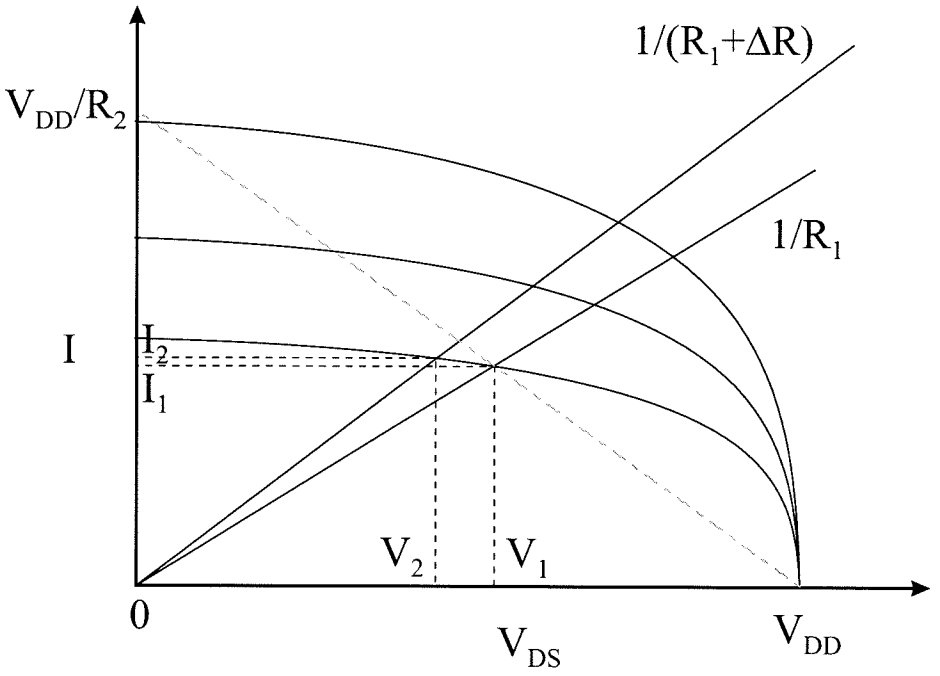


Figure 3.3: Load lines for the transistor-based balancing scheme.

3.3 Integrated Active Bridge Circuit

In the previous section I showed how a transistor can be used as an element of a ratiometric sensing circuit. However, there is still the necessity to balance the bridge, just as there was with the resistor based scheme. With the resistor balancing the problem was the adjustment of the balancing resistor was difficult and for practical purposes irreversible. While we still need a means to adjust the balance of the bridge, at least we have traded off a physical parameter, the resistance, with an electrically adjustable parameter: the gate voltage.

One method to adjust the gate voltage is to bias the transistor with an offchip source, such as a potentiometer. For even a small number of sensors on a chip this is intractable. Instead, a digital to analog converter could be used to set the necessary bias voltage. Another method is to store the necessary gate voltage on a capacitor, and then refresh the capacitor voltage from offchip, similar to the refresh scheme discussed in the previous chapter. However, we then have to contend with the droop on the capacitance voltage affecting the bias point of the sensor, and although circuit designs could minimize its impact, for systems with amplification even this droop may be intolerable.

One solution without these drawbacks is to use the memory cell described in the previous chapter to store the gate voltage. This eliminates the need for refreshing the voltage, and the small area makes it very suitable for array implementation. However, this system could still not tolerate drift in the sensor resistance, so periodically the bias would have to be checked and adjusted if necessary. Fortunately the time scale of this update is measured in hours if not days, so that requirement is not as demanding as in the case of refreshing.

3.4 Continuous Time Ratiometric Biasing

An alternative to the schemes discussed in the previous section is to employ feedback to attempt to keep the bridge balanced. One possible implementation is shown in Figure 3.4. An amplifier adjusts the gate voltage of the sensor biasing resistance at the desired set point. The capacitor is used to slow the response of the amplifier. Since we are trying to observe the change in the resistance due to the presence of odors, it is necessary to give the sensors a chance to react. If the response of the amplifier is too fast, it will not only adapt out any baseline drift, but also the signal we are trying to observe.

The speed of the amplifier's response to a sudden change in the voltage across the sensor is limited by the rate at which it can charge the capacitor on its output. This is the current in the bias leg of the amplifier.

The circuit of Figure 3.4 includes a clocked transistor so that the bias current can be further reduced by a PWM scheme. This enables the amplifier to be switched off, so that the only current impacting the capacitor is the leakage through the nfet and pfet at the output. The nfet leakage current will be toward ground, while the pfet leakage current will be towards VDD. Depending on a variety of factors including capacitor voltage and operating temperature, the net current could be either towards VDD or towards GND.

Because of the leakage currents it is very difficult to make the response of the circuit slow enough so that slow sensor responses, which may take minutes to manifest, are not nulled out by the feedback. While modern technologies, with their thin gate oxides, allow large onchip capacitances, this system is unsuitable for slow sensor responses.

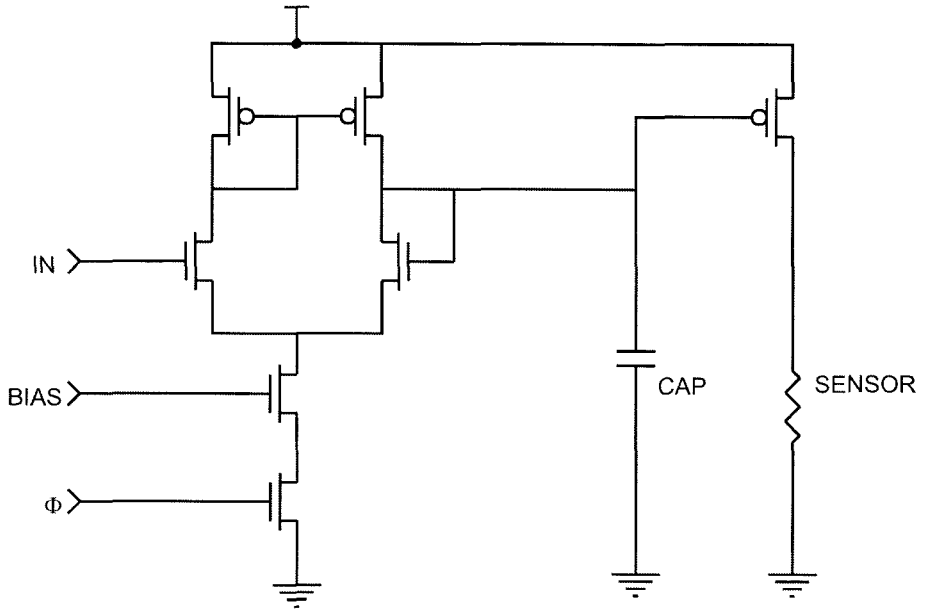


Figure 3.4: Continuous time adaptation of the bias.

3.4.1 Adaptive Ratiometric Circuits based on Oxide Currents

Minch proposed interesting continuous mode approach to the ratiometric biasing of resistance sensors [13], shown in Figure 3.5. The circuit consists of an autozeroing amplifier where the active load is replaced by the sensor. To improve current range a bipolar transistor is employed.

The circuit consists of a tunneling node at T1, and a pfet injector at Q1. The circuit dynamics are controlled by the competition of these two complementary oxide currents. Recall from Chapter 2 that the injection current in the pfet injector is controlled by the voltage across its source and drain. In this circuit the source of the transistor is at VDD, and the drain of the transistor is offset from the output voltage by the base-emitter voltage drop of the bipolar transistor.

The major drawback of this circuit is that the DC operating point of the

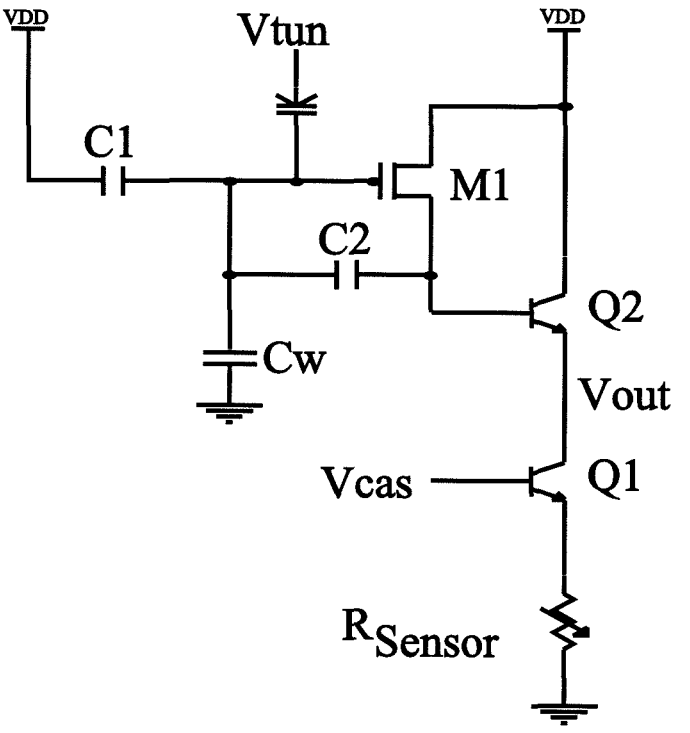


Figure 3.5: An adaptive continuous circuit for ratiometic sensing.

circuit is set by the output voltage required for the two currents, injection and tunneling, to equilibrate. Since the currents are heavily dependent on the oxide trapping, process parameters, and temperature, the DC voltage of the amplifier will not likely match well between cells on an array. While the sensor response can be extracted from the DC signal, the response time of the sensors makes that impractical.

This benefit of this circuit is that now we are not limited by leakage currents to perform the adaptation, but by oxide currents, which are many orders of magnitude smaller that picoampere range leakage currents (see Chapter 2). This circuit allows time contacts from seconds to days. These time constants are exactly what are required for the sensor biasing.

Figure 3.6 shows an alternative solution to the problem. Instead of having the output voltage and the balancing of the injection current be dependent, the circuit uses feedback to decouple the two voltages. As before, a constant tunneling voltage is present; however, the injection current is controlled by feedback from the output node, rather than by the output node directly. This allows the DC output voltage to be set arbitrarily, and the feedback circuit will determine the correct Vds on the pfet injector to equilibriate the currents.

A full stability analysis of this circuit is beyond the scope of this thesis. It is a difficult problem, as the time constant of the system changes exponentially with the feedback error. Empirically, if we model the tunneling response as a low-pass filter, we can set the location of the zero at a frequency low enough such that the gain is rolled off well before the system is unstable, at least for a low gain. The higher the gain of the feedback, the slower the necessary response of the tunneling element.

Figure 3.7 shows the time response of the circuit. Like the Minch circuit,

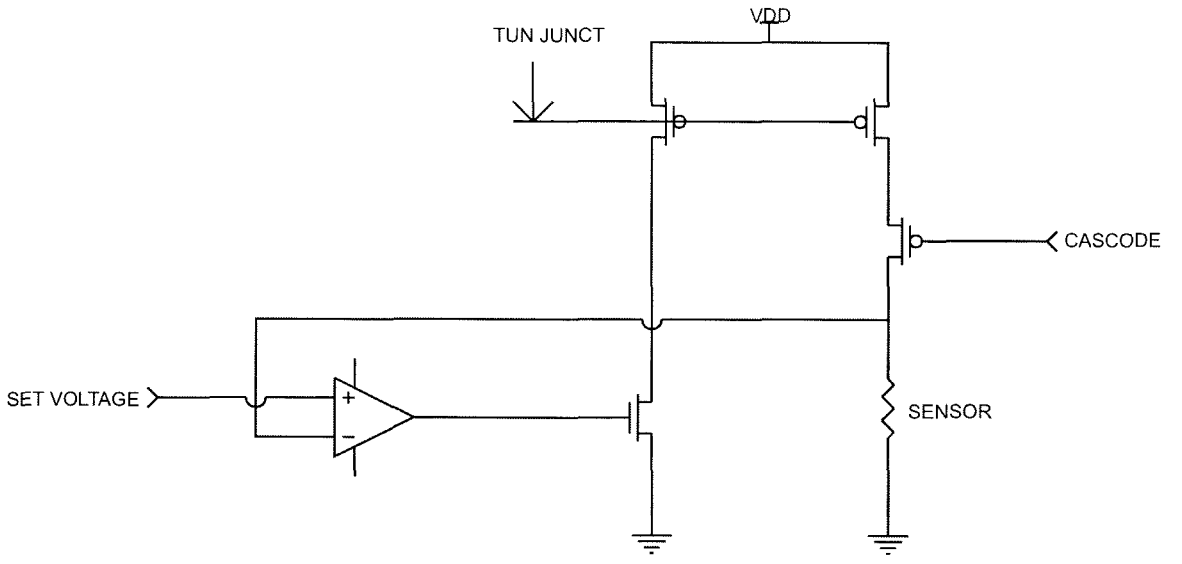


Figure 3.6: An adaptive feedback method of ratiometric balancing.

the system exhibits long time constants. However, note that the DC operating point is set arbitrarily. This circuit is therefore better suited to be followed by a differential amplifier to provide gain to the system. The offset voltage of the feedback amplifier determines the error in the desired setpoint and the actual DC output value. Adding gain to the system would reduce this error, although this can quickly make stabilizing a system difficult, since the dominant pole of the feedback system has to be at a frequency low enough to roll off the loop gain before the system is unstable.

This circuit has the desirable characteristics of a known DC operating point. In addition, the DC operating point will not be affected by changes in the injection or tunneling current due to trapping. The feedback will set the V_{ds} of the injection transistor to compensate for the change in current.

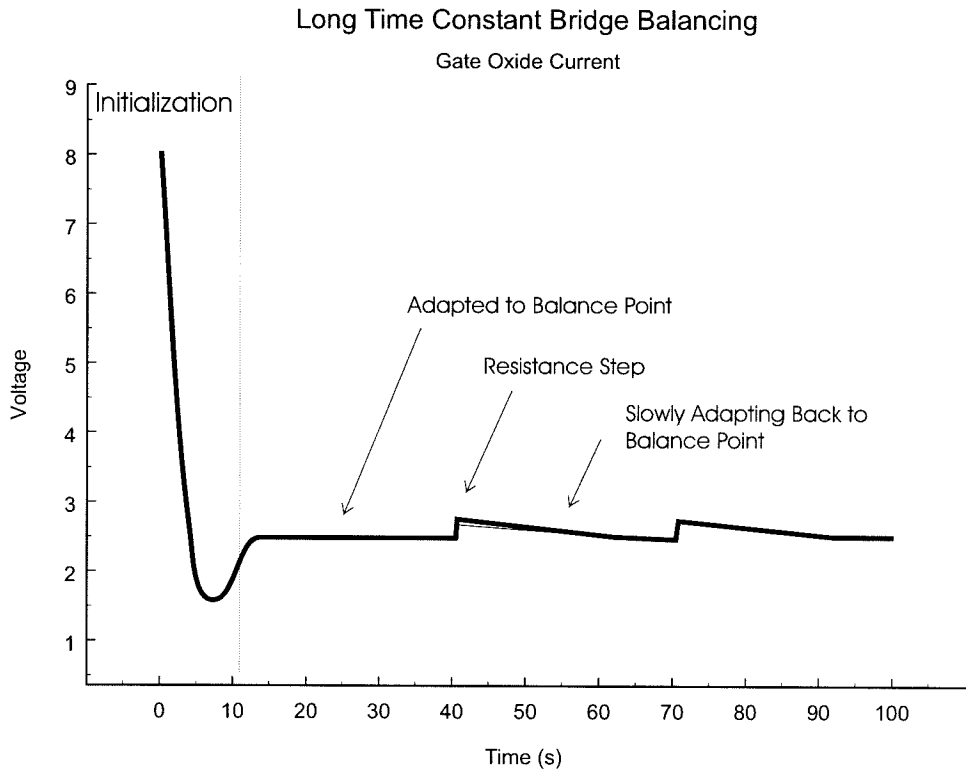


Figure 3.7: Time response of the adaptive feedback circuit.

3.5 Conclusion

In this chapter I have presented circuits for the biasing and amplification of chemical sensors. The continuous time balancing circuit can be applied to a variety of systems where a slow response time is required. For the chemical sensors that are the focus of this thesis it provides a method of biasing the sensor, compensating for drift while allowing the output to respond to chemical responses. The post-processing of integrated circuits to allow this circuit to be integrated with sensors on-chip is the focus of the next chapter.

CHAPTER

FOUR

Postprocessing of Integrated Circuits for Chemical Sensors

Integration of the chemical sensors with a standard integrated circuit process has a number of major advantages. A CMOS process is a complicated series of steps, and requires sophisticated and expensive facilities. At this writing there are very few companies who are able to maintain fabrication facilities. For a university setting, a CMOS process is a very difficult undertaking, and the few universities that have fabrication facilities are far from the state of the art. While having a dedicated fabrication facility would make prototyping and research much easier, it is unfortunately totally impractical. In this chapter I describe the difficulty in integrating the carbon black thin film sensors with integrated circuits, and discuss the results of an electroless plating process for its solution.

4.1 The Fabrication Process

Since the goal of this research is to ultimately combine chemical sensors and integrated circuits, in this section I will present a brief overview of integrated circuit fabrication. For this simple overview, I will discuss a p-substrate (n well) process with two metal layers, similar to the processes used for the circuits in this thesis.

4.1.1 Wafer

The transistors and wires that make up an integrated circuit, and ultimately the chemical sensors, are fabricated in or on a thin slice of silicon. Typically the wafers are cut from an ingot of single-crystalline silicon, fabricated using the Czochralski method. A single crystal of silicon, the seed, is drawn slowly from melted metallurgical-grade silicon. This produces an ingot, usually 1 meter in length and 10 to 20 cm (4 to 8 inches) in diameter, depending on the speed of the pull and the rotation rate of the seed. To create a p-type substrate, silicon heavily doped with boron impurities is added to the molten silicon. This ingot is sliced into wafers approximately 1 mm thick, and then the wafers are polished.

