IC Fabrication

Fall 2003

Josh Fryman, Chad Huneycutt
{ fryman, chadh }@cc.gatech.edu

December 9, 2003

Contents

List of Figures

List of Tables

Chapter 1

Basic Characterization

1.1 Theory and Methodology

The growth of silicon dioxide for semiconductors is necessary to provide a high-quality insulator. The layer is called the passivation layer, as it isolates device components – thus it is grown at various stages of circuit fabrication for this isolation property. This isolation is used, for example, to create a barrier between metal layers to support more complex circuit designs. The layer also serves as a barrier during impurity diffusion or implantation.

1.1.1 Field Oxide Growth

In order to grow a good barrier, the wafer must be etched clean of any oxide buildup. The technique for scouring the wafer, developed by the RCA company, is called the RCA Cleaning Process. This is performed first in the oxide growth process. Once the wafer has been cleaned and dried, it is transferred to a furnace where dry, wet, or a combination of oxide growth techniques are applied. By controlling the time, temperature, and oxidation method, oxides of arbitrary thickness can be grown.

RCA Cleaning Process

The RCA process is used to thoroughly clean the wafer of all surface contaminants before critical points of the fabrication process. The RCA process, developed by the RCA company, uses three stages of cleaning: organic cleaning, oxide stripping, and ionic cleaning. This is followed by a rinse and dry, after which the oxide growth process can begin.

Organic Cleaning

In a quartz tank, a solution of 1000 mL DI (H_2O), 200mL H_2O_2, 200mL NH_4OH is prepared. The solution is heated to $80^\circ C$ while the wafers are submerged in the solution. When the solution begins to bubble, the wafers are left in for exactly 15 minutes. After 15 minutes, the wafers are rinsed in a DI H_2O, N_2 bubblier rinse tank for 5 minutes. Organic cleaning removes basic particles, grease, oils, etc.

Oxide Stripping

After the 5 minute rinse following the organic cleaning, the wafers are submerged in a solution of 4500 mL DI H_20, 50 mL HF for 15 seconds. Immediately after this, the wafers are transferred back to the bubblier tank for 30 seconds. This stage of cleaning is to remove oxide buildup.

Ionic Cleaning

For the final cleaning step, the wafers are transferred to the ionic solution, which is prepared by heating 1000 mL DI H_2O, 200 mL H_2O_2, and 200 mL HCl to 80° C. The wafers are kept in the solution for 15 minutes. The wafers are then transferred to the bubblier rinse tank for a final 15 minutes.

Drying

Finally the wafers are transferred via the RCA carrier box to the spin dryer for automated drying.

Oxide Growth Process

Two different methods for oxidation are used: wet and dry. The primary reaction extracts oxygen and combines it with silicon. The dry method provides a denser layer, while the wet method is much faster. The reactions for both wet and dry oxidation are shown below.

$$\begin{aligned} dry &: \mathrm{Si} + \mathrm{O_2} \rightarrow \mathrm{SiO_2} \\ wet &: \mathrm{Si} + 2\,\mathrm{H_2O} \rightarrow \mathrm{SiO_2} + 2\,\mathrm{H_2} \end{aligned}$$

The oxide growth is approximated by the Deal and Grove Model, which represents oxidant diffusion from an ambient gas through the oxide and into the silicon wafer. This model uses a steady-state assumption, equating the values of the flux through each interface. The thickness of the oxide, d_0, can be calculated from equation 1.1. The values A, B, and τ are calculated based on time, temperature, etc.

$$d_0^2 + Ad_0 = B(t + \tau) \tag{1.1}$$

The wafers were loaded into the furnace when the furnace was at 900° C. After insertion into the "mouth" of the furnace in a quartz carrier, the wafers were allowed to warm up from room temperature for 5 minutes. After the initial 5 minutes, the furnace was ramped up to 1100° C at a rate of 20° C per minute. Once the furnace reached 1100° C, the wafers were dry oxidized in the center of the furnace for 5 minutes. Steam (wet) oxidation then took place for 60 minutes. Finally, steam was turned off and dry oxidation continued for 5 more minutes. The wafers were then allowed to cool and returned to storage.

1.1.2 Wafer Substrate Characterization

This section describes tests to measure the properties of a wafer.

Hot Probe Test

A hot probe is applied to a wafer to determine the movement of the majority carrier type. The majority carriers under the probe have a higher thermal energy than those in the cooler regions of the wafer. As they diffuse away from the hot probe, a net charge is left behind with a polarity opposite of the majority carrier type.

A control wafer was selected at random. Using a Weller non-temperature regulated soldering station, heat was applied approximately in the center of the wafer. A microammeter was then used to probe for current flow near the tip of the hot soldering iron. When current flow is positive, the material is *n*-type. When current flow is negative, the material is *p*-type.

Wafer Thickness and Four Point Probe

Wafer resistivity is a measure of the electrical resistance in the wafer. Resistivity is the inverse of conductivity, or the ease with which electrons (or holes) can move through the material.

Another control wafer was selected at random. A micrometer was used to repeatedly sample the thickness of the wafer at different locations on the edge. The values for wafer thickness were smoothed by discarding the maximum and minimum, with the remaining values averaged to create the smoothed average wafer thickness (W_A). After thickness was determined, the wafer was then placed into a Veeco Four Point Probe, Model FPP-100 [1], to measure resistivity (ρ). The smoothed average thickness was dialed into the front panel, and then repeated measurements were made at different locations on the wafer. The results of the four point probe were also smoothed and averaged to create a final value (ρ_A) for use in calculations (The unit was power-cycled between measurements due to machine problems.) The output was collected and converted to resistivity according equation 1.2, provided by the lab instructor and in textbook section 6.1.3. The trailing scalar (0.991) is a correction factor applied due to the size of the wafer.

$$\rho\ [\Omega \cdot \mathrm{cm}] = \rho_A\ [\Omega]\ \cdot W_A\ [\mathrm{mil}]\ \cdot 0.001\ \left[\frac{\mathrm{in}}{\mathrm{mil}}\right] \cdot 2.54\ \left[\frac{\mathrm{cm}}{\mathrm{in}}\right]\ \cdot 0.991\,[\Omega \cdot \mathrm{cm}] \tag{1.2}$$

To find the dopant concentration, the primary equation of conductivity is used as shown in equation 1.3, assuming that the background concentration of the wafer is constant. In our case, with *n*-type material, the *p* term is negligible, and we can reduce the form to that shown in equation 1.4 for *n*-type dopant concentration. Here, q is the charge of an electron (1.602×10^{-19} [C]) and μ_n is the mobility of electrons in the material (given 1800 $\mathrm{cm}^2/\mathrm{Vs}$).

$$\frac{1}{\rho} = \sigma \quad = \quad q\,(\mu_n \cdot n + \mu_p \cdot p) \tag{1.3}$$

$$n \quad = \quad \frac{1}{\rho \cdot \mu_n \cdot q} \tag{1.4}$$

Crystal Orientation

The crystal orientation in terms of Miller indicies can be determined based on making small cuts on the wafer edge and observing the breakage pattern.