4.1.2 Well Formation

The n-type transistors are formed in the p-substrate. Formation of the p-type transistors require the creation of n regions on the substrates, called wells. These wells, and many of the subsequent steps, are formed using a photolithographic process. The entire wafer is coated in a photoresist, a light sensitive

polymer, which is then exposed to radiation through a mask. The mask is a sheet of glass with a pattern to selectively expose regions of the photoresist to the radiation. Where the light hits the photoresist the polymer chains cross-link and make it insoluble to an organic solvent. An organic solvent then removes the unexposed areas of photoresist Baking then hardens the remaining photoresist. Another type of photoresist, called positive photoresist, allows solvent to remove the areas exposed to the light.

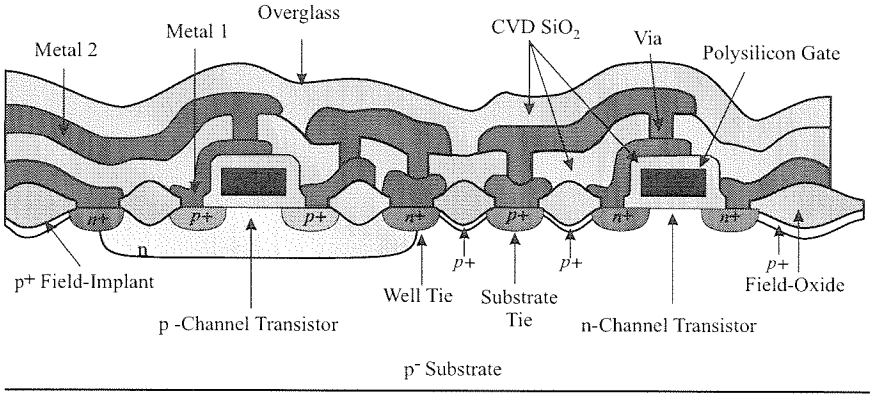


Figure 4.1: A typical CMOS process.

For well formation, a thin layer of S_iO_2 is grown on top of the wafer, by heating the wafer in the presence of water vapor.

$$Si + 2H_2O -> SiO_2 + 2H_2$$

Or oxygen gas:

$$Si + O_2 -> SiO_2.$$

Note that in both cases the formation of the oxide consumes Si from the Silicon wafer. The photoresist is then patterned. The areas of the S_iO_2 that are not covered by photoresist after the patterning are then etched away, and then the

photoresist is removed from the wafer. The S_iO_2 that remains protects the p-type substrate, and the openings in the S_iO_2 define the location of the n-type wells. There are two methods to introduce the necessary n-type dopants to create the wells. Ion implantation accelerates dopant molecules through an electric field into the wafer. The diffusion process exposes the wafer to gas containing phosphorus, and the phosphorous atoms will diffuse into the wafer.

4.2 Post Fabrication Processing

Ultimately the sensor material must make electrical connection to the underlying contacts and circuitry. The metal layers in an integrated circuit are typically composed of aluminum, although the current state of the art technologies are beginning to switch to copper interconnects. There are two common methods to connect an integrated circuit die to the outside world: wire bonding and, for the higher density chips, bump bonding. In both cases etching a window in the top layer of glass exposes the top layer of aluminum to allow connection.

Unfortunately, we are not able to simply take our chips and deposit the sensor material on the exposed aluminum. While aluminum is an adequate conductor of electricity, it also rapidly grows a native oxide on its surface in air. In many applications, this oxide is a beneficial property of aluminum. The formation of the oxide increases the chemical resistance of aluminum. Unfortunately, as the oxide is an electrical insulator, it is also resistant to our efforts to fabricate chemical sensors. The inability to make electrical contact with the aluminum was one of the major challenges to fabricating sensors on chip.

One solution is to etch the oxide off the aluminum with acid, and then deposit the sensor material before the oxide can grow. This would require keeping

the chips in an inert environment, such as Nitrogen gas, while depositing the sensor material. However, the sensor material would not prevent oxygen from diffusing to the aluminum surface, and the oxide would soon reestablish itself.

Therefore, aluminum is not an appropriate metallization for making integrated chemical sensors. Palladium or gold plating is usually applied to surfaces that are used in switches. These metals do not form oxides with exposure to the atmosphere, at least at reasonable temperatures, and so are ideally suited for our application. In order to fabricate integrated sensors, we must find a way to deposit a better material on top of the aluminum. The difficulty increases because we are constrained to dealing with individual die.

One method I explored for making sensor contact is related to the wire-bonding employed to connect an integrated circuit to the outside world. A gold or aluminum wire is placed onto an aluminum pad and vibrated at ultrasonic frequencies. The heat causes the wire to melt and bond with the underlying aluminum. Figure 4.2 shows a ball of gold placed on two aluminum chip pads. These balls were placed by an ACME ball bonder. In a typical application a wire would extend from the ball of gold and make contact with the chip package for connection to external circuitry; however, here the wire is not included. Deposition of a chemical sensor between these two pads successfully fabricated a sensor. However, this approach is not suited for arrays because of the large number of required bondings. In addition it is not suitable for the ring structure to be described later in Section 5.1.2.

Another approach is to sputter gold onto the die and pattern it. This technique is used for TAB bonding applications, where instead of wires ultrasonically attached to a chip pad, a small ribbon cable with conductors is lined up on the pads and soldered. Low profile applications such as portable equipment

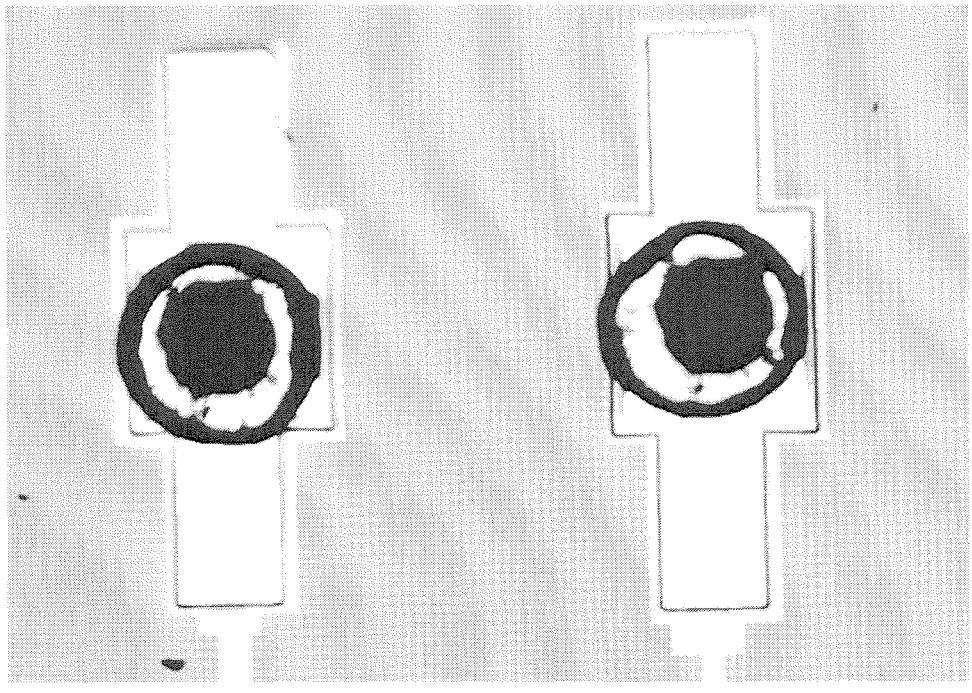


Figure 4.2: Balls of gold placed by bonding equipment on aluminum pads.

employ this technique. The process is shown in Figure 4.3. The process begins with a sputtering of a chrome/tungsten adhesion layer on the wafer, followed by a thin layer of gold. The entire wafer is covered by this operation. Then a masking layer is patterned so that only the pads are exposed, and gold is electroplated onto the exposed gold areas to create a thick gold layer. An organic solvent then removes the resist. An acid etches the thin gold and seed layer, leaving the isolated gold pads. While it would solve our contact problem, unfortunately this process is impractical, as it requires patterning steps. On an individual die this would be extremely difficult to perform. Since for prototyping an entire wafer is impractical, especially on research budgets, another solution is required.

4.3 Electroless Plating

While electroplating is an attractive option, the requirement of sputtering on a seed layer for adhesion, and the need to pass current to facilitate the plating mechanism is a major drawback. Electroless plating, discovered in 1944 by A. Brenner and G. E. Riddell, uses an autocatalytic reduction to perform the plating. Thus there is no need to provide a path for electrical current to all the sensor sites.

While electroless plating of aluminum is not a new technology, formulations compatible with integrated circuits are a recent development. Two companies have recently introduced electroless gold plating for integrated circuits. The use of an electroless solution is suitable for this application, as a masking step is not required. In this section I discuss my results with the proprietary formulation from Stapleton Technologies.

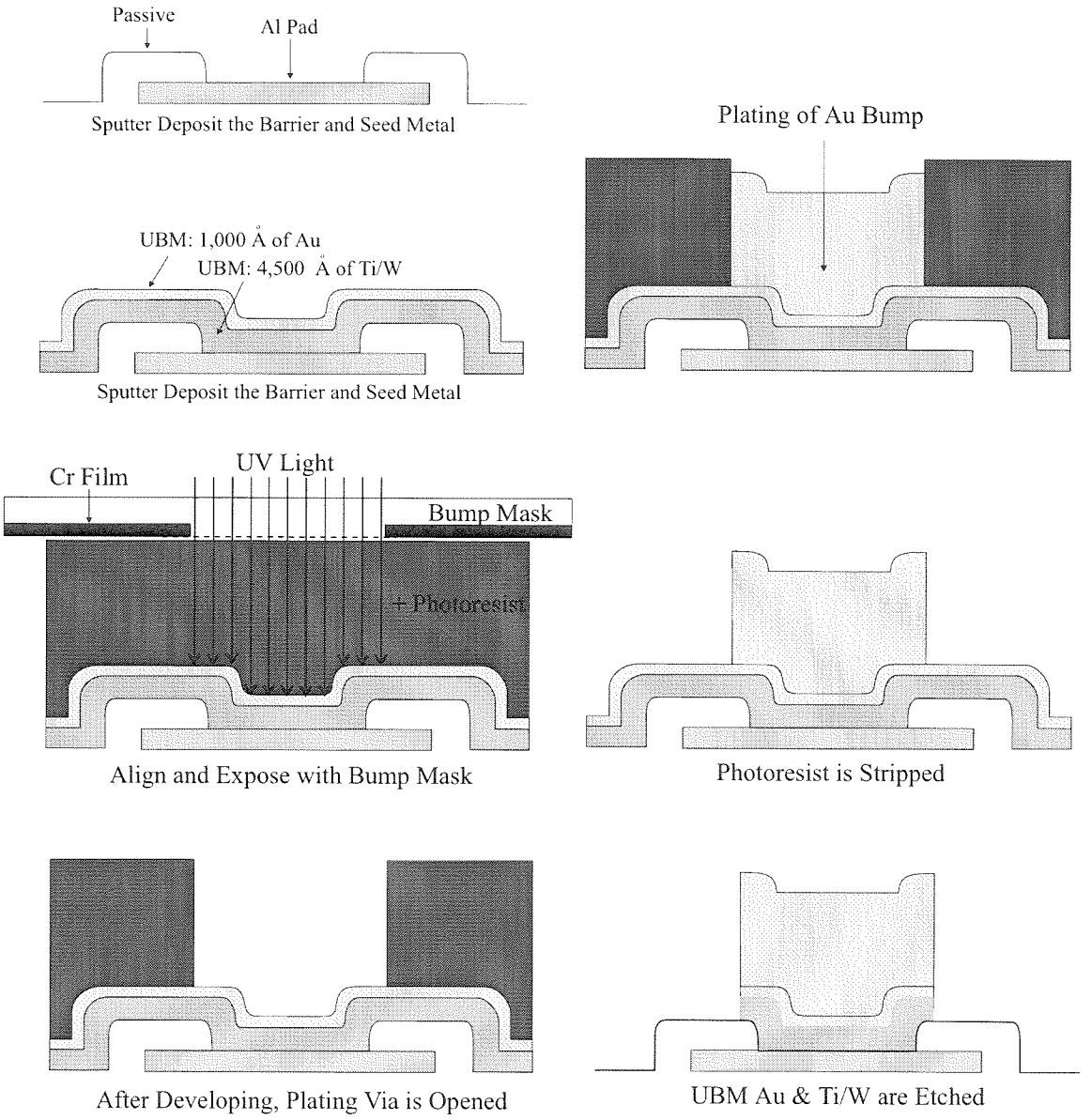


Figure 4.3: Process flow for physical deposition of gold on aluminum pads.

4.3.1 Equipment

The plating process is intended for a large scale application with sophisticated process control; however, this is beyond the scope of our capabilities. The small-scale prototype plating process described in this system used very basic instruments. All plating operations were performed under a fume hood using 250 ml Pyrex beakers. Two hotplates with quartz temperature probes provided the heating and magnetic stirring. An Acmetech Static2000 provided a balanced stream of ions over the entire workplace to protect from static charge, which could damage the sensitive integrated circuits. The ionizer was turned off while plating the chips to prevent interference with the reduction reactions.

4.3.2 Experimental Procedure

This section describes the procedure I followed to plate the die. The procedure is summarized in Table 4.1.

Die Preparation

The first step of the plating process is to attach the individual die to another substrate so that processing is possible. I investigated a couple of approaches to solve this problem, including machining a polystyrene holder with a Viton gasket to hold the individual die. However, because of the many different sizes of the die I adopted a simpler and more flexible approach. I used photoresist to glue the die to small sections of silicon wafer, which were in turn glued to glass microscope slides. The adhesion of the photoresist to silicon was much better than to glass. Treatment of the glass slides with HMDS improved the adhesion of the photoresist by making the surface of the glass hydrophilic; however,

Step	Solution makeup (ml)			Additive	Temp(C)			PH	Adj Sol	Note
0	Surface cleaning (descum....)									
1	M209A		DI water	N/A	T,start	T,set	time(min)	N/A	N/A	Al cleaner
	10		190		40	40	1			Change daily
2	Nitric Acid		DI water	N/A	T,start	T,set	time(min)	N/A	N/A	Deox for Al
	100		100		Room Temp		0.5			
3	M262		DI water	N/A	T,start	T,set	time(min)	3.1	N/A	Acid Zincate
	180		20		Room Temp		1-1.5			
4	M282SX	NH_4OH	DI water	M282B	T,start	T,set	time(min)			Ni plating
	100		100		91	91	7.5um/10min	5.3		
low temp	100	PH adj	100		40	65	4.5um/10min	**7.3**	NH_4OH	
5	M290		DI water	M290	T,start	T,set	time(min)			immersion gold
	24		176		80	80	5	6.5		
low temp	24		176		40	65	5	6.5		
6	M294A	M294B	DI water	M294B,C	T,start	T,set	time(min)			autolytic gold
	100	20 (37.5)	80		65	65	0.5um/10min	7.8	$10\%H_2SO_4$	( ) indicates the makeup
low temp	100	20 (37.5)	80		40	65	0.5um/10min			when t > 1um

Note

1. The growth rate won't vary too much for fresh solution within 1 to 2 weeks (for not-very-often usage).
2. Surface cleaness and step 4-6 are the critical steps. Additives and Ph adjustment are needed especially for used solution after weeks.
3. Low temperature processes are for vulnerable devices and weak passivation layer. Normal operation temperature is more predictable.
4. Growth rate are faster in the edges than in the center, but low temp recipe can improve the uniformity.
5. Between each step, DI water rinse 1-2min and spin dry.

Table 4.1: Process flow for plating of aluminum.

the benefit was not considered worth the added risk of dealing with HMDS, a volatile and carcinogenic substance. The photoresist was ACME Resist alpha3 and baked at 120°Ffor 20 minutes. Photoresist is an ideal material as it is compatible with the chemicals and temperatures of the plating process, and removes easily with acetone.

Cleaning

The next steps are cleaning steps. First, a cleaning with a surfactant (Stapleton MICRO 209) for 60 seconds at 60°C, followed by a DI rinse. Second, a quick etch with 35% Nitric Acid (Stapleton MICRO 242) at ambient temperature for 60 seconds removes oxygen from the aluminum surface. These steps prepare the aluminum surface for the plating process that follow.

Aluminum Activation

The third plating step is an acid zincate process (Stapleton MICRO 262), which coats the surface of the aluminum with a thin layer of nickel and zinc. This step provides the seed layer for the subsequent plating steps. There is some physical change that can be observed after this process. The process is performed at ambient temperature for 60 seconds.

Nickel Plating

The fourth step is the nickel plating and I have found it the most critical for a successful plating operation. There are three very important considerations. Because many of the aluminum pads to be plated are connected to pn-junctions, it is critical that this step is performed in darkness. Light will produce a pho-

tocurrent which will impede or reverse the plating operation, and can result not only in the blocking of the nickel plating, but also the etching of the underlying aluminum. Secondly, the solution is heated to 90°C and the part is in the solution for 10 minutes, and significant evaporation will occur. Some form of reflux condenser is required or the solution's pH will increase during the plating, and can attack the aluminum. Finally, this solution must be fresh or metal dissolved in the solution will precipitate out of solution and result in the growth of dendridic nickel. As always, a rinse in DI at room temperature follows.

Gold

The last two steps plate the nickel with gold. The first step is to deposit a near monolayer of gold on the Nickel using an immersion gold process (Stapleton MICRO 290). This process is a displacement reaction, exchanging nickel for gold. The process terminates when the gold replaces all the surface nickel. As before, this process is performed in darkness to prevent interference from photocurrent. The die are left in the solution for five minutes at 82°C. After a DI rinse this gold is plated further using an autocatylitic gold process (Stapleton MICRO 294). This process plates gold on the thin layer of gold from the previous step.