While supplied with the orientation of the crystals by the lab instructor as <110>, a diamond scribe was applied to another random wafer to observe the pattern on breaking. Repeated application of the scribe to different locations was performed, with the breaking pattern recorded. The initial cut was performed on the primary flat. In 45° increments clockwise, additional cuts were performed. Based on the material handed out in lab, the breakage pattern was then used to determine the crystal orientation.

[1] Only two switches were toggled on: Auto Penetration and Auto Type.

1.2 Results

1.2.1 Field Oxide Growth

RCA Method

The lab instructor indicated that the RCA process on the first-time cleaning of a wafer must be followed exactly for all time limits. During later processing we may be able to shorten certain steps.

Field Oxide Growth

The oxide growth predicted by the Deal and Grove Model, shown in equation 1.1, suggest a thick oxide creation. By this model, we calculate a predicted oxide thickness of the first 5 minute dry growth as d_A:

$$\begin{aligned} d_A^2 + (0.090\ [\mu m])\, d_A &= \left(0.027\ \left[\frac{\mu m^2}{hr}\right]\right)(0.083\ [hr]) \\ d_A &= 0.0204\ [\mu m] \end{aligned}$$

Now modeling for the next 60 minutes of wet growth, we find d_B including d_A as:

$$\begin{aligned} d_B^2 + (0.110\ [\mu m])\, d_B &= \left(0.510\ \left[\frac{\mu m^2}{hr}\right]\right)\left(1.000 + \frac{d_A^2 + 0.110\ [\mu m] \cdot d_A}{0.510\ \left[\frac{\mu m^2}{hr}\right]}\ [hr]\right) \\ d_B &= 0.663\ [\mu m] \end{aligned}$$

Finally, for the last 5 minute dry growth, we find the total oxide growth d_C as:

$$\begin{aligned} d_C^2 + (0.090\ [\mu m])\, d_C &= \left(0.037\ \left[\frac{\mu m^2}{hr}\right]\right)\left(0.083 + \frac{d_B^2 + 0.090\ [\mu m] \cdot d_B}{0.027\ \left[\frac{\mu m^2}{hr}\right]}\ [hr]\right) \\ d_C &= 0.665\ [\mu m] \end{aligned}$$

Using another random wafer, we observed a direct color of green to blue-green. This would correspond to an oxide thickness of between 0.50 - 0.54 μm. No ellipsometer measurements were taken during this lab period.

As reported by the lab instructor[2], our control wafer had a measured oxide thickness via ellipsometer of 0.5309 μm. The measured thickness more accurately reflects the observed colors on the random control wafer.

The lab instructor also reported that our wafer had a refractive index of 1.46. This value corresponds well to a good SiO_2 oxide on the wafer.

1.2.2 Wafer Substrate Characterization

There was insufficient time to permit students to test the group control wafer. Instead several wafers of the same type as ours were used to measure the characteristics and become familiar with lab equipment.

Hot Probe Test

A current of -5 μA was observed. The negative current indicates an *p*-type material of the random wafer.

[2] Due to our Thursday lab schedule, we were unable to measure these ourselves as our wafers were unfinished by the end of the lab period.

Wafer Thickness, Resistivity, Dopant Concentration

The measurements by micrometer for the random wafer were 11.0, 11.0, and 10.5 μm. The value used in our calculations was 11.0 μm. The measured sheet resistivities in Ω per square via the four point probe test were 298, 292, 282, and 285. The value used in our calculations was 288.5 Ω. By equation 1.2, we calculate bulk resistivity as:

$$\begin{aligned} \rho &= 288.5\ [\Omega] \cdot \frac{11.0}{1000}\ [in] \cdot 2.54\ \left[\frac{cm}{in}\right] \cdot 0.991 \\ \rho &= 7.988\ [\Omega \cdot cm] \end{aligned}$$

The dopant concentration, as per equation 1.4 is calculated as:

$$\begin{aligned} n &= \frac{1}{7.988\ [\Omega \cdot cm]\ \cdot 1800\left[\frac{cm^2}{Vs}\right]\ \cdot 1.602 \times 10^{-19}\ [C]} \\ n &= 4.336 \times 10^{14}\ \left[\frac{carriers}{cm^3}\right] \end{aligned}$$

Crystal Orientation

The crystal was observed to break cleanly on a random wafer in the directions perpendicular to the primary flat and parallel to the primary flat. Other cuts resulted in curving breaks. This suggests a basic crystal orientation of <100>, but as indicated by the lab instructor, we are using crystal orientations of <110>.

1.3 Conclusion

Our calculated oxide thickness was 0.665 μm, while the measured was $0.531 \mu m$. The error between these is approximately 20%. Sources for error include errors in procedure (inexact time recordings due to the low resolution timing device); assumptions in the model that each flux rate is equal for each flux; and lack of consideration for ambient impurities in the environment. A larger source of error was the time the wafer sat in the "mouth" of the furnace warming up, and then waiting for the furnace to reach peak temperature. Similarly, the cool-down period was also unaccounted for.

The wet oxidation method provides rapid growth and is useful for the limited lab time available. The instructor performing the RCA cleaning for us was unexpected, but understandable given the risky nature of the process and the amount of time involved. Were students to individually carry out the tasks, the procedure would take too long, and if one student is designated to perform the task for all wafers, the risk is that a mistake in procedure could destroy all wafers.

Chapter 2

Photolithography

2.1 Theory

In CMOS fabrication, photolithography refers to the process of transferring a pattern from a mask to the wafer by employing a photosensitive, chemically resistant layer. This "photoresist" is *exposed* with light, resulting in regions that will wash away when the wafer is *developed*. These regions can then be etched and the remaining photoresist will provide a barrier to the rest of the wafer.

The steps involved in photolithography are broken down into the following simple procedures, each of which we explain briefly:

1. Surface Preparation: *Dehydrate and prime wafer surface*
2. Photoresist Spin: *Apply a thin layer of photoresist to wafer*
3. Softbake: *Partial evaporation of resist*
4. Align & Expose: *Precise pattern alignment and UV exposure*
5. Development: *Removal of unpolymerized resist*
6. Inspection A: *Verify proper alignment*
7. Hardbake: *Final solvent evaporation*
8. Etching: *Strip oxide in open areas*
9. Inspection B: *Verify outcome as desired*
10. Stripping: *Remove hardened resist*

2.1.1 Surface Preparation

This step aims to provide a dry and clean wafer to apply photoresist to. The wafers, previously RCA cleaned and placed clean into N_2 storage boxes, are warmed to $200^\circ\,\mathrm{C}$ for several minutes (5+). This heating is to evaporate any bulk water on the surface of the wafer that may have accumulated. A dry surface will promote adhesion between the photoresist and the wafer. Once dehydrated, the wafer is transferred to a photoresist spinner system. After placement in the center of the spinner chuck, a primer is applied in sufficient volume to cover the entire surface. The primer's purpose is to bind any remaining water molecules on the surface of the wafer in order to further promote adhesion of the photoresist. The most common chemical used is Hexamethyldisilazane (HMDS), a patented compound developed by IBM. After the wafer is coated and sufficient time elapses, the spinner is activated to spin-off the excess primer.

2.1.2 Photoresist Spin

There are two types of optical photoresist that we consider: positive and negative. Both are made of polymer compounds, which are kept in liquid state by addition of solvents. The key difference between the resists is the

mechanism by which polymer chains are linked together or broken apart. For optical resists, a light source such as a UV lamp radiates intense energy into the resist polymers. How that light affects a resist determines whether it is negative or positive. To prevent accidental light exposure, facilities generally use filtered light (e.g., yellow filters on all lights).

Negative photoresists are based on polyisopreme, which occurs naturally in rubber. Prior to light exposure, they exist in a form without cross- links. When sufficient light energy is applied, the polymers become cross-linked, or *polymerized*. Figure 2.1, on the left, shows the negative photoresist system. (Figures from *Microchip Fabrication*, Peter van Zant, 2nd ed.)

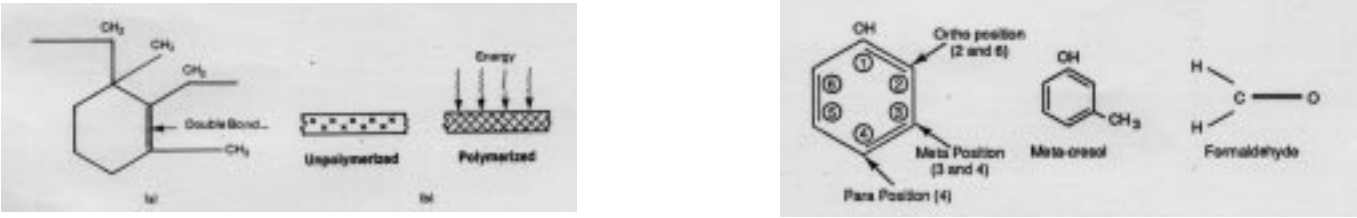

Figure 2.1: **Basic negative (left) and positive (right) photoresist structure and processing.**

Positive photoresists are based on phenol-formaldehyde, or phenol-formaldehyde novolak resin. In the unexposed resist, the polymer is insoluble. The originally cross-linked polymers are broken into short chains during exposure. The light energy causes *photosolubilization*, or the conversion to a more soluble state. Figure 2.1, on the right, shows the positive photoresist system.

The photoresist is applied in a small volume to the top of the wafer. The wafer is then spun for a short period of time to shed excess resist material while keeping a thin layer of resist bonded to the wafer. Ideally the resulting layer is uniformly thin and defect free, with a thickness on the order of 1μm with variation of approximately 1%.

2.1.3 Softbake

After application of the photoresist, the wafer is transferred to a Softbake oven. This oven will heat the wafer to a relatively low temperature (100°C) and will only be used for a short period of time. The primary purpose of this stage is to partially evaporate the excess solvents in the remaining photoresist. This evaporation will enable the photoresist to sustain more contact from the following steps without distorting the desired mask image.

2.1.4 Align & Expose

Masks control which regions of the photoresist get exposed. Many masks are required to complete the fabrication process, and it is crucial that each mask is aligned identically with respect to the wafer. Very small misalignments can render the entire wafer useless. For example, a contact might end up going into the bare Silicon substrate instead of into the gate of a transistor. As the number of masks required for a chip increases and the feature size decreases, the tolerance for misalignment also decreases.

A common technique for assisting with mask alignment is to put reference marks on the masks; hash marks, rectangles, and cross-hairs are common. These reference marks are transferred to the wafer, and then can be used as a reference for the next mask, which will have the same mark in the same place.

As discussed in section 2.1.2, exposure of the photoresist by a high-power light source such as UV lamps causes a chemical reaction to take place within the photoresist. The amount of exposure, measured in Watts, controls how strongly the reaction takes place. Too much or too little exposure will result in an incorrect pattern, which will then have to be removed before starting over on the photolithography process. Negative photoresists require 5 to 15 seconds of exposure, while positive photoresists can require two or three times as long.

2.1.5 Development

After exposing the photoresist to the light source, all unlinked polymer resist material needs to be removed. This is accomplished by transferring the wafer to a bath of developing solution, the contents of which are dependent on the type of resist used. Negative resists develop in readily available solvents and have a wider tolerance in processing. Positive resists require carefully constructed solvents and tight control over temperatures used. At the end of the development process, only the desired pattern of resist is left on the wafer. All unexposed areas will be etched later for additional processing.