Final Cleaning

After the final plating step the die are DI rinsed in room temperature for 5 minutes, followed by a DI rinse at 90°C for 5 minutes. These cleaning steps are required to remove the many salts from the final gold plating step. A clean surface is critical for permitting bonding to the interface pads of the chip, and

to avoid contamination of the sensors and ultimately the underlying circuitry. An acetone rinse then removes the photoresist and frees the die. Another DI rinse and the die are epoxied into packages and the pads are wirebonded.

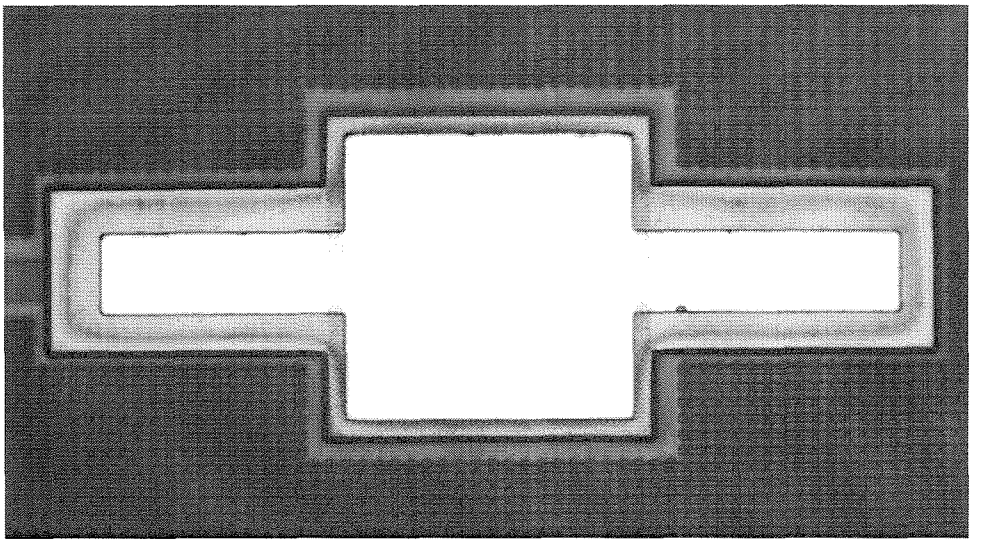
4.4 Plating Results

In this section I present results of the chip plating. I plated the chips according to the procedure outlined in the previous section. The results of plating chips in the context of sensors will be discussed in the next chapter.

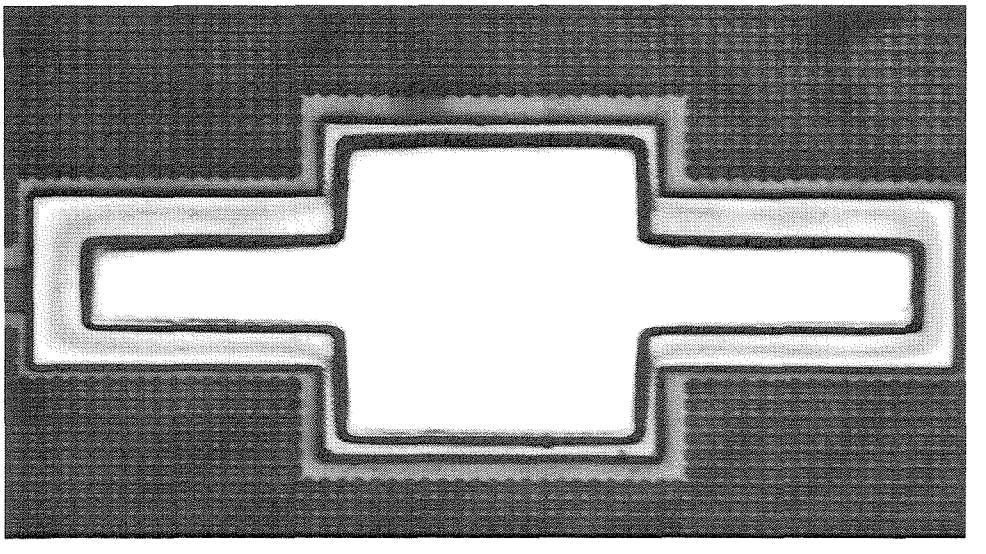
4.4.1 Plating of an Aluminum Pad

Figure 4.4 shows the surface of unplated aluminum pad, and the same pad after completion of the plating procedure. After plating, gold covers the previously exposed aluminum. This pad is the same test structure used shown in Figure 4.2 for the ball bonding example. The plating produced a smooth layer of gold over nickel. Contact to the plated pad with deposition of the sensor polymer was successful.

Profilometry, shown in Figure 4.5 and Figure 4.6, indicates that the surface of the plating is 4.7μmhigher than the underlying metal. The profilometry data in the region surrounding the test pad is from a patterned stack of all metal layers. This is to satisfy the metal density rules of the .35 fabrication process. Spacing of this pattern from the pad results in the deep cavity that surrounds the test pad in the profilometry output. The plated surface is very even and smooth.

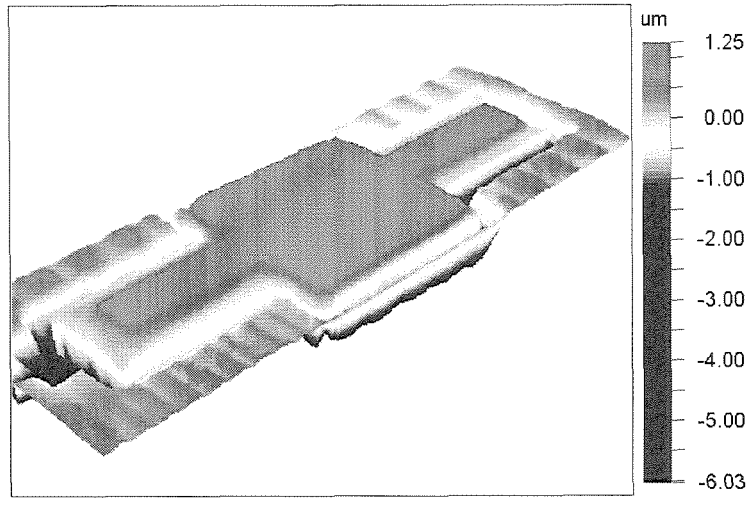


(a) Before Plating

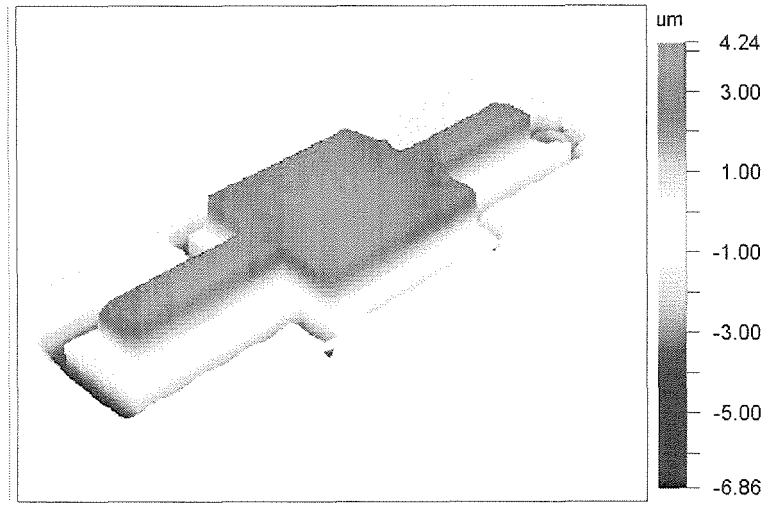


(b) After Plating

Figure 4.4: A before and after shot of the pads.



(a) Before Plating



(b) After Plating

Figure 4.5: Profilometry of the plated aluminum pad.

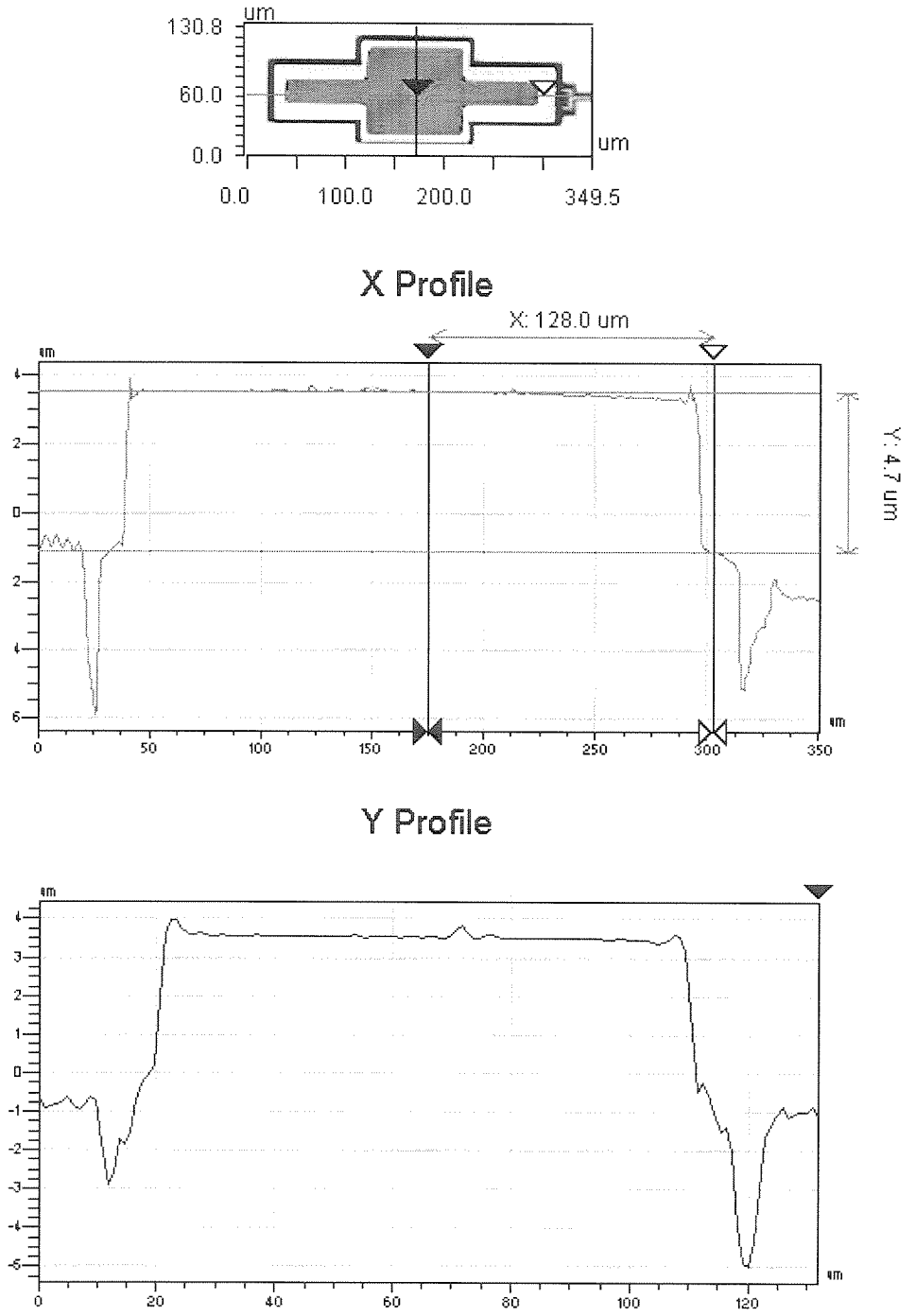


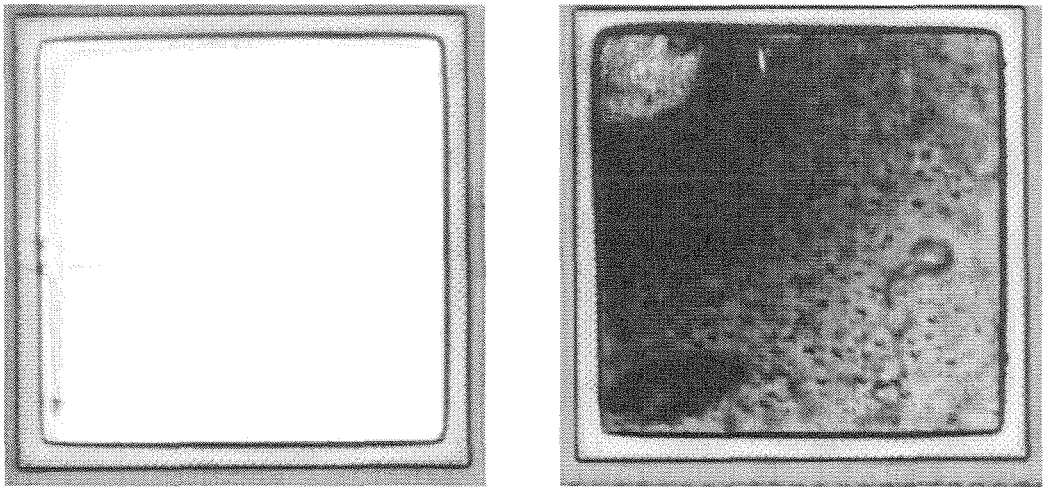
Figure 4.6: Profilometery results on plated test pad.

4.4.2 The Effect of Light

To investigate the impact of light on the plating process I plated a chip without taking any steps to shield the process from light. The fluorescent light in the fume hood was left on at all times. After the plating process I compared two pads, one pad which was connected to the substrate by protection diodes, and a second pad whose metal was floating. Figure 4.7 shows photomicrographs of the two pads. The pad with no substrate connection plated normally, with a nice smooth shiny gold surface. The pad with the connection to the diodes, however, is dark and its surface is pitted. Profilometery, shown in Figure 4.8, shows that there is a difference of 4.4μmacross the surface of the pad. The pads were adjacent to each other on the chip, so localized pH or temperature differences are unlikely to have had an impact. Pads connected to the substrate were consistently harder to plate, even in darkness, for all chips that I plated. A smooth and even surface is important for subsequent successful bonding of the chips when packaged into their chip carriers.

4.4.3 Edge Effects

Another aspect of the plating is that the edges of the aluminum plate faster. This effect can be seen in Figure 4.9, which shows the profilometry of a circle 35 microns in diameter. The edges of the circle have plated faster than the center, causing the formation of a dimple in the center of the circle. However, this effect has little impact on the fabrication of the sensors, and can be mitigated by using the lower temperature process described in Table 4.1.



(a) Pad with no substrate connection

(b) Pad connected to substrate

Figure 4.7: The effect of light on the plating of the pads on a test chip. The pad whose output is floating has a smooth plating of gold, while the pad with a substrate connection has a darker, pitted surface.

4.4.4 When Things Go Wrong

The plating process was certainly not 100% successful. There were many factors the contributed to an unsuccessful plating operation.

The pH and concentrations of the plating solution must be carefully controlled. For the purposes of the prototyping work I did for this thesis, approximately 50 chips were plated. For such a small number compared to even a small scale production, the solution was used for a few chips and then discarded. I performed no filtration or pH control.

The small scale of the plating, while having the benefit of not requiring the sophisticated control of a production process, had its disadvantages. The biggest problem is that the small solutions were difficult to control, the temperature was likely very uneven, and there was no buffering that a larger volume

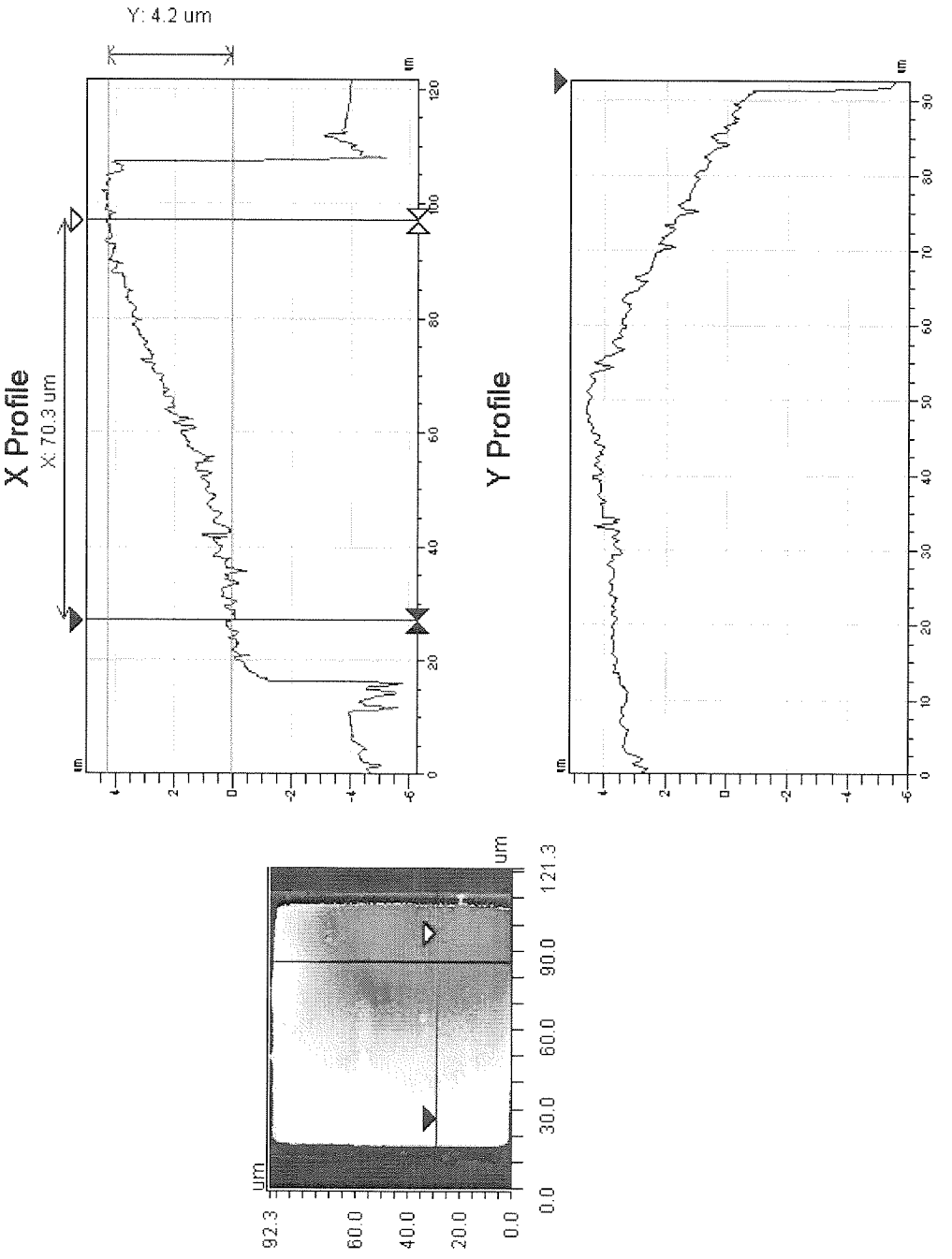


Figure 4.8: Profilometry of pad etched by the light effects. The pad is very rough, and very uneven.