2.1.6 Inspection A

Stopping at this point and placing the wafer in a Quality-Control (QC) microscope station permits the verification of the remaining photoresist pattern. At this stage, alignments are verified and consistency of the resist layer is observed across the entire wafer. Sufficient deviation from the intended results can cause failure in the wafer dies, and this QC check ensures that proceeding will not waste time. If necessary, the wafer may be cleaned and the photolithography process started over.

2.1.7 Hardbake

Much like the Softbake discussed in section 2.1.3, the Hardbake is designed to remove solvents from the photoresist layer. Unlike the Softbake, however, this step aims to remove all of the solvents remaining. Therefore it is conducted at a higher temperature ($120^{\circ}\mathrm{C}$) for a longer period of time. Some resists can also be activated by heat, and this baking stage could further bond the resist to the wafer surface.

2.1.8 Etching

After the Hardbake, the wafer has areas covered by bonded polymer photoresist, and areas that are exposed to the underlying wafer surface. At this stage, a Buffered Oxide Etch (BOE) is used to strip the exposed wafer surface regions. This process can be used for diffusion, metallization, or other needs of the multiple stages during wafer processing.

The BOE will not etch those areas covered by the hardened resist material initially. However, the amount of surface stripped is entirely dependent on the time the wafer is left in the BOE solution. If left in too long, the BOE can actually etch below the resist layer and begin removing material that is desired. Therefore, careful monitoring of time is essential to have a successful etch.

2.1.9 Inspection B

After etching the wafer, a final QC check is made to verify the results are as desired. At this stage we can evaluate whether to proceed or re-perform the prior stage(s) in order to obtain a more precise result. Compounded errors in the registration over multiple masks may yield non-working components within the wafer.

2.1.10 Stripping

After final QC is complete, the only material left for removal is the hardened photoresist. During this stage, appropriate solvents at the correct temperature are used to break away the cross-linked polymers. The final result will be a wafer ready for the next phase of processing. It is important to remove all of the resist material, as it will combust in later stages such as high- temperature diffusion. This combustion will cause severe problems with contamination of furnaces and wafers, and therefore it is imperative that all resist be removed.

2.2 Procedure and Observations

For this report we have done six photolithography steps: P-wells, P-FETs, N-FETs, and contacts. The steps are the same, except that no mask alignment was necessary for the P-well pre-dep mask – only an approximate initial registration based on the primary flat of the wafer was used.

2.2.1 Surface Preparation

Prior to the photolithography process, the wafers are thoroughly dried in the Blue M Electric Stabil-Therm Gravity Oven, which we used as a *dehydration oven* at 200° C . Our wafers were already in the oven, placed there by the lab instructor, and we were informed they had been heating for sufficient time. The minimum recommended time is 5 minutes, but longer is better.

We moved the wafers to the Integrated Technologies P-2604 photoresist spinner. Placing the wafer as close to center as possible on the elevated chuck, we then covered the wafer entirely with Shipley HMDS primer. This was allowed to stand for 10 seconds. The spinner was then activated to remove the excess developer while keeping a thin layer on the surface.

Note: A visual cue to tell when the primer has spun to its target level is when a **rainbow** *pattern much like oil in water is observed.*

2.2.2 Photoresist Spin

With the wafer still attached to the photoresist spinner, sufficient volume of Shipley 1813-Positive photoresist was then applied such that the center 2/3 of the surface was covered. The spinner was then activated for an automated 30 second spin at 3500 rpm. This resulted in a clean thin layer of photoresist without excessive buildup on the underside of the wafer due to excessive application of the liquid resist. Figure 2.3 shows the spin speed for PR thickness chart (from Shipley datasheets). Figure 2.2 is from a datasheet that characterizes the exposure time and PR performance.

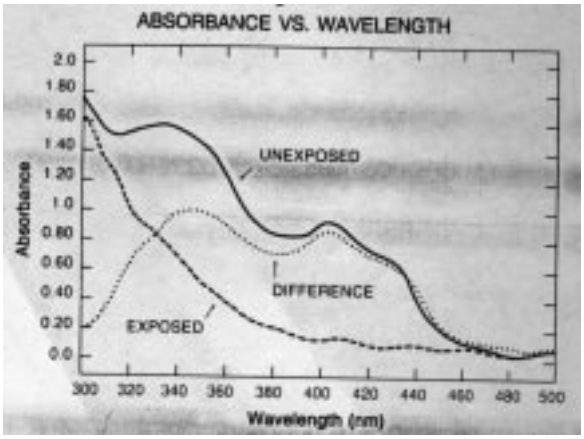

Figure 2.2: **Shipley 1813P photoresist exposure data.**

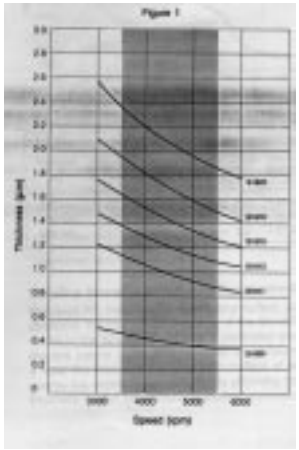

Figure 2.3: **Shipley 1813P photoresist energy data.**

2.2.3 Softbake

For the Softbake stage, the wafers were again transferred to a Blue M Gravity Oven. This oven is typically set for 95 to 100° C . For our wafer processing, 95° C was used, heating the wafers for 5 minutes.

2.2.4 Align & Expose

After the Softbake, each wafer was loaded into the Karl Suss MJB3 Mask Aligner and exposure system with a UV lamp. The lamp power was set to 200W, while exposure time was only 30 seconds for an exposure energy of 6mW. The UV wavelength was $\lambda = 365$. Of the three types of exposure systems (contact, proximity, and projection), the Karl Suss unit we used uses contact exposure.

The first mask (P-wells) required no alignment, although we did make sure to begin the convention of orienting the wafer on the chuck so that the primary flat was on the left side (9 o'clock). The second mask (P-FETs) required alignment to the transferred pattern of the first mask on the wafer. The lab instructor pointed out how useless the mask registration patterns were, and instructed us to use the primary horizontal and vertical features on the mask pattern directly. Alignment of successive masks to the first mask is attained via careful manipulation of the $< x,\ y,\ \theta >$ controls on the mask aligner. By experimental process the bulk of correction went toward the angular rotation, θ, before performing fine $< x,\ y >$ adjustments.

Figure 2.4, reproduced from *Fundamentals of Semiconductor Fabrication, May and Sze*, illustrates how the exposure energy affects positive or negative photoresist layers.

Note: It is important to warn others in the lab when exposing so no one inadvertently looks at the UV light.

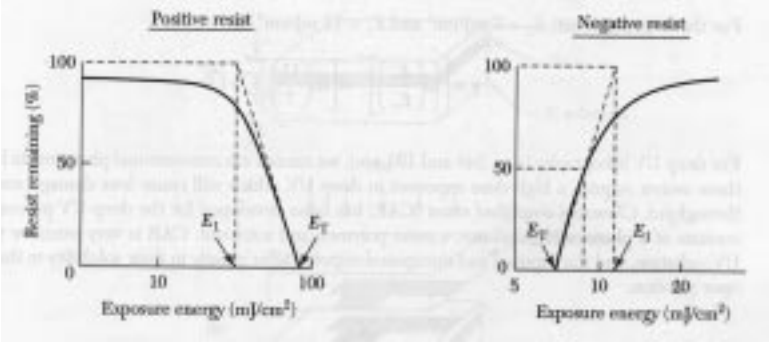

Figure 2.4: **Exposure effects on positive and negative PR.**

2.2.5 Development

Immediately after exposure, the wafers were transferred to a bath of Shipley MF319 Developer. This solution was used for 1 minute and 30 seconds, with automated agitation via magnetic spinners in the solution. In order to place the wafers in the developer, they were first transferred to plastic/Teflon wafer holders.

Note: To verify the wafer is installed correctly into the wafer holder, shake it gently while holding it above and close to the surface of the lab bench. If a faint rattling sound is heard, tighten further.

After the developer solution, the wafers were DI rinsed for 10 seconds and dried with pure N_2 air guns.

2.2.6 Inspection A

The underlying wafer oxide appears blue-green, as reported previously. The photoresist covered areas appear dark red or burgundy. For the first mask, the QC test is simply whether the pattern appears clearly and the oxide is visible. For the other masks, the verification is to ensure that all features are aligned properly with the preceding layer. No reworking was necessary.

Several pictures of different features illustrate this QC inspection. In figure 2.5, a transistor with 40 μm channel width. Figure 2.6 shows that even though teflon tweezers are used, extreme care is essential to avoid damaging the PR coating. Figure 2.7 shows several resistors Finally, figure 2.8 illustrates how difficult it is to use the default hash registration pattern. By using the long vertical and horizontal features on the masks, a good alignment was achieved even when the mask aligner was broken in the Y-axis during one phase.

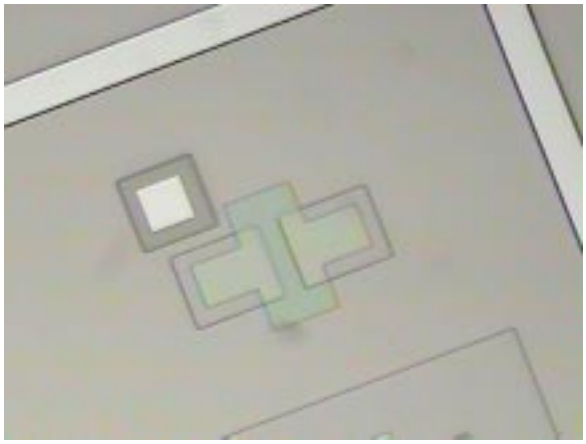

Figure 2.5: **Transistor features under PR layer.**

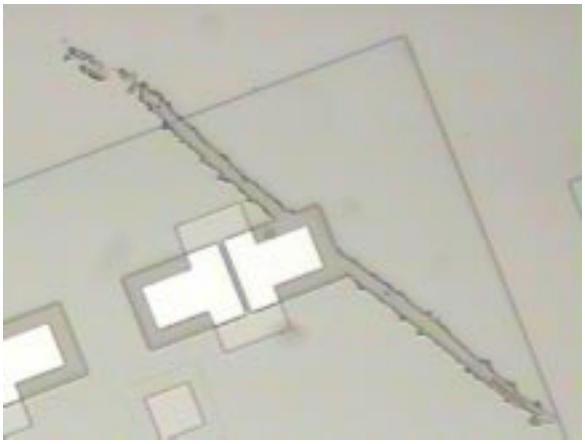

Figure 2.6: **Damage to PR layer from tweezers.**

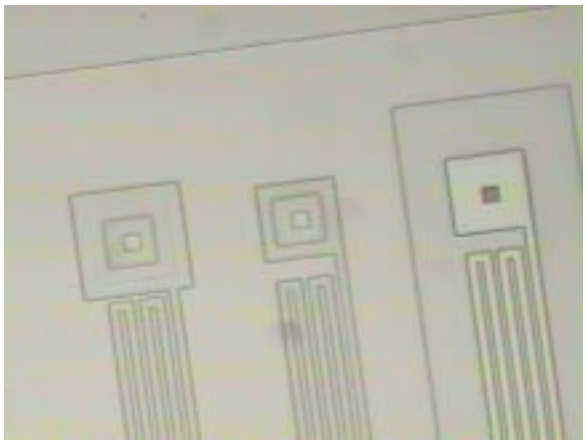

Figure 2.7: **Resistors in PR pattern.**

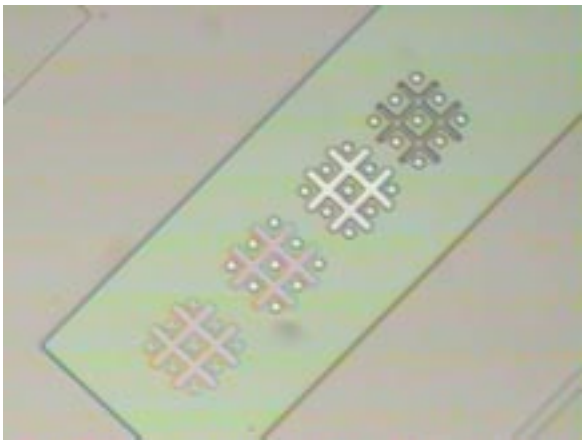

Figure 2.8: **Default hash mark registration is unusable.**

2.2.7 Hardbake

After QC inspection, the wafers were transferred to a third Blue M Gravity Oven for the Hardbake process. These ovens were set to 120° C and the wafers were baked for 5-30 minutes, depending on the thickness of the oxide being stripped, to evaporate the remaining solvents in the resist material.