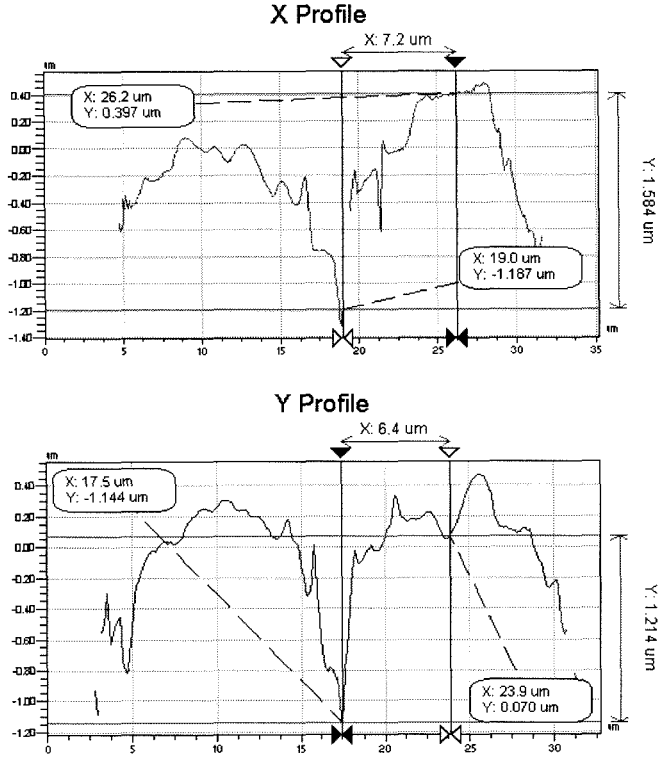


Figure 4.9: Profilometry of a circular aluminum pad showing the faster plating on the edges.

would provide to changes in concentration. I believe a larger plating bath is desirable, even for a small scale prototyping operation. Constant filtering is probably unnecessary, and I performed filtration of the solution only before storage.

One common problem was the growth of nickel in areas of the chip where aluminum was not present. Figure 4.10 shows an example, where small nodules of Nickel have deposited and subsequently been plated with gold. This situation occurred with older solution, where nickel and other metals have contaminated the solution.

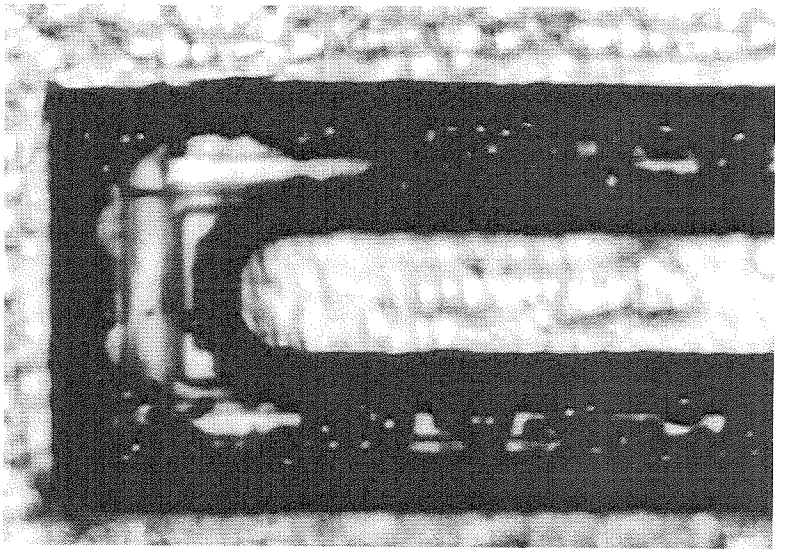


Figure 4.10: Dendritic nickel growth of old solution.

Another problem was the complete etching of the aluminum. Instead of a nice shiny plate of gold, the end of the plating process ended with the aluminum having disappeared from the chip. The exact step of this problem was never determined, due to the lack of proper monitoring equipment.

Surface cleanliness was observed to be a critical factor for successful oper-

ation. If I observed no surface changes after the zincate step, returning to step 1 and repeating usually produced excellent results. A few platings resulting in poor adhesion to the aluminum, and the plated nickel and gold peeled off the underlying aluminum in some areas of the chip.

4.5 Packaged Die

Although in all cases the pads could be bonded, the difficulties in consistently producing easily bondable pads led to the investigation of plating packaged chips. Plating a packaged chip with the bond wires already in place, would not only eliminate the time required to have the plated die packaged, but also alleviate the need to plate thick enough to provide a bondable surface. One potential problem is chemical interaction between the chemicals and the chip package or the epoxy used to secure the die in the package. In addition, the pins and the package cavity are plated with gold, which would certainly consume plating solution. In addition the fragile bonding wires might be physically damaged by the stirring or thermal expansion.

I plated the packaged chips according to the procedure described above; however, I extended the rinse cycles to compensate for any debris from the chip packages, or chemicals leaching into the package. Since the chips were already bonded, the length of the plating processes could have been reduced. The plating was successful, although the packaged chips were much more sensitive to light, and some pads were more problematic to plate. The back of the chip is epoxied to the gold plated chip carrier core with a conductive epoxy. It is possible that a reaction between the plating process and the large conductive area created a more efficient electrochemical cell that prevented the plating

process. However, in complete darkness the plating was successful.

In order to minimize the effect of the chip package on the plating process, I investigated encasing the package in wax. AMCE Wax product AceWax is a wax with a melting point of 90°C and dissolves in acetone. The plating operation proceeded and I observed no detrimental effect of the wax on the plating procedure. However, removal of the wax was not as simple as advertised, even in heated acetone. Another alternative is to use photoresist to cover the gold areas of the package. Ultimately, however, for the small scale and scope of the plating used in this research, I concentrated on plating bare die or packaged chips without preparation.

4.6 Conclusion

This chapter has investigated the postprocessing of individual die to enable the fabrication of chemical sensors on the surface of the chips. Despite come difficulties, electroless plating was very successful. For a small scale operation with limited controls, it exceeded my expectations. More importantly, it solved the contact problem in an efficient and cost effective manner. This was a critical step toward the integration of an electronic nose, which is the topic of the next chapter.

CHAPTER

FIVE

An Integrated Chemical Sensor Array

Arrays of chemical sensors have a number of advantages over discrete or small numbers of sensors. By enabling the creation of a large number of sensors we can approach the sensor count of the biological system. In addition, multiple sensors for the same odorant improves the redundancy of the system. By integrating chemicals sensors we can reduce its size, enabling the implementation of larger number of sensor sites. Moreover, a smaller sensor should be more sensitive, as a single odorant molecule will affect a larger percentage of the sensor area. This chapter describes the development of an array of chemical sensors using the carbon black sensors.

The biggest problem with making arrays of sensors is that as the number of elements increase, it becomes intractable to communicate with them. A two-dimensional array of 32X32 elements would require 1024 wires to measure each sensor, plus one wire for current return that could be shared amongst all sensors. With the current size of the discrete carbon black sensors of .25 cm^2, both the size and wiring become prohibitive. Multiplexing the sensors onto a common bus will eliminate this problem, but the addition of a switch on each sensor will require more area.

A solution to this problem is to reduce the size of the sensors and multiplexing circuitry. An integrated approach for sensor fabrication promises great size reduction, as integrated circuitry has done for discrete implementations of circuitry. In addition to the smaller size, and the practicality of fabrication large arrays, integrating sensors with circuitry has a number of advantages. Fabricating amplifiers in close proximity to the sensors enables addition of gain to the signal with minimal noise interference. Signal processing performed in parallel across the array can minimize extraneous inputs or detect features in the sensor responses, minimizing the data requirements.

Arrays of sensors present different signal processing challenges from those of individual sensors, whose goal is not ambitious. A single sensor fabricated specifically for Carbon Monoxide does not demand much to process its signal. An array of broadly tuned sensors that present a unique multidimensional "fingerprint" for a specific analyte clearly requires much more sophisticated processing. Instead of considering a single sensor output, the system must interpret an entire family of responses. As discussed in Chapter 1, the brain performs this type of processing, trading off sensor specificity for processing complexity.

5.1 Design of Sensor Site

In this section, I discuss the considerations into the design of the individual sensor sites for the integrated sensor array. The requirements of the individual sensors are:

- amenable to array layout
- compatible with standard integrated circuit fabrication
- compatible with deposition techniques.

One layout of a thin film resistor is shown in Figure 5.1. It consists of t-wo contacts of width W, separated by a distance L. The sensor material that comprises the resistor lies between the contacts.

5.1.1 Coffee Ring

The deposition of the sensor material is complicated by non-uniformity of the carbon black distribution within the sensor material. After deposition of the solvent, polymer, and carbon black solution, the carbon black particles redis-tribute as the solvent evaporates. The carbon black particles are pushed to the edges of the drop, leaving a ring of carbon black around the perimeter of the drop.

The "coffee ring" has a serious consequence for the rectangular sensor topol-ogy of Figure 5.1. The higher density of carbon black in the ring of the sensor will form a low resistance path between the contacts. Since this path is a very small portion of the resistor's area, it will not contribute greatly to the area

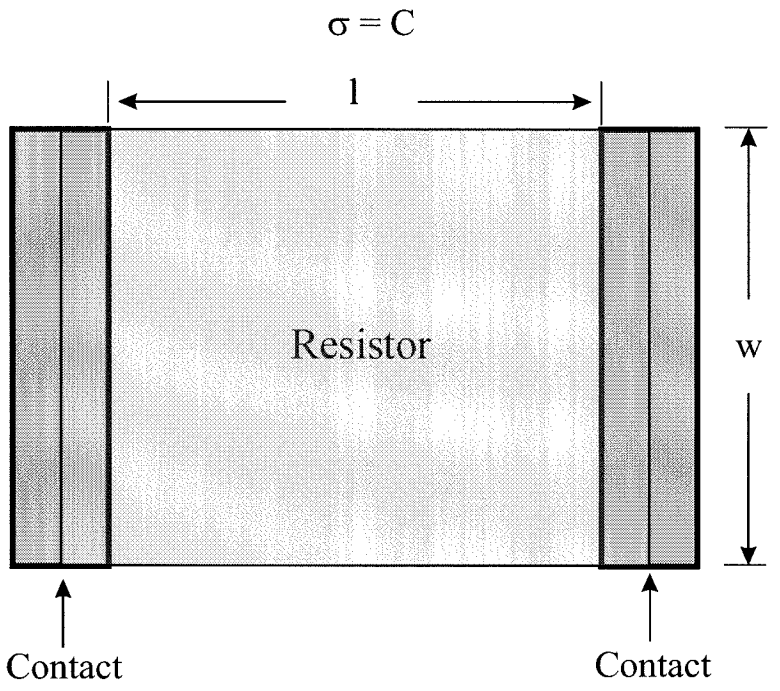


Figure 5.1: Linear sensor layout.

available for sensing. However, the low resistance path will dominate the resistance, reducing the signal from the sensing area.

Another drawback of this resistor layout is that it is not ideally suited for application in densely packed sensor arrays. Figure 5.2 illustrates this situation. While two sensors can share a common terminal, once they are arrayed sharing on a large scale is not possible. Therefore, there must be space left between each of the sensors, and the sensor material must be deposited so as not to contact adjacent sensor sites.

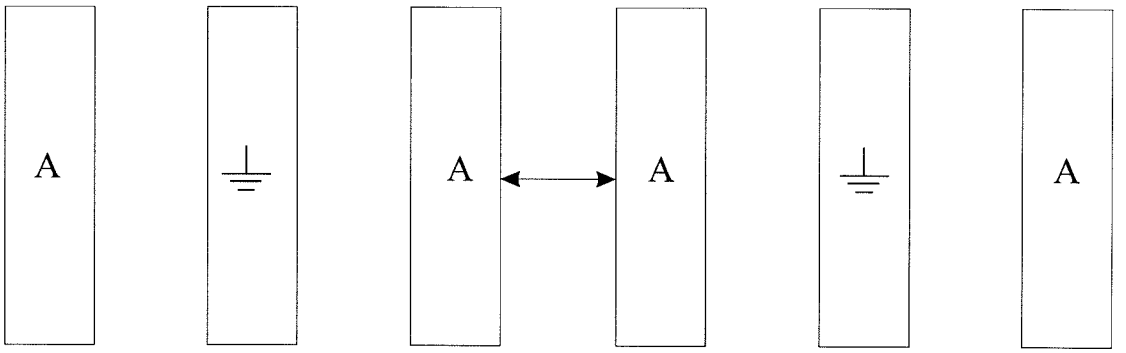
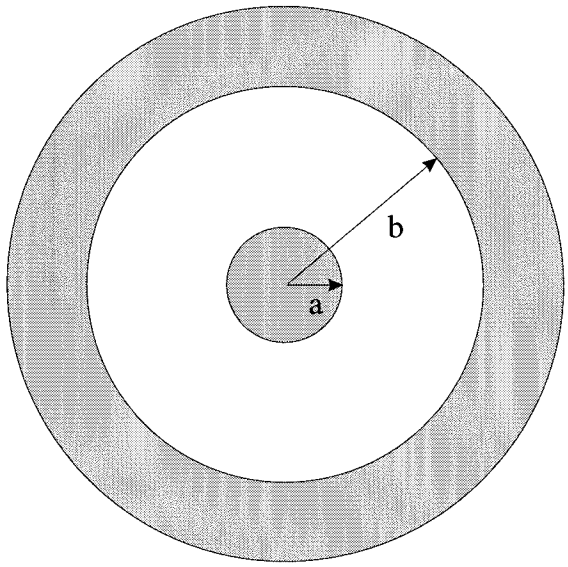


Figure 5.2: The linear sensor array layout. The difficulty is that there can be interference between adjacent nodes (labeled A). The individual sensors must be placed far enough apart for electrical isolation, and this layout is not amenable to a mixture array, where different sensor materials are allowed to mix.

5.1.2 The Ring Sensor

An alternate sensor topology is shown in Figure 5.3. There are two terminals organized in a ring structure. One sensor contact is located at the center of the ring, while the peripheral ring is the other. This sensor has a number of benefits over the layout previously discussed. By making the exterior terminal common to all sensors, this conductor can be shared among all sensors. In

addition, the aggregation of carbon black particles on the exterior of the sensor changes from a problem to a benefit. The higher carbon black density should decrease the contact resistance between the sensor and the conductor.



$$R = \frac{1}{2\pi\sigma x} \ln(a / b)$$

Figure 5.3: The ring sensor topology for integrated arrays.

The current density $J_r(r)$ across a differential with cross section $t(r)rd\theta$ is given by

$$J_r(r) = \sigma(r)E_r$$

where E_r is the radial component of the electrical field E and $\sigma(r)$ is the ohmic conductivity and $t(r)$ is the thickness of the sensor material. In the absence of free charge, Gauss' law requires

$$\nabla.(\varepsilon E) = 0$$

or

$$(1/r)(\delta(rE_r)/\delta r) = 0$$

with solution

$$E_r = dV(r)/dr = c/r$$

where c is determined from:

$$Vab = V(b) - V(a) = \int_a^b E_r dr = c \ln(b/a)$$

or

$$c = V_{DS}/\ln(b/a)$$

Combining these equations allows us to express the current density is expressed as

$$J_r(r) = \sigma(r) V_{DS}/(r \ln(b/a))$$

The differential of resistance dR is

$$dR = (1/I)[dV(r)/dr]dr = [1/(2\pi r t(r) J_r(r))][dV(r)/dr]dr$$

$$dR{=}dV(r)/I{=}dr/[2\pi\sigma(r)t(r)r]$$

Therefore the resistance R of the sensor with a constant ohmic conductivity $\sigma(r) = \sigma$ and constant thickness $t(r) = t$ is given by:

$$R{=}\frac{1}{2\pi\sigma x}\ln(a/b).$$

5.1.3 The Geometry of Noise

Current noise accompanies the flow of current in all materials: semiconductors, metals, thin films, ionic solutions. The choice of the geometry of the resistor affects the noise characteristics. The impact of the geometry of a thin film resistor on noise has received a lot of interest, as laser trimming of precision resistors typically cuts a slit in the resistor to tune the impedance to a desired value.