2.2.8 Etching

After the hard bake, the wafers were directly transferred to a 6:1 Buffered Oxide Etch with HF for 6 minutes. After BOE completion, the wafers were rinsed in DI H_2O for 15 seconds, and then dried in the N_2 air guns.

Note: It is important that no agitation occur in the BOE stage.

2.2.9 Inspection B

During this second inspection stage, observations are made that the desired effect has occurred. For example, in the case of oxide etching, it was necessary to ensure that the oxide had been stripped and the non-resist regions appear clear (where the bare Si substrate is now visible). If further etching is required before stripping, this QC test will indicate it. No additional work was necessary.

Several pictures of different features illustrate this QC inspection. In figure 2.9, the over-etching ((in this case, under the surface) causes a rainbow-like pattern in reflected light from the QC station. Figure 2.10 shows that these reflected light patterns give a 3-D feel to the resulting stages, which are used to continue the alignment process. The continual buildup of these shadows, however, causes precise alignment difficulty as the shadows become thicker due to multiple oxide growths. Figure 2.11 shows the same resistors from figure 2.7 after B.O.Etch application. Finally, figure 2.12 illustrates that repeated contact alignment is slowly damaging the features of the glass mask. Here, what would be straight edges are rough from slight tearing on the glass mask pattern.

Figure 2.9: **High magnification of 10μm transistor gate with over-etching.**

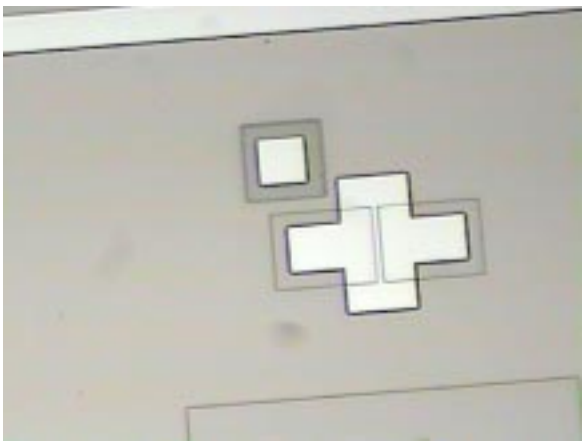

Figure 2.10: **A three-D appearance results from shadows and color variations.**

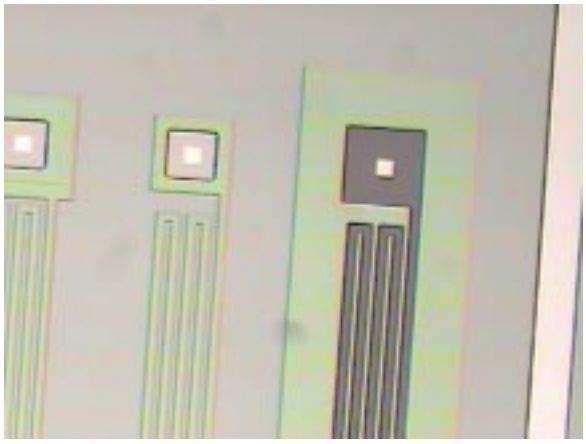

Figure 2.11: **Resistors from QC A after B.O.Etch.**

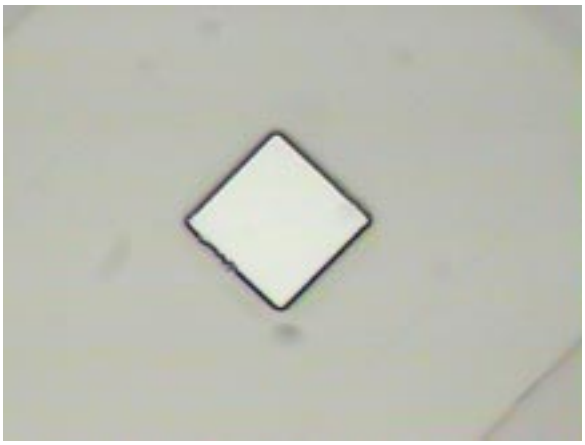

Figure 2.12: **Rough edges from glass mask damage by repeated contact alignment.**

2.2.10 Stripping

The final stage is removal of the hardened photoresist. We placed our wafers in Shipley 1165 Stripper, which was heated to 70° C , for 3 minutes. Following this, we rinsed the wafers with DI H_2O and dried a final time with the N_2 air guns. After drying, the wafers were returned to secure storage. Figure 2.13 shows the Shipley datasheet graph of performance for the 1165 stripper.

Note: No agitation is to occur during the stripping stage. Moreover, the final DI rinse is complete when the DI water running over the wafer no longer appears "foamy" as soap bubbles do.

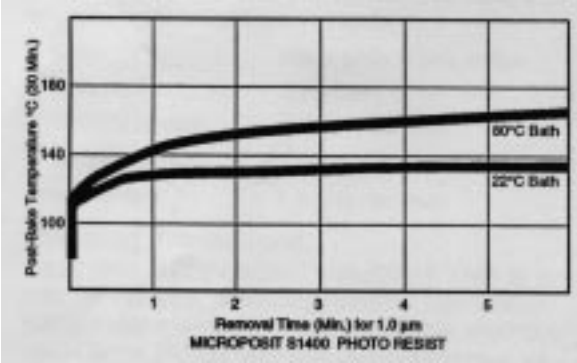

Figure 2.13: **Performance of the 1165 stripper by Shipley, copied from company datasheet.**

2.3 Conclusion

The most critical stage of the photolithography process is the mask alignment. Rotation or X-/Y-errors can render some or all of the dies on a wafer useless. While extra time should be spent during this stage, it is imperative to build up sufficient skill that accurate alignment can occur in short time periods. Given that fabrication facilities are very expensive, too many errors could easily make any project financially unsustainable.

Given that the mask alignment is the most critical stage, it was surprising to find the mask patterns provided in our lab have almost useless registration marks. The provided masks use a sequence of "#" marks on the first layer. Successive layers fill in small dots in a tic-tac-toe pattern on the first registration mark, while providing a new registration "#" to the side. As each layer is aligned, compound errors are easily introduced from this system.

Instead, we use the primary vertical and horizontal features on the mask. By careful alignment to these, it was quick and easy to calibrate each layer. As a side effect, the original mask registration patterns were near-perfect in final alignment.

The other key stages are the careful QC inspections and the close monitoring of time spent in Developer or BOE solutions. Too much or too little of either will result in errors and starting over.

Chapter 3

Diffusion

3.1 Theory

Conductive regions on a wafer are essential for making operational circuits. *Doping* is the process of inserting ions into the silicon substrate to generate conductive regions. These regions, whether P- or n-type, form critical junctions that determine electrical behavior.

A p-type region is rich in electron holes (h^+), whereas a n-type region is rich in free electrons (e^-). Depending on the dopant material used (B, P, As, etc.), either free electrons or electron holes are inserted by replacing silicon atoms in the lattice. The two most common methods for doping are *diffusion* and *ion implantation*. In this lab, we only use diffusion. The dopant atoms move by either *vacancy* or *interstitial* diffusion. Vacancy diffusion fills holes in the lattice with impurity atoms, while interstitial diffusion fills gaps in the lattice structure.

The primary characteristics of doping are determined by the *junctions* or interfaces between two regions of different doping characteristics. The junctions can be classified as either from n-type material to p-type material (NP) or *vice-versa* (PN). The *junction depth* is defined as the point where the two regions have equivalent concentrations (*i.e.*, $N_{h^+} = N_{e^-}$).

3.1.1 Diffusion

Diffusion is the process whereby the random movement of atoms in adjacent materials results in a *flux* (F) of the atoms from the higher concentration material (*source*) to the lower concentration material (*target*). The flux can be expressed as the rate the atoms from the source pass into the target:

$$F = -D\frac{\partial C}{\partial x} \tag{3.1}$$

where F is the flux, C is the concentration of the source, and D is known as the diffusion coefficient or *diffusivity*. If we assume Intrinsic diffusivity (*i.e.*, D is constant for a given temperature), we can model diffusion for time t by Fick's Law:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{3.2}$$

Diffusion is used during doping to achieve three goals: (1) insert a specific number of dopant atoms on our wafer; (2) create a junction at a target depth; and (3) create a specific distribution of dopant atoms within the region. These goals are achieved by two primary stages, *deposition* and *drive-in*. Each stage is briefly described next.

Deposition

In the deposition stage (sometimes referred to as predeposition), the dopant atoms are introduced into the wafer. Diffusivity increases with temperature, so high temperature diffusion ovens speed the movement of dopant atoms.

The *solid solubility* of the dopant limits the total concentration of atoms that can be introduced at a given temperature, much like adding sugar to iced-tea.

Deposition occurs when the dopant concentration exceeds solid solubility, at a sufficiently high temperature to permit the diffusion process. Using such a high concentration of dopant eliminates the variable of concentration and simplifies the result to be a function of temperature only. Thus the deposition occurs at *solid solubility*. The characteristic plot of the predeposition diffusion concentration and depth is defined mathematically by the complementary error function.

The sequence of steps in a deposition process are:

1. Preclean & Etch: *Remove particles, stains, minor oxides*
2. Tube deposition: *Pre-heat, diffusion, and post-cool*
3. Deglaze: *Removal of thin oxide grown during deposition*
4. Evaluation: *Test electrical properties of result*

Drive-in

The drive-in stage is performed without a dopant source material, but is otherwise identical to the deposition stage. The primary function of drive-in is to distribute the dopant atoms deeper into the wafer structure. During deposition a high-concentration but shallow layer was introduced just below the surface. The drive-in will rearrange the dopant atoms according to a *gaussian* distribution.

A secondary process during drive-in is the oxidation of the wafer surface. This oxidation affects the underlying dopant atoms. With n-type dopants, the resulting effect is called *pile-up* as the dopants segregate into the silicon region and do not enter the oxide. With p-type dopants, the dopants are more soluble in the oxide and are therefore drawn into it. The n-type behavior increases the surface concentration of dopants and the p-type behavior lessens the surface concentration. Either result alters the electrical characteristics over what was originally intended.

In general, a period of "dry" drive-in will occur in a furnace with N_2 gas as a transport. At the end of the drive-in time, a thin oxide will be grown by insertion of O_2 gas into the chamber.

3.1.2 Models

As stated earlier, diffusion can be modeled by Fick's Law. The two doping stages yield different solutions to Fick's Law because they have different initial and boundary equations. Deposition is modeled as a *constant surface concentration* or *infinite source* diffusion, since the concentration of dopant at the wafer surface stays constant at the solid solubility of the dopant. Drive-in is modeled as a *constant total dopant* or *limited source* diffusion, since no additional dopant is introduced.

Infinite Source

To solve Fick's diffusion equation, we need an initial condition $C(x,0)$ and two boundary conditions $C(0,t)$ and $C(\infty,t)$. For deposition, we are depositing either N- or p-type dopant into a region of opposite type, indicating that initially the concentration of the dopant atoms in the wafer is much lower than the concentration in the source, so $C(x,0) = 0$. As the name constant surface concentration suggests, the first boundary condition is satisfied by $C(0,t) = C_s$, and since the point of deposition is to deposit dopant at the *surface* of the wafer, $C(\infty,t) = 0$. The general solution of Fick's Law given these conditions is

$$C\left(x,t\right) = C_s \; erfc\left(\frac{x}{2\sqrt{Dt}}\right) \tag{3.3}$$

with the total number of dopant atoms per unit area given by

$$Q\left(t\right) = \int_0^\infty C\left(x,t\right) dx \tag{3.4}$$

and by substituting Equation 3.3 into 3.4 we find

$$Q\left(t\right) = \frac{2}{\sqrt{\pi}} C_s \sqrt{Dt} \cong 1.13\ C_s \sqrt{Dt} \tag{3.5}$$

Limited Source

For drive-in the initial condition is the same. We also maintain the boundary condition $C(\infty, t) = 0$. The constant source boundary condition no longer holds, though. We replace it with a boundary condition that reflects the constant total dopant invariant of this model: $\int_0^\infty C(x,t) = S$, where S is the total amount of dopant. The solution of the diffusion equation that satisfies these conditions is given by

$$C\left(x,t\right) = \frac{S}{\sqrt{\pi Dt}}\, e^{\left(-\frac{x^2}{4Dt}\right)} \tag{3.6}$$