The lower noise of the rectangular resistor arises from the constant electric field. In our ring transistor the electric field is not constant, but increases closer to the center conductor. Any resistance noise fluctuation in this region will have a greater impact on the resistor voltage. From a current noise viewpoint, therefore, the ring structure is not as desirable as the rectangular sensor.

An alternate topology that maintains the array practicality of the ring sensor, but attempts to minimize the current noise due to the electric field distribution, is shown in Figure 5.8. This rectangular shape has a number of advantages. By incorporating a longer central conductor the electric field distribution in the middle of the sensor is constant, and this will minimize the current noise in that region. The longer central conductor also provides increased contac-

t area for the sensor node, improving reliability. In addition, any noise due to the contacts will be smaller since the area is increased. The design is also flexible, as depending on the deposition method the ends of the sensor can be left free of sensor material, avoiding any of the higher noise regions due to the nonlinear electric field.

5.1.4 Wells

The plastic and carbon black composing the chemical sensors are suspended in a solvent. After deposition, the solvent evaporates leaving the sensing plastic and the carbon black. While the final sensors are very small and thin, the volume of solvent can be many orders of magnitude larger. This causes some problems for the deposition, since small volumes of liquid are difficult to deposit. I felt it would be necessary to form a constraining well to contain the solvent while it evaporated. In this section I will briefly describe approaches to well formation in CMOS. Ultimately, however, it was impractical to form wells constrained to individual die from an outside foundry.

5.2 Deposition Techniques

The deposition of the polymers presents a significant challenge. Fabrication of small sensors requires a method of dispensing small amounts of the solvent, polymer and carbon black combination. While the development of this capability was beyond the scope of this thesis, in this section I will briefly review possible techniques, and discuss the methods used in this thesis.

An airbrush is a pneumatic device that is used by artists to spray a fine

atomized mist of paint. A physical mask, or frisket, is used to protect selective areas from the deposition of paint. Substituting the sensor material for paint, and fabricating physical masks for deposition, is one possible approach for sensor deposition. Figure 5.4 is a composite picture of two masks I fabricated for this purpose. The nickel mask (Photostencil Incorporated, Boulder, CO) is electroformed: nickel is electroplated from an initial seed. This allows small features to be fabricated. Due to its low surface energy, Nickel makes an excellent material for a mask, as the sensor material is not likely to stick to it. The other mask is a laser cut polyamide material (KJ Marketing, Toronto, ON, Canada). Polyamide is an interesting material for a mask as it is flexible, and can make a better gasket with the surface of the chip.

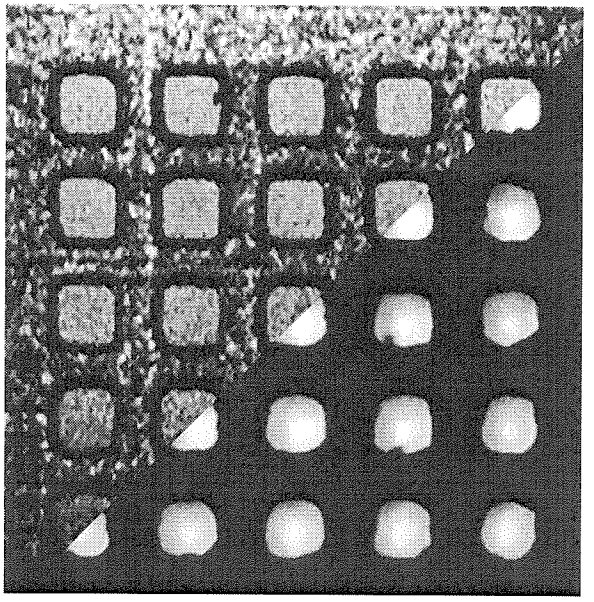


Figure 5.4: Stencils for sensor fabrication.

Because of the small droplets of sensor material that are produced by the airbrush stream, the sensor material dries quickly. This has two significant

benefits. First, wetting of the material to the surface glass of the chip is not an issue, because there is no time for the material to run. Secondly, the uneven distribution of the carbon black particles that led to the coffee ring doesn't occur.

There are some potential drawbacks with the airbrush technique. The greatest of these is the possibility of damaging the underlying circuitry. As the polymer particles are ejected from the airbrush and are propelled towards the surface of the chip, the physical interactions are likely to cause the build up of static charge. This charge could destroy the sensitive transistors integrated beneath the sensors. Fortunately, the size of the sensor sites allow for inclusion of protection circuitry, which should mitigate the impact.

5.3 Integrated Chemical Sensor Arrays

The goal of this research was to investigate the feasibility of fabricating chemical sensors on standard integrated circuits. Therefore, I decided to keep the unit cell simple, so that this phase of the project can be evaluated without the complication of gain or adaptive circuitry. Future work can address these issues.

Figure 5.5 shows the architecture of the integrated sensor chips. The main feature is a core array of unit cells consisting of an individual sensor and the circuitry to connect its output to the column bus. X and Y selection units situated along side the array address and select individual sensors. There is also some minimal circuitry associated with the column multiplexing. In the following sections I will discuss the individual circuit blocks.

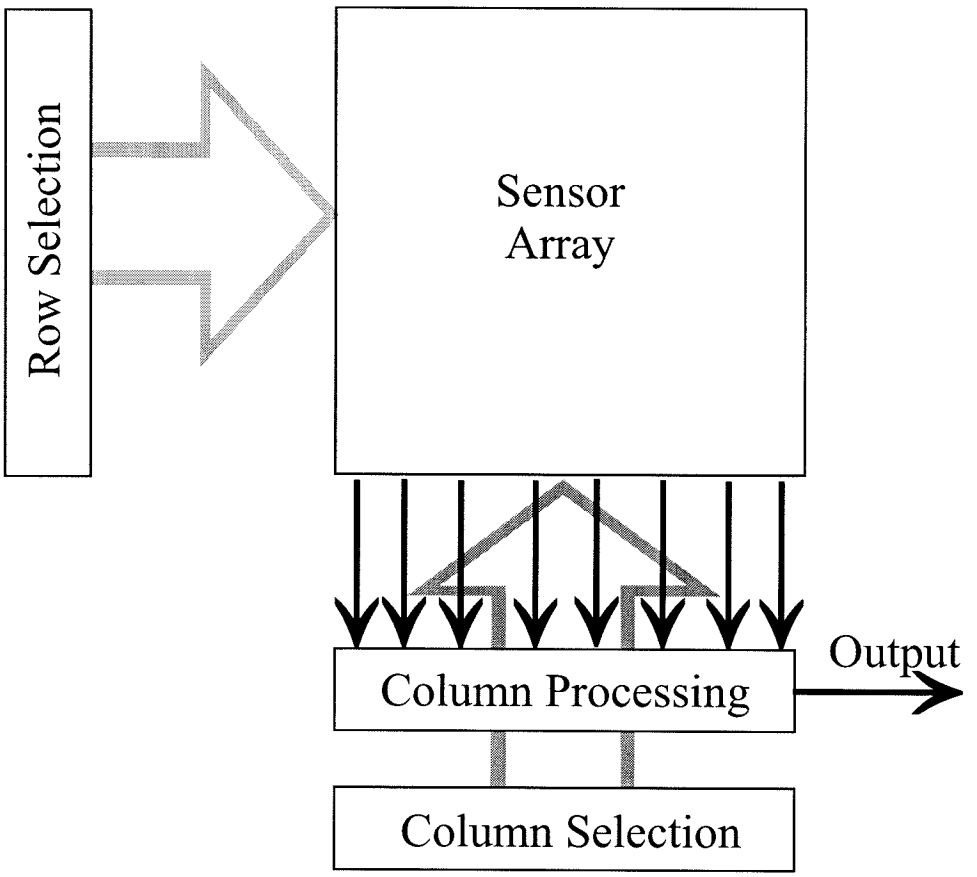


Figure 5.5: Block diagram of the sensor array chip.

5.3.1 Circuit Cell

The function of each unit cell is to connect the sensor node to the excitation signal generated off-chip. I fabricated arrays of two variety of unit cell, a two-wire sensing scheme that requires the sensor measurement via the excitation line, and a three-wire version that multiplexes the sensor voltage to a separate bus.

Two Wire Scheme

Figure 5.6 shows the schematic for the two wire sensing cell. The cell receives three inputs: X and Y selection signal and the excitation signal. The X and Y selection signals are active low. Normally the signals are held at the high voltage supply and the p-fet transistors are off. When both the X and Y selection signals are low, then the two p-fets turn on and connect the sensor resistance to the excitation signal. An alternative scheme would have the excitation signal selected on a ROW basis, and to have a single transistor switched on a column basis. However, the fabrication processes employed for the majority of the chips I fabricated possessed only two levels of metal. With the sensor contacts fabricated in the top layer of metal, completely surrounding the sensors, it was difficult to use the only remaining metal layer for connections in both the X and Y directions. I ran the Y select signals in polysilicon, which is not generally a desirable material for signal transmission due to its inherent resistance and capacitance. However, since the Y select signal had to source no DC current and switched at a much slower rate than the X select line, this was a reasonable tradeoff.

To select an individual sensor the both gates have to be pulled down low.

Since one gate is a column select line and the other is a row select line, this enables a individual sensor to be selected. When selected the two transistor have the same gate voltage, and can be considered one long transistor with an effective L (which will be approximately the sum of the two transistor lengths). For voltages of V_{ds} <V_{gs}-V_t the compound transistor will act like a resistor whose resistance is given by

$$R_{on} = \frac{L}{W\mu_n C_{ox}(V_{GS} - V_T)}$$

where:

L = length of the transistor

W = width of the transistor

μ_p = surface mobility in the channel

C_{ox} = capacitance per unit area of the gate oxide

V_{GS} = Gate source Voltage

V_T = Threshold Voltage of pmos transistor

This equation is only valid for small V_{ds}. This circuit gives us the ability to select each individual sensor via the two selection lines. A drawback of this circuit is the inability to measure the voltage across the sensor directly. For voltages V_{ds} >V_{gs}-V_t the transistors will saturate, and their current will most be a function of the gate voltage. In this range of operation the change in sensor resistance will not be detectable, since the switch transistors will not be affected. Therefore this circuit will only operate properly when the switch transistors are in their active region, V_{ds} <V_{gs}-V_t.

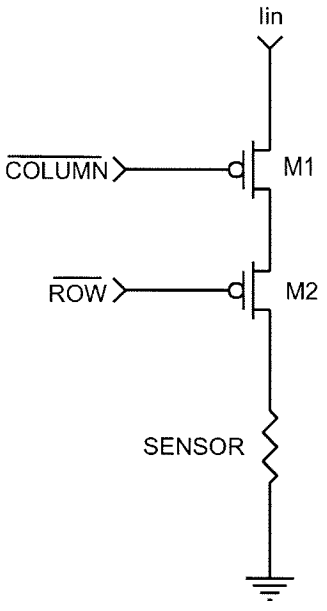


Figure 5.6: Schematic of the two wire sensing unit cell.

Three Wire Scheme

The three wire scheme adds a separate voltage output selection to the two-wire line. Figure 5.7 shows the schematic of this version of the unit cell. Logic at each cell decode the select lines, which are now active high. This decoded signal then drives a single transistor connected to the sensor, in addition to a transmission gate that selects the voltage output. Using two transistors in series to select the sensor, as in the two wire scheme, is not desirable as two transistors do not use efficient space. Moreover, since to minimize noise the switch transistor should be as large as possible, the gate capacitance that the X and Y selection lines have to drive is much larger. With this scheme the selection lines only have to drive the smaller capacitance of the decoding logic.

A simple two transistor transmission gate is included to connect the output voltage of the sensor to a column bus. The NAND gate selection signal and

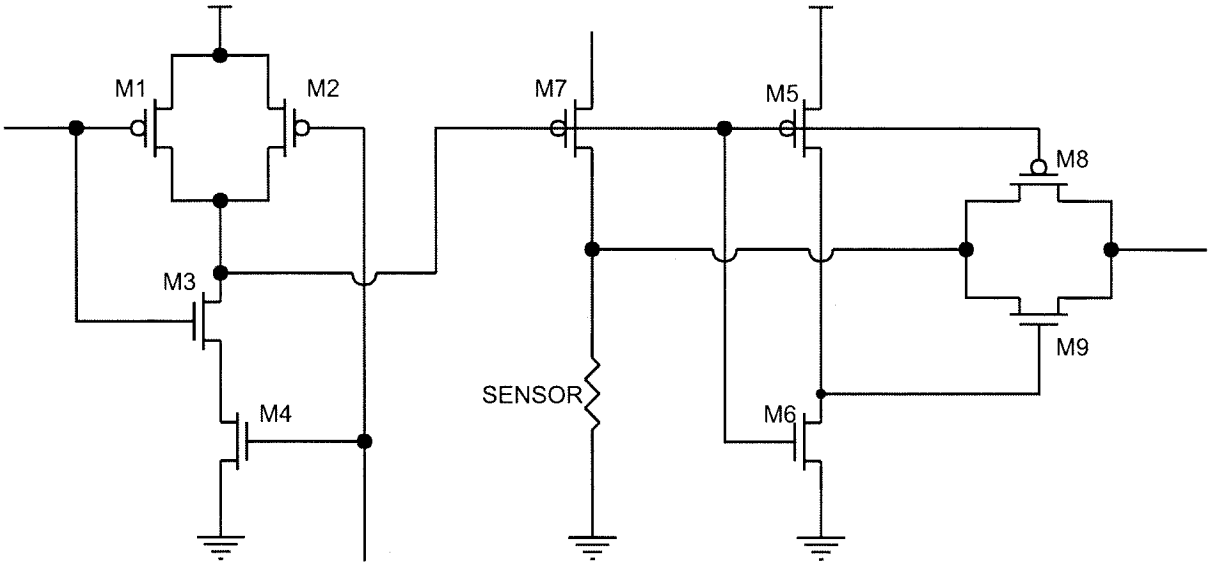


Figure 5.7: Schematic of the three wire sensing unit cell.

an inverter to provide the complementary signal activate the switch. The $1/f$ noise of the transistor is of the greatest concern for our system. It is given by [ref]

$$\overline{i_D^2} = \left[\frac{8\mathrm{kT}\mu_n \mathrm{C_{ox}} \frac{\mathrm{W}}{\mathrm{L}} V_{DS}(1 + \frac{\gamma}{\sqrt{2(2|\phi_F| + V_{SB})}})}{3} + \frac{(KF)I_D}{fC_{OZ}L^2} \right] \Delta f$$

where:

k = Boltzman's constant $.38 \times 10^{-23}\mathrm{J/^\circ K}$

T = Temperature (K)

V_{DS} = Drain source voltage (V)

V_{SB} = Source bulk voltage (0V)

γ = Bulk threshold parameter $.2\mathrm{V}^{1/2}$)

φ_F = surface inversion potential

KF = flicker noise coefficient $10^{-28}\mathrm{F} \cdot \mathrm{A}$

f= frequency (Hz)

Δf = bandwidth (Hz)

Since there is no DC current through the switch, it is not as important to make them large for noise considerations. Each column V_{out} line connects to a simple voltage amplifier, which is also selected by the column line. Each amplifier will have an associated offset, this will introduce an offset on each column.

Figure 5.8 shows the layout of this cell. The perimeter is the top layer of metal and is the ground contact for the sensor. The middle of the cell is the drive contact for the sensor. Along either side of the contact is the switch transistor. The decoding logic is at the top of the cell, and the voltage switch and inverter are located at the bottom.

5.3.2 Column Amplification

Since the three wire cell has a voltage output, the addition of an amplifier will speed up the readout of the array voltages. For the initial chips I employed a very simple amplifier, shown in Figure 5.9. The amplifier is configured as a simple voltage follower. Each column of the array has its own amplifier.

The column amplifier is selected by a clock line generated by the shift register cells. When the column is not selected, the bias current of the amplifier is turned off, and the amplifier has no drive. This allows all the outputs of the column amplifiers to be tied together; only the selected amplifier will be able to drive the output line.

This amplifier has many drawbacks in exchange for its simplicity. Its open

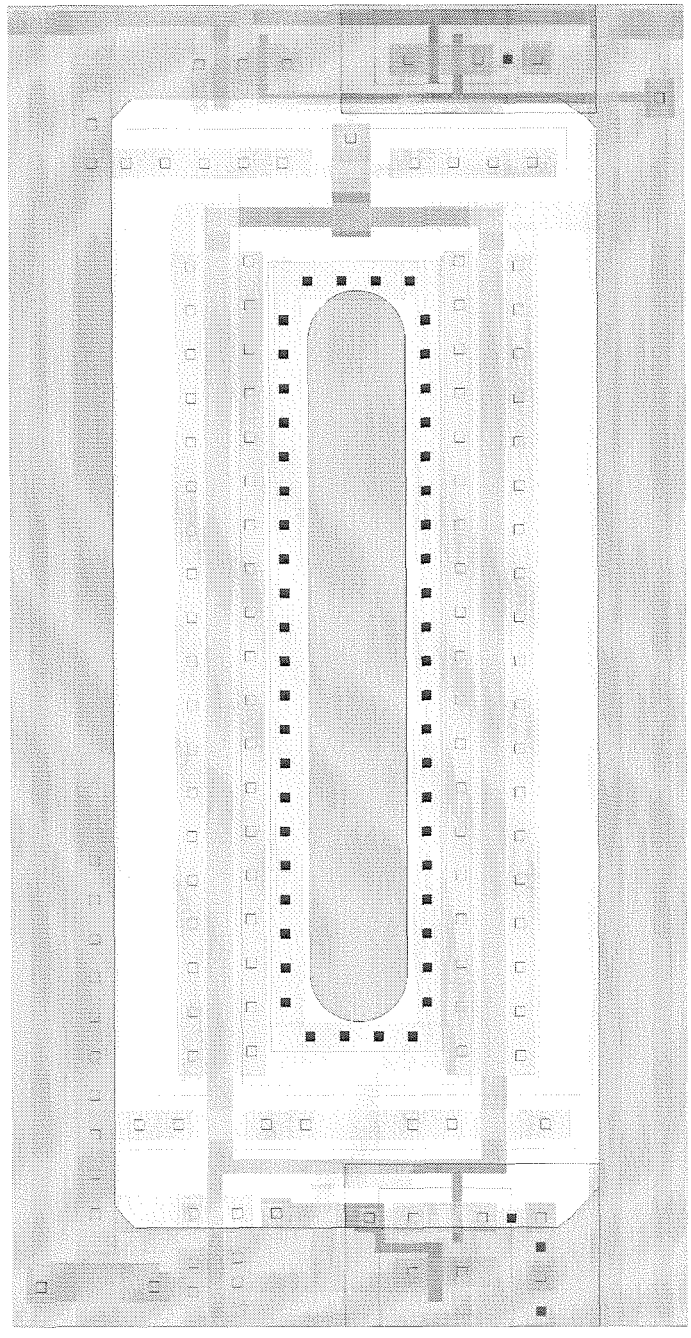


Figure 5.8: Layout of the three wire unit cell.

loop gain is not large, so the accuracy of its unity transfer function is low. Moreover, it does not have rail-to-rail outputs. The output has to remain high enough to keep the n transistor on the output side in saturation. For this application the amplifier was good enough; however a better amplifier is clearly desirable. The selection was made for its simplicity and its compact size, allowing it to be laid out within the pitch of the columns of the array in a 2.0 micron process. Smaller feature size processes will easily permit the fabrication of more sophisticated amplifiers. A switched capacitor amplifier would be an ideal candidate for future implementations. At the time of this writing such an array is currently in fabrication.

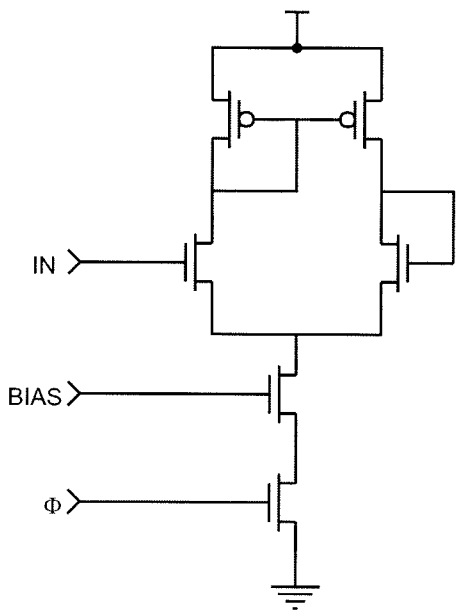


Figure 5.9: The Column Amplifier.

5.3.3 Interface Amplifier

In addition to the individual column amplifiers, the three-wire chip has a larger buffer amplifier for driving off-chip loads. While it too will introduce an offset voltage, since all the column outputs pass through this amplifier, the output will be constant for all columns. The amplifier does have the benefit of rail-to-rail outputs, however its transconductance is not constant over its operating range, and will have gain errors as a result. There is still the problem of increased noise and temperature drift, so a more sophisticated amplifier is recommended for future work.

5.3.4 Digital Design

The digital portion of the device controls the addressing of individual sensing elements in the array. Speed and reliability were the main considerations in selecting an addressing scheme.

One approach for generating the selection bits is a counter, and decoding the binary value at each row or column. One benefit of this approach is that the decoding requires very little area, and would be a suitable choice where the pitch of the cells does not allow much room for circuitry. However, the smallest of the sensors I designed is 90 microns, leaving plenty of room for circuitry. A drawback of the counter is that expansion of the array size would require changes to all decoders and the counter. In addition, one must implement large decoders in multiple levels of logic for speed, as long series chains of transistors can suffer from slow performance due to the body effect. While a decoder per row or column would allow for random access to the array, I did not consider that a necessary feature for the initial attempts at an integrated array system.

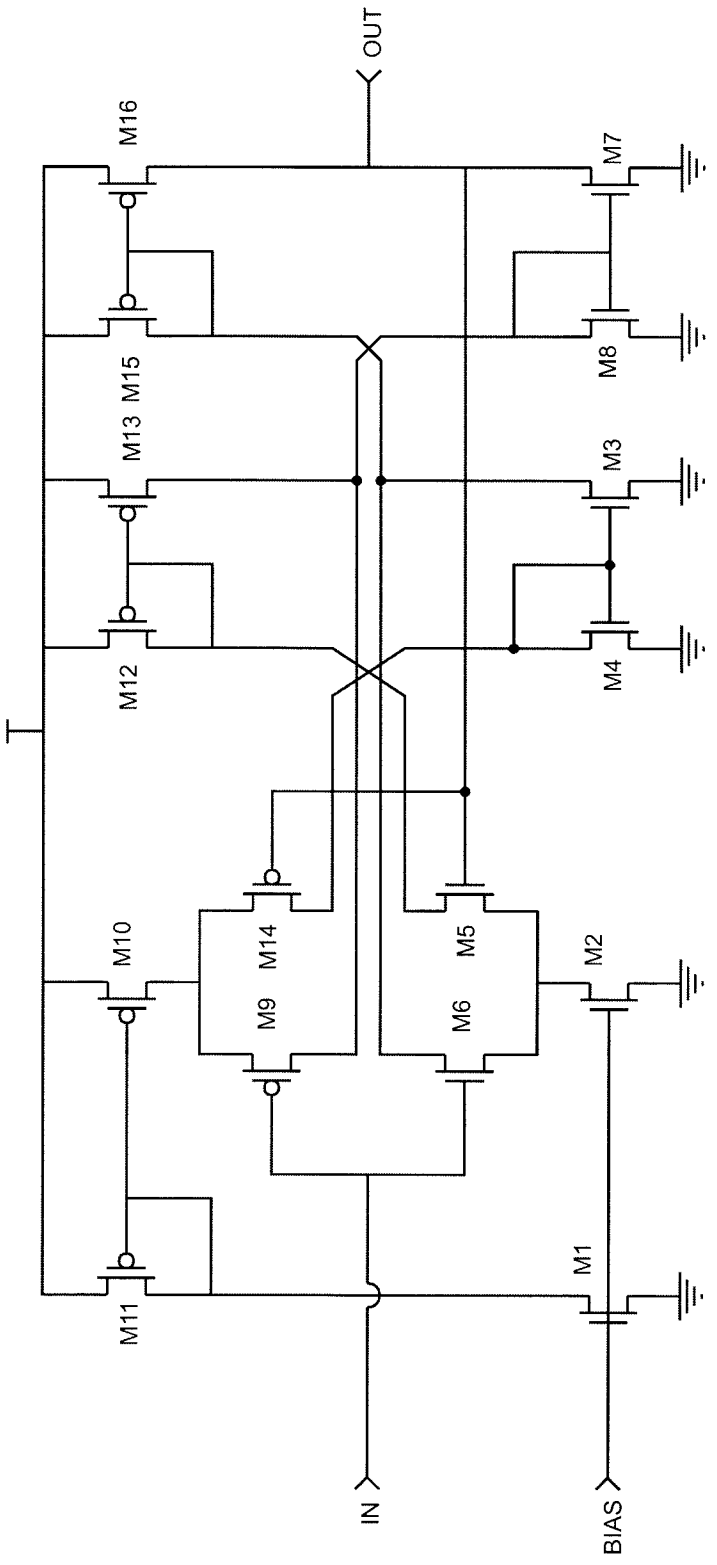


Figure 5.10: The widerange amplifier for driving the signals off-chip.

Since scanning through the array in a sequential manner is satisfactory for this design, the X and Y decoders in Figure 5.5 are simple shift registers. The selection bit is loaded into the first stage of the register and is clocked along to select each consecutive row or column. Although I used a conventional master/slave D flip flop in one version of the array implemented in a .8 micron technology, for the 1.2 micron and 2.0 micron technologies I used the scanner frame developed for silicon retinas [5]. The scanners have the advantage of an elegant and compact design, and most importantly have many years of successful operation in the technologies I used.

Figure 5.11 shows one stage of the scanner shift register. It has a typical pair of latches in the master/slave configuration, but elegant design simplifies the amount of logic in each cell. Moreover, since it is designed as a shift register, the scanners employ a dual rail signaling scheme to eliminate the need to compute inverted logic signals at each shift register cell.

In order to improve the speed of the scanning function, the selection lines of the scanner are run through buffer amplifiers consisting of dual inverters.

5.3.5 Noise Reduction

In order to reduce the amount of noise introduced into the system, there are a number of design techniques to apply. The most susceptible points for the introduction of noise are the bonding wire and package lead, as they can operate as antennae. In order to minimize this pickup I placed the analog lines on the opposite side of the chip from the digital lines. In addition, two analog ground pads placed on either side of the sensitive analog output minimized inductive coupling.

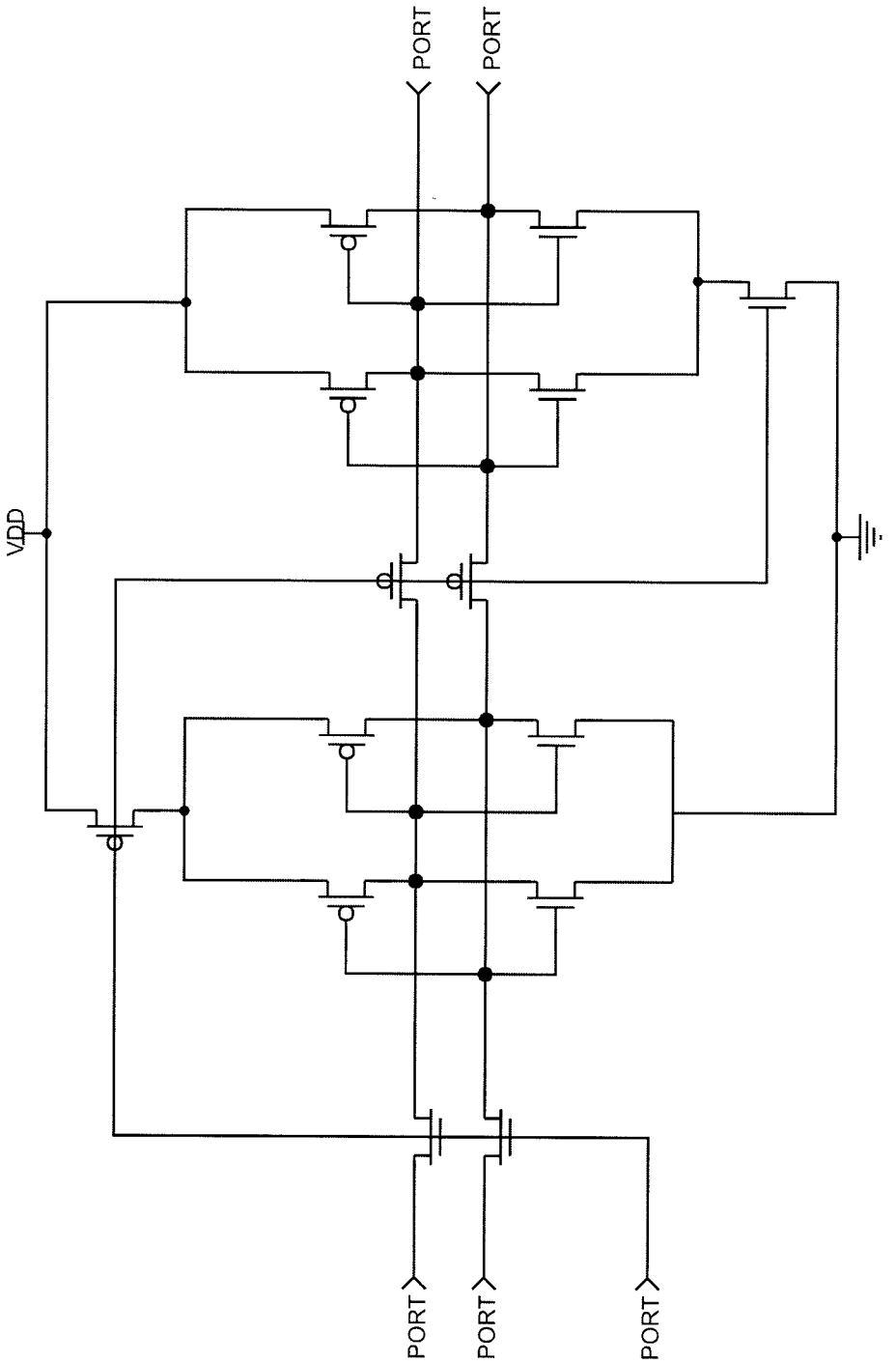


Figure 5.11: Schematic of the shift register cell.

In order to isolate the analog sensing core and the digital shift registers I placed substrate contacts around the periphery of the array core. In addition, I placed many well and substrate contacts in the cells themselves. These contacts help capture any stray carriers injected into the substrate.

5.4 Postprocessing

I fabricated a number of array chips with different ring topologies. In this section I present the results of the postprocessing of the sensor chips.

Figure 5.12 shows a photomicrograph an individual plated two-wire sensor, fabricated in the AMI 1.2 micron process. The sensor has a perimeter contact and a central circular contact. The perimeter contact, as discussed in the section describing the ring sensor, is ground. The central contact is the selection contact; the two transistors that make up the switch are located under the contact.

The results of profilometry on the sensor is shown in Figure 5.13. The plating has is approximately 6.5 microns high. Figure 5.14 shows a three dimensional view of this profilometry data.

Measurements of the plating height are shown in Figure 5.15. The x-axis profile cuts through the the space between the central conductor and the outer ground ring. The y-axis profile cuts through the center of the central conductor. These graphs show the plating to be approximately 6.5 microns high. The step at y-axis position -2 microns is the level of the metal before plating. The data at position -4.8 microns reflects the underlying silicon substrate.

As the plating is only 25 microns wide, this reflects an aspect ratio of .26. The height of the plated line is determined mainly by the nickel plating. It is possible to plate the nickel higher, although there is some possibility of stress

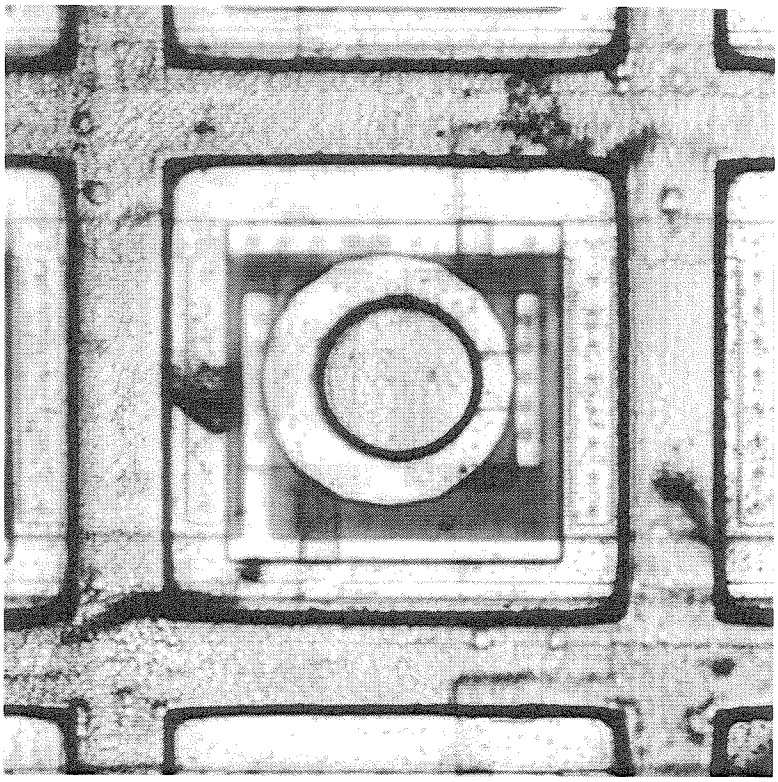


Figure 5.12: Closeup of plated sensor.

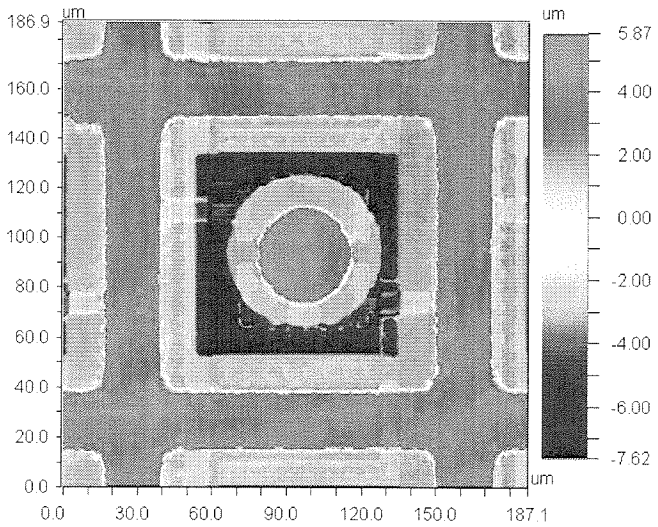


Figure 5.13: Profilometry of single cell.

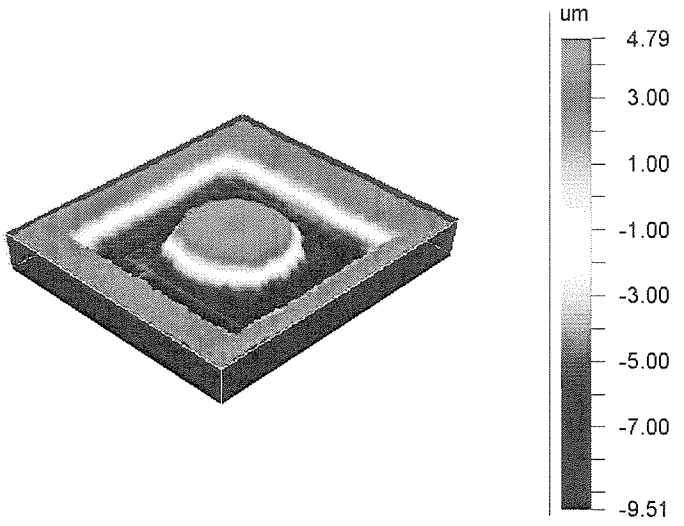


Figure 5.14: Three dimensional projection of the sensor cell.

in the nickel causing adhesion problems. The highest nickel I plated was approximately 12 microns.

A 3X3 section of the plated array and its profilometry data is shown in Figure 5.16. Figure 5.17 shows the X and Y profiles taking through the array, showing the uniformity in the sensor height.

A section of the array is shown in Figure 5.18. The circuitry along the bottom of the figure are part of the row multiplexers.

5.5 Sensor Formation

With the sensor arrays successfully fabricated and plated, the next step is the deposition of the polymer material. For these initial tests a very simple deposition apparatus was employed: a paintbrush. Careful application of the sensor material resulted in deposition in as little as three sensors. However, the

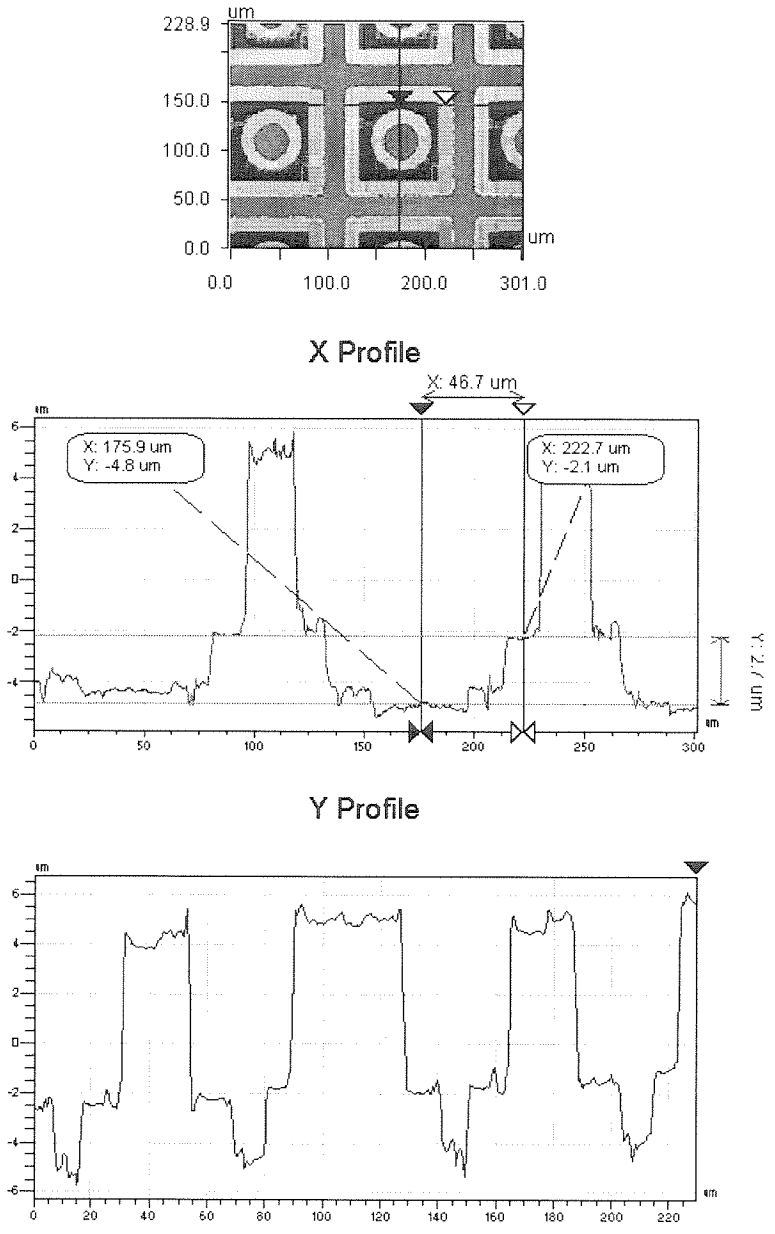
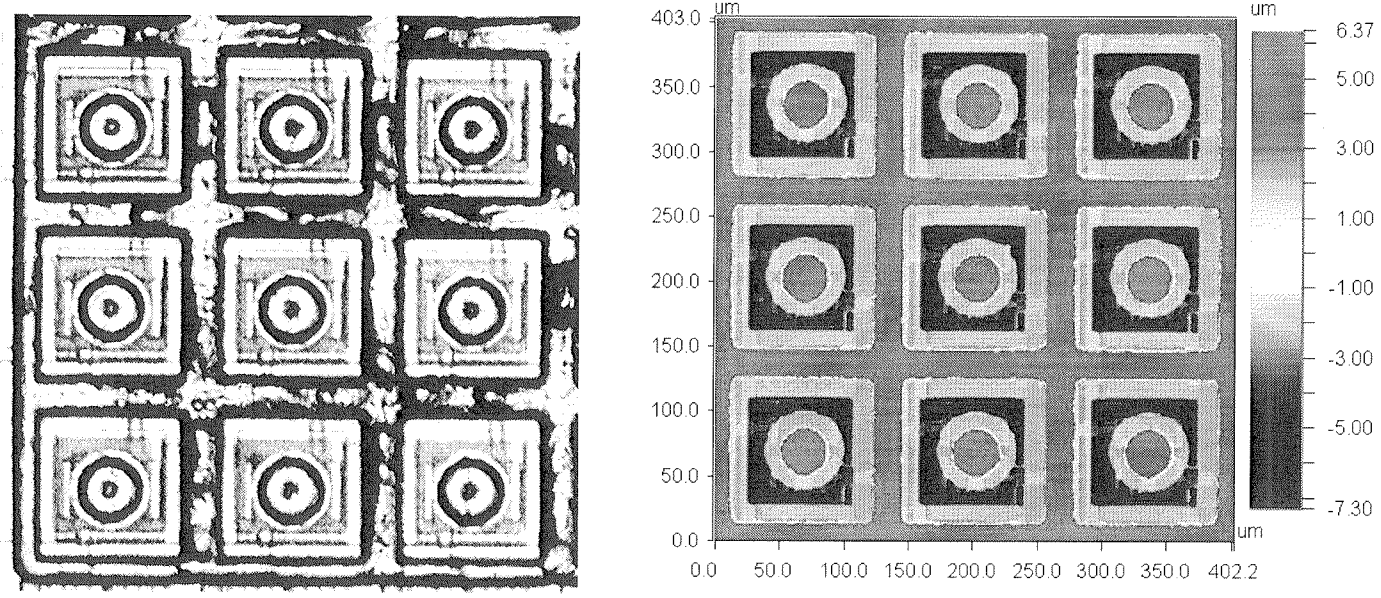


Figure 5.15: Profilometry data on sensor.



(a) a

(b) b

Figure 5.16: 3X3 section of array and profilometry.

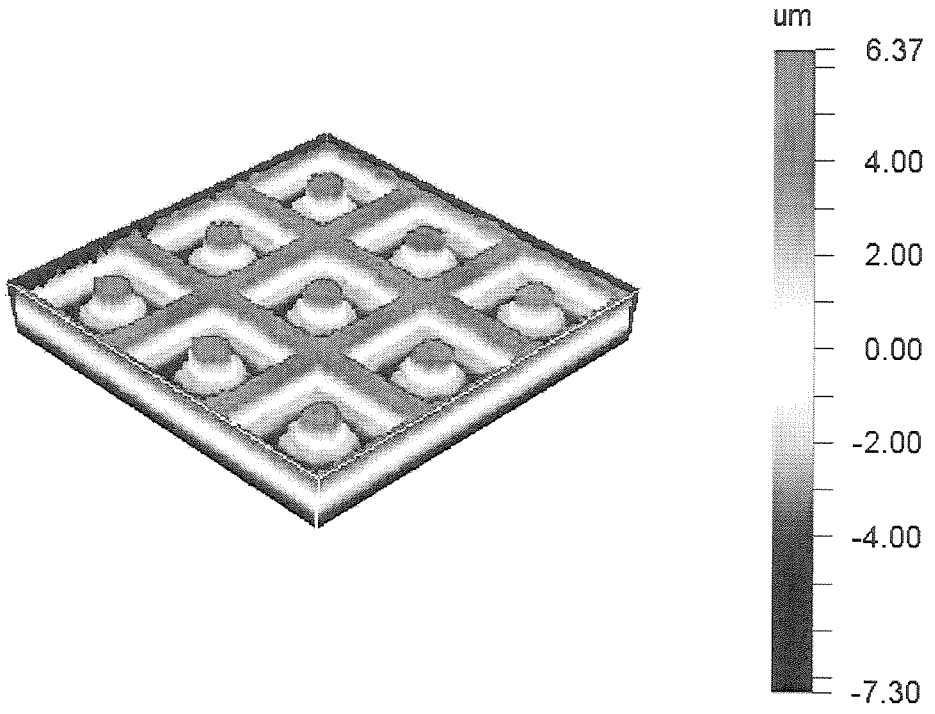


Figure 5.17: X and Y profilometry of sensors.

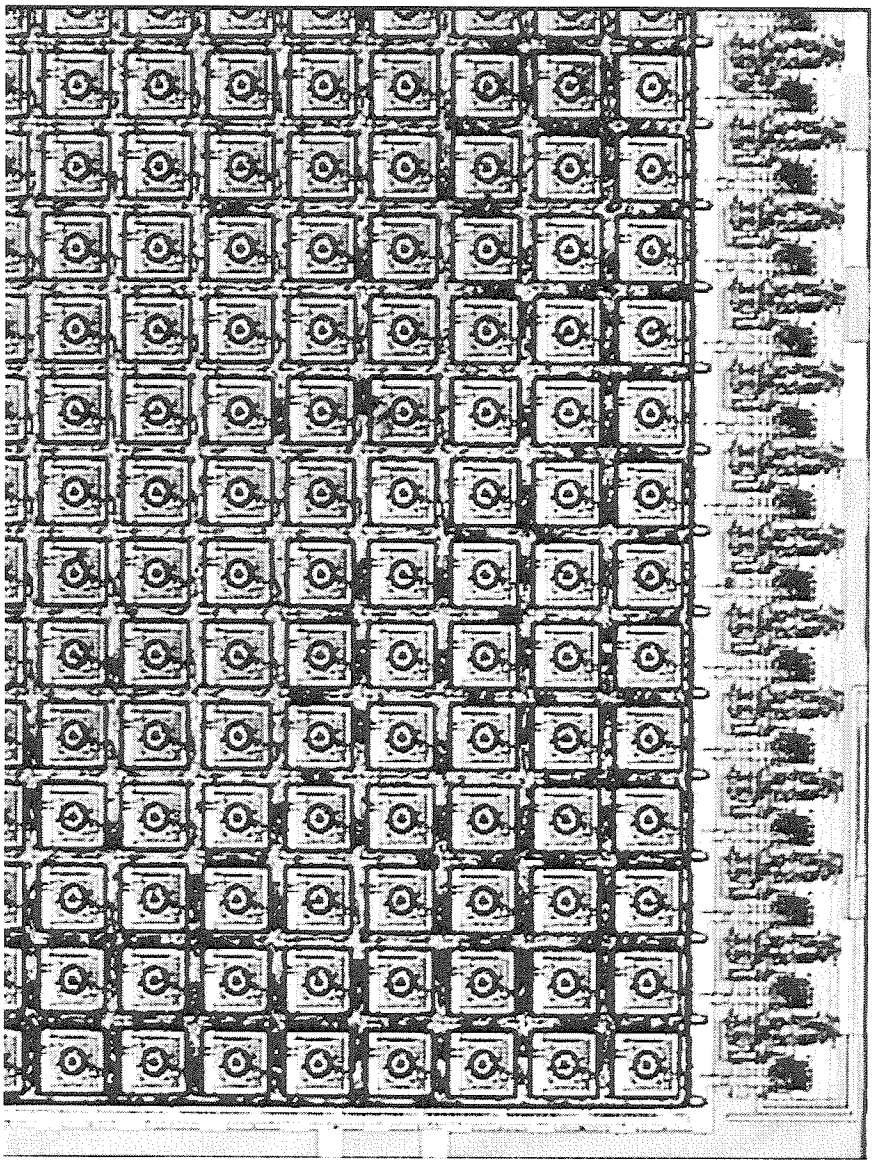


Figure 5.18: Picture of fabricated array.

method is not reliable and is prone to mistakenly depositing over large sections of the chip. For confirmation of the sensor operation, however, it is adequate.

The carbon black used in the sensor fabrication was Black Pearls 2000, a furnace black from Cabot Co. (Billerica, MA). To prepare the polymer-carbon black solution I added 20 mg of carbon black and 80 mg of the desired polymer, into 10 mL of the solvent. I sonicated the mixture for five minutes before using, to suspend the carbon black particles.

Figure 5.19 shows a section of 3X3 sensors that have had a mixture of PVVA and carbon black deposited on them. The deposition resulted in a uniform appearing sensor. The roughness of the sensors shown in Figure 5.19 are the result of a subsequent air brush experiment. The "coffee ring" effect is clearly evident around the edges of the gold plating. As discussed earlier, in this instance it is actually a benefit, as the higher density of carbon black likely improves the contact to the gold.

5.6 Testing Results

In order to test the sensors with a controlled concentration, a simple flow system was employed. The apparatus provided known partial pressures of the three test chemicals: acetone, hexane, and ethanol. A glass bubbler provided a means of creating a vapor with a known partial pressure. The carrier gas, nitrogen, was introduced into the bottom of the bubbler via a ceramic frit. The resulting vapor was diluted by combining it with a controlled flow of the nitrogen carrier gas. The carrier gas was obtained from a pressurized supply tank. The flow rates of the bubbler and background supply were monitored using flow meters from Gilmont Instruments, and controlled using manual valves.

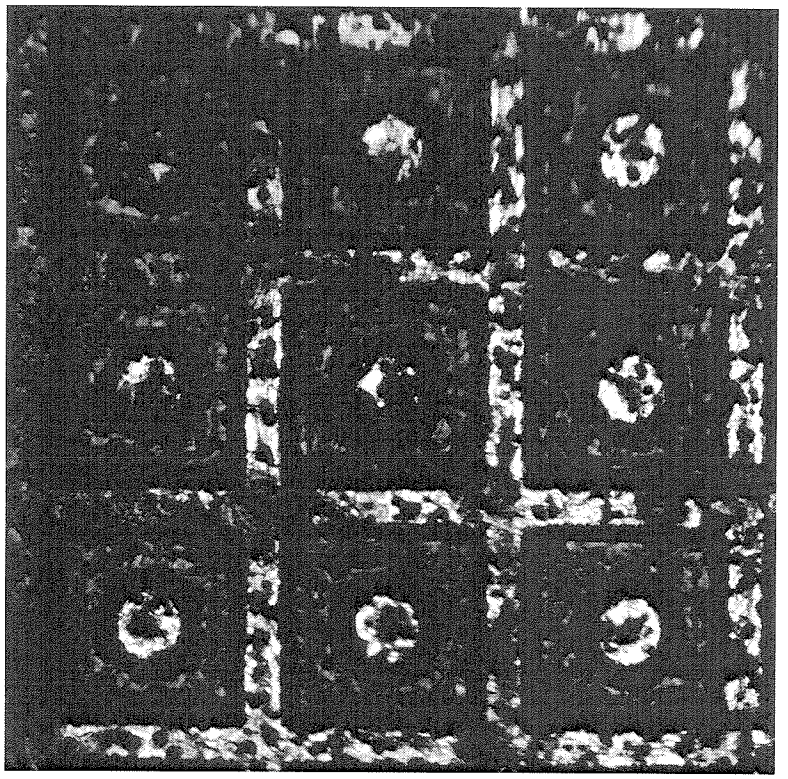


Figure 5.19: Deposited sensor material.

Time constraints prevented the reporting of the results of this detailed tests in this version of the document, although the results should be ready by defense and certainly by June 9th. Limited testing consisting of depositing a large drop of polymer on a array chip, shown in Figure 5.20, produced a map of conductivity shown in Figure 5.21

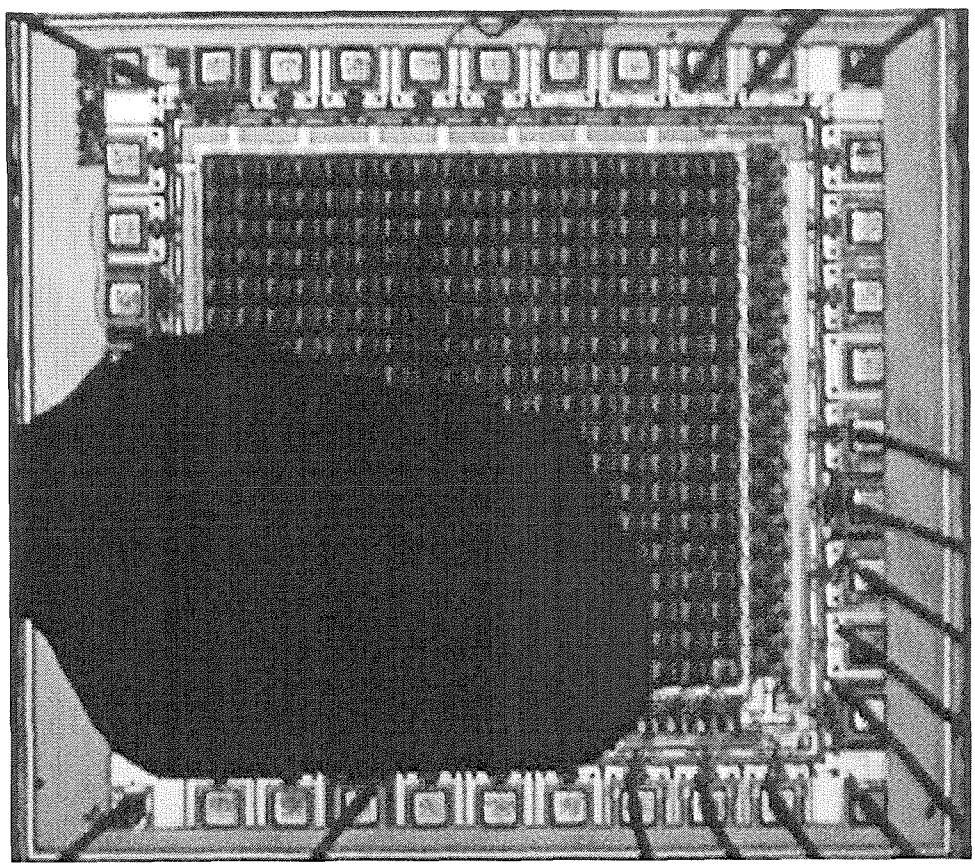


Figure 5.20: Preliminary deposition test. The sensor material is deposited on top of the chip after the gold plating.

Qualitative observation of the sensor response should a change in response to stimulus. The scanning operation of the multiplexers and the circuitry cell were confirmed. The chips successfully survived the plating operation, and

electrical contact was made with the sensor material.

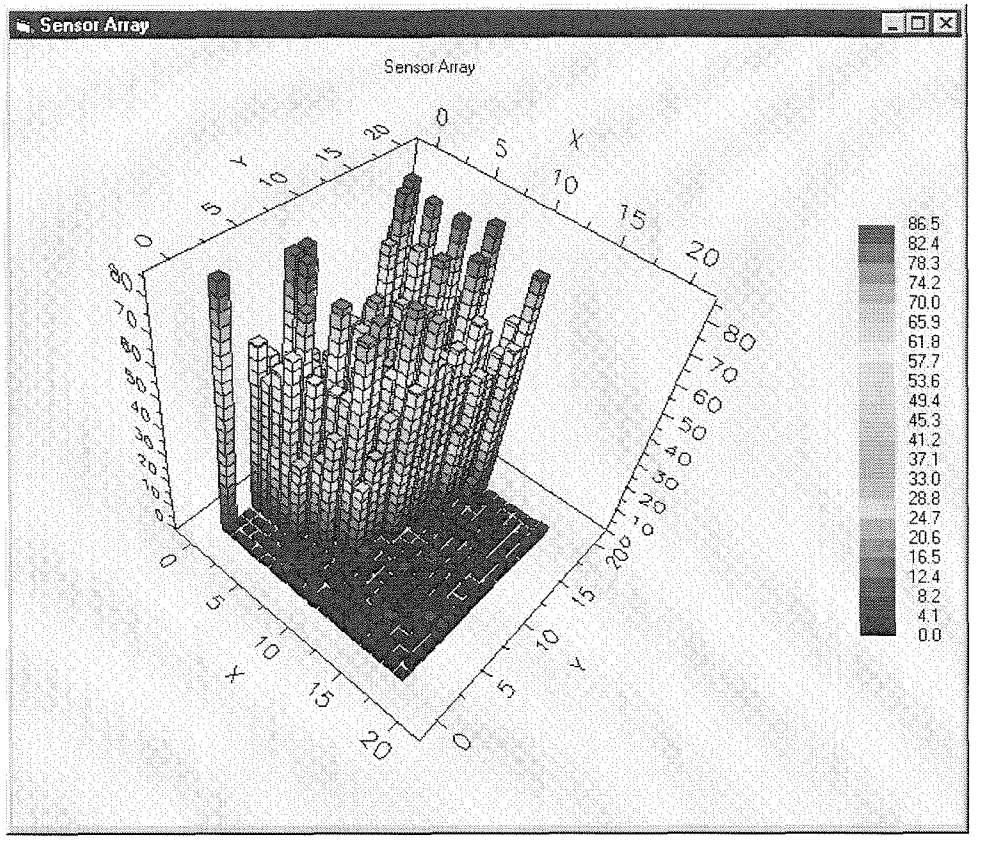


Figure 5.21: Array data showing conductive map.

The current vs. voltage curve for an individual sensor nose is shown in Figure 5.22. The data show that the output is essentially linear. The deviation from linear is due to the nonlinearity of the output buffer, and not the sensor itself.

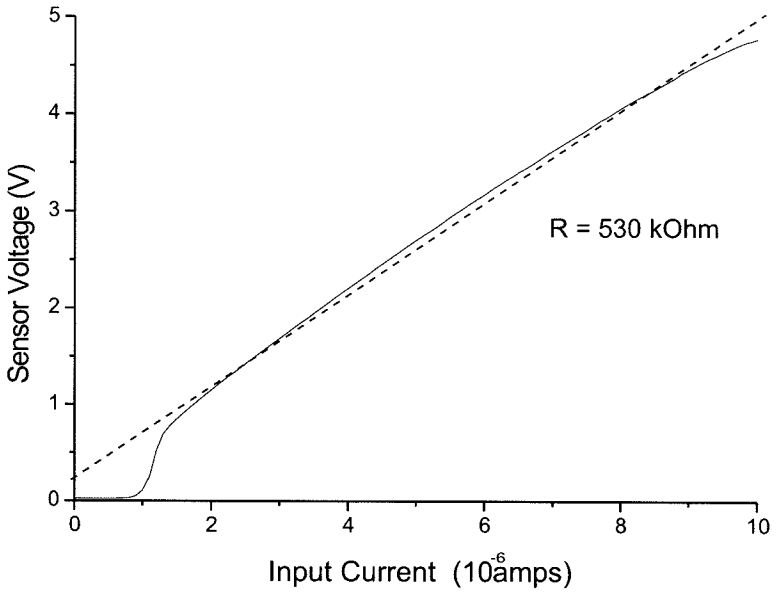


Figure 5.22: I vs V curve for sensor.

5.7 Conclusion

In this chapter I have demonstrated an array based chemical vapor sensing system. The sensor topology selected is compatible with the carbon-black based sensors. The sensor array survived the plating operation discussed in Chapter 4, and was successfully demonstrated an integrated chemical sensor array.

CHAPTER

SIX

Conclusions and Future Work

The work of this thesis was intended to make a foundation on which base the development of an integrated chemical sensing system. I have demonstrated the creation and fabrication of an array of sensors using a standard CMOS process followed by a post processing step. The arrays were simple in function, but demonstrate the feasibility of the system. In addition I have presented building for the creation of a system that will adaptively bias the sensors. A nonvolatile memory cell suitable for the implementation of on chip classifiers was also presented.

Future work should extend the results of this thesis and investigate robust methods for the deposition of arrays of different sensors material. This will enable the creation of a large set of diverse chemical sensors that can perhaps mimic the performance of biological olfaction systems.

In order to aid in the further development, a detailed analysis of the response of the sensors should be performed. These sensors are much smaller than the sensors used in the initial chemical studies of their sensing proper-

ties. Their response must be properly characterized to enable the circuitry to be targeted properly. The noise characteristics of the sensors should be investigated, in order to better understand the design requirements and to evaluate deposition methods. (note - i hope to include this in the final thesis)

Alternative sensing methods based on the swelling of polymers can also be investigate. One method is to employ capacitive sensing. It may be possible to perform chemical sensing by detecting the change in permittivity either due to the swelling of the polymer, or by a change in the permittivity of the polymer due to absorbed vapor molecules. I have already begun preliminary work into this topic. A picture of one of the capacitive test circuits in shown in Figure 6.1. The interdigitated lines comprise a capacitor over which the desired polymer will be deposited. The use of capacitive sensing has some appealing advantages. Since the sensing is effectively performed by an electrical field, electrical contact is no required. Thus we do not have to perform any postprocessing on the chips. Moreover, since the sensors are not resistive, thermal noise generated by the sensors will not exist.

Another investigation would be to control the temperature of the sensors on the chip. By controlling the temperature it may be possible to bias the resistive sensor at its percolation threshold. At this point a small swelling will cause the resistance to go from a finite resistance to an infinite resistance. It may be possible to exploit this region of operation for the use of high sensitivity chemical detection. A small microhotplate can be fabricated using a standard CMOS process, as shown in Figure 6.2. The hotplate consists of a polysilicon resistor, which will increase in temperature by passing a current through it. A metal plate over the resistor distributes the heat evenly. On the top second layer of metal I have fabricated a ring sensor described in Chapter 5. The structure

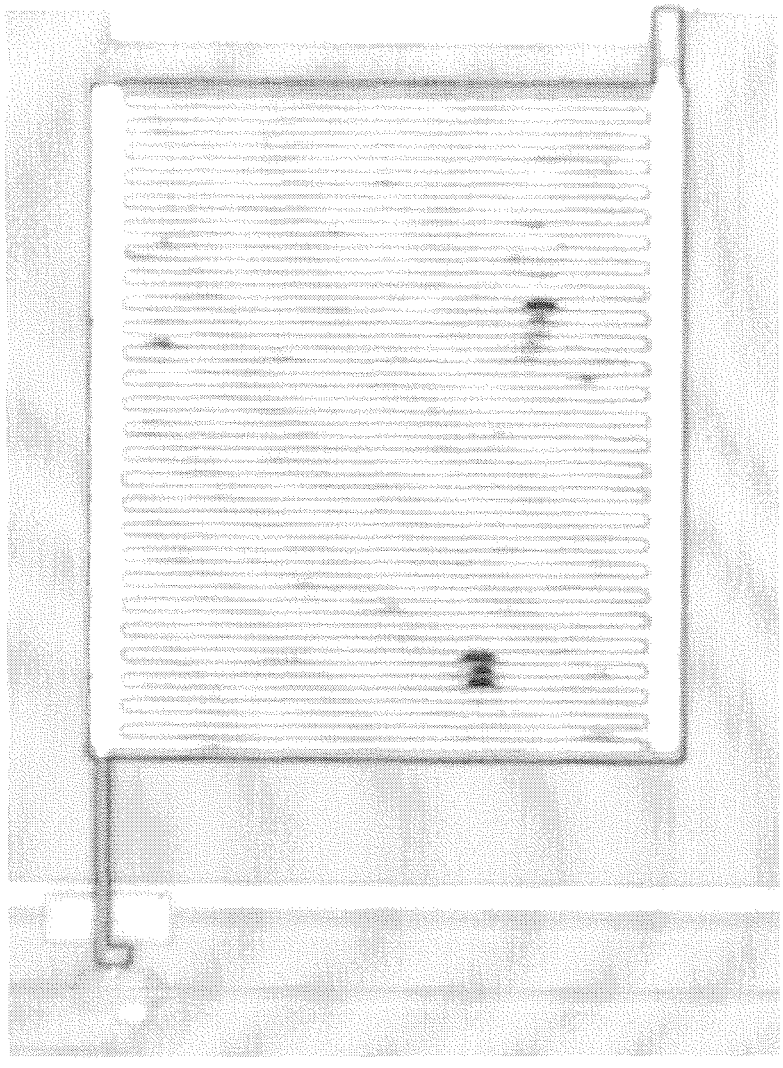


Figure 6.1: Capacitive sensing chip

will be thermally isolated from the underlying bulk silicon substrate by etching with a suitable etchant.

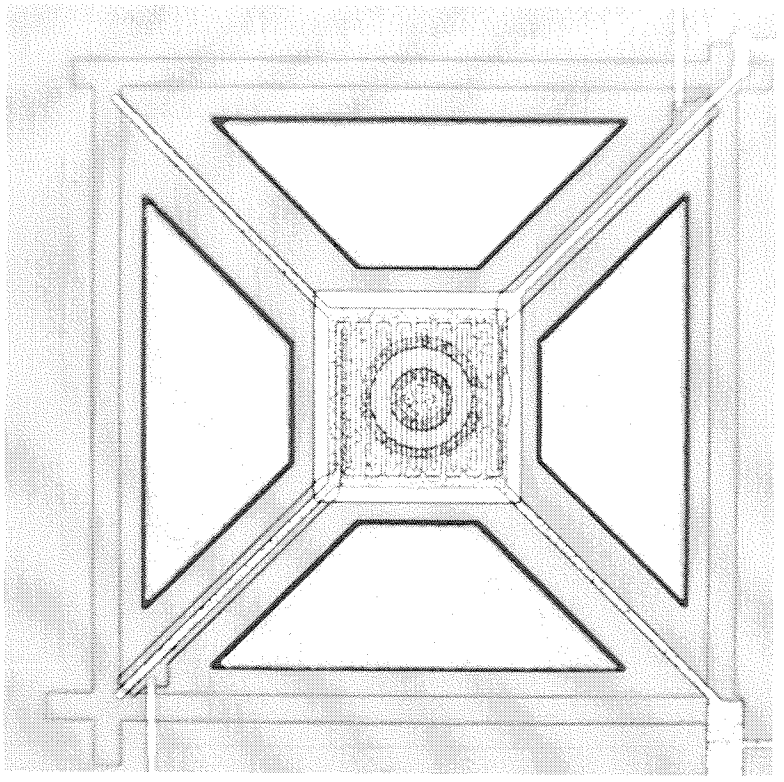


Figure 6.2: Picture of the ring sensor with integrated heater.

One interesting avenue of future research is the method of reading out the sensor. The size of the sensors is limited by the deposition technology, and the requirement of being able to make good contact with the underlying circuitry. The decrease of feature sizes available in modern CMOS processes will not likely have a great impact on the size of the sensor. This means that there is the opportunity to put a large amount of circuitry underneath each sensor cell.

A possible use of this area is to integrate a per-pixel analog to digital converter. By incorporating the converter at each cell the noise bandwidth of the

system can be reduced. As discussed in Chapter 5 (note to committee - I haven't yet!) the noise bandwidth of the multiplexed system increases with the square root of the number of elements. By incorporating a nyquist rate A/D at each cell this increase in noise bandwidth is avoided. Fowler [16] used a pixel level A/D converter in a CMOS image sensor. The drawback in that application is that the size of the circuitry compromises the pixel density. In an optical system that is a drawback; however, in the chemical sensor array we have larger pixel sizes and we build the sensors on top of the circuitry. The overall required bandwidth of an olfactory array is probably much lower than an imager, at least in certain applications. An arbitration scheme could be employed so that only sensors whose input has changed output their bit stream.

REFERENCES

[1] L. B. Buck. Unraveling chemosensory diversity. *Cell*, 83(3):349–352, 1995.

[2] L. B. Buck. Information coding in the mammalian olfactory system. *Cold Spring Harbor Symp. Quant. Biol.*, 61:147–155, 1996.

[3] L. B. Buck, S. L. Sullivan, and K. R. Ressler. Receptor diversity and spatial patterning in the olfactory system. *J. Neurochem.*, 62:S32–S32, 1994.

[4] G. Cauwenberghs and A. Yariv. Fault-tolerant dynamic multilevel storage in analog VLSI. *IEEE Trans. Circuits Syst. II-Analog Digit. Signal Process.*, 41(12):827–829, 1994.

[5] T Delbruck and C.A. Mead. Scanners for visualizing activity of analog VLSI circuitry. 1:93–106, 1991.

[6] T. A. Dickinson, J. White, J. S. Kauer, and D. R. Walt. A chemical-detecting system based on a cross-reactive optical sensor array. *Nature*, 382(6593):697–700, 1996.

[7] B. J. Doleman, M. C. Lonergan, E. J. Severin, T. P. Vaid, and N. S. Lewis. Quantitative study of the resolving power of arrays of carbon

black-polymer composites in various vapor-sensing tasks. *Anal. Chem.*, 70(19):4177–4190, 1998.

[8] B. J. Doleman, R. D. Sanner, E. J. Severin, R. H. Grubbs, and N. S. Lewis. Use of compatible polymer blends to fabricate arrays of carbon black-polymer composite vapor detectors. *Anal. Chem.*, 70(13):2560–2564, 1998.

[9] B. J. Doleman, E. J. Severin, and N. S. Lewis. Trends in odor intensity for human sind electronic noses: Relative roles of odorant vapor pressure vs. molecularly specific odorant binding. *Proc. Natl. Acad. Sci. U. S. A.*, 95(10):5442–5447, 1998.

[10] M. S. Freund and N. S. Lewis. A chemically diverse conducting polymer-based electronic nose. *Proc. Natl. Acad. Sci. U. S. A.*, 92(7):2652–2656, 1995.

[11] J. W. Gardner and P. N. Bartlett. Performance definition and standardization of electronic noses. *Sens. Actuator B-Chem.*, 33(1-3):60–67, 1996.

[12] B. Hochet. Multivalued MOS memory for variable-synapse neural networks. *Electron. Lett.*, 25(10):669–670, 1989.

[13] Brad Minch. Personal communication, 1996.

[14] E. J. Severin, B. J. Doleman, R. D. Sanner, M. C. Lonergan, R. H. Grubbs, and N. S. Lewis. Array-based vapor sensing using chemically sensitive carbon black polymer resistors. *Abstr. Pap. Am. Chem. Soc.*, 213:35–BTEC, 1997.

[15] E. J. Severin, R. D. Sanner, B. J. Doleman, and N. S. Lewis. Differential detection of enantiomeric gaseous analytes using carbon black-chiral poly-

mer composite, chemically sensitive resistors. *Anal. Chem.*, 70(7):1440–1443, 1998.

[16] D. X. D. Yang, B. Fowler, and A. El Gamal. A nyquist-rate pixel-level ADC for CMOS image sensors. *IEEE J. Solid-State Circuit*, 34(3):348–356, 1999.