3.1.3 Evaluation

After each diffusion or drive-in stage, the wafer can be evaluated for resultant electrical characteristics. The three primary evaluation metrics are junction depth, sheet resistance, and dopant profile. For this lab, we only measure the sheet resistance, and we determine junction depth based on two equations for R_s. Using that ρ is resistivity measured by the four-point probe, it becomes possible to calculate the junction depth x_j by

$$R_s = \frac{1}{q \int_o^j \mu C\left(x\right) dx} \tag{3.7}$$

where the carrier mobility μ is dependent on $C(x)$. When $C(x)$ is constant, equation 3.7 is simplified to equation 3.8.

$$R_s = \frac{\rho}{x_j} \tag{3.8}$$

When $C(x)$ is not constant, such as after diffusion, μ must be calculated using equation 3.9[1]

$$\mu = \mu_1 + \frac{\mu_2 - \mu_1}{1 + \left(\frac{N}{N_{\text{ref}}}\right)^{\alpha}} \tag{3.9}$$

Application of equation 3.9 requires several values, which are dependent on whether the carrier is n-type or p-type. The n-type values are shown in Figure 3.1, and the p-type values are shown in Figure 3.2. N is actually $C(x)$ for the depth being examined. That is, the resistance varies as the depth changes.

Symbol	Value	Units
μ_1	92	cm^2 / Vs
μ_2	1380	cm^2 / Vs
N_{ref}	1.3×10^{17}	cm^{-3}
α	0.91	

Figure 3.1: **Constants used for n-type carriers in solving $\mu(x)$.**

Symbol	Value	Units
μ_1	47.7	cm^2 / Vs
μ_2	495	cm^2 / Vs
N_{ref}	6.3×10^{13}	cm^{-3}
α	0.76	

Figure 3.2: **Constants used for p-type carriers in solving $\mu(x)$.**

The solution for equation 3.9 can be found in multiple ways. The method used in this report is numerical integration approximation via spreadsheet calculations. Tables were generated for $\mu\left(N\left(x\right)\right)$ for both n-type and p-type, and were applied to $N(x)$ calculation approximations as x was varied from our calculated junction depth to the surface.

[1] from Introduction to VLSI Circuits and Systems, Uyemura. Note that this is an empirically derived equation.

3.2 Procedure and Observations

In this report, we describe the actual steps involved and measurements made from four diffusion steps: (1) pre-deposition of P-well, (2) P-well drive-in, (3) P^+-source/drain deposition, (4) field oxidation, (5) N^+-source/drain deposition, and (6) gate oxidation. In each stage, the basic process is identical. The only process variations are in temperature and time. For the drive-in stage, an additional deviation is that no sources are loaded into the diffusion furnaces. Likewise, during field oxidation and oxidation process discussed previously is executed, which uses different furnaces and no sources.

3.2.1 Basic Process

The basic sequence of steps for every diffusion step is to first RCA clean the wafers, as discussed in earlier reports. This stage will remove particle build up and any thin oxides which may have grown while in storage or during handling. After RCA cleaning, the wafers are loaded into an automated N_2 drying machine, the PSC-101 by Semtool. After the drying is complete, the wafers are loaded into a quartz carrier. The carrier has several notched slots for holding wafers. The wafers are loaded into the carrier side-by-side with a source wafer "sandwiched" between two target wafers (*i.e.*, <(Wafer,Dopant,Wafer),(Wafer,Dopant,Wafer),...>). Each (Wafer,Dopant,Wafer) group is inserted such that the *processed* side is facing the Dopant. This method is called the *solid neighbor* arrangement.

Once loaded into the quartz carrier, the carrier is placed inside the furnace chamber but not in the quartz furnace tube. The furnace is idling at $800^\circ\mathrm{C}$ during this period. After sitting inside the chamber for two minutes, the carrier is then loaded into the mouth of the quartz furnace tube. The tube is then closed, and the chamber sealed. After five minutes of "warm-up" time, the carrier is pushed into the center of the furnace. After re-sealing the tube and chamber, the furnace is ramped to the target temperature at a rate of $20^\circ\mathrm{C}$ per minute. When the furnace reaches the target temperature, the gas flow is enabled by an Electromantle controlled source (no model number was discernible). The diffusion process then occurs for the designated time. After the diffusion time is complete, the process is executed in reverse to eventually extract the wafer carrier. The slow process of heating the wafers before inserting into the center of the furnace is to prevent damage or destruction of the wafers due to thermal stress.

During the diffusion process, a thin oxide typically develops on the wafer. After the diffusion period is over, this oxide is then stripped in an HF bath for a short period, before returning the wafers to storage. Every stage used a 1:50 HF etch for two minutes with no agitation.

For each P-type deposition stage, the source material was a 2 inch diameter Techneglas GS126 BoronPlus solid source wafer. For the N-type deposition stage, we used a 2 inch diameter Techneglas TP250 PhosPlus solid source wafer. Every stage had the same N_2 gas flow rate at 1000 sccm. The variations in timing and temperature are shown in Table 3.1.

Stage	Time [m]	Temp [$^\circ$C]
P-well predep	15	850
P-well drivein	360	1175
P^+src/drain	30	935
Field Ox	150	1100
N^+src/drain	15	890
Gate Ox	55	1000

Table 3.1: **Time and temperature settings for each stage of the diffusions performed.**

3.2.2 Characterization

In this section, we construct analytical models for the junction depth, x_j, and resistivity, ρ, for each diffusion stage. After each process, a Veeco Four-Point Probe, Model FPP-100, was used to measure R_s. This value of R_s is used

according to the equations in section 3.1.3. Table 3.2 lists all the constants used in computations for each stage, and where they were obtained from.

Stage	Symbol	Meaning	Value	Units	Source
P-well predep	C_s	Surface Concentration	4.5×10^{19}	cm^{-3}	Fig 2.18 in text
	D_p	Diffusivity pre-dep	8×10^{-16}	cm^2/s	Fig 6.4 in text
P-well drivein	D_d	Diffusivity drive-in	8×10^{-13}	cm^2/s	Fig 6.4 in text
P^+src/drain	C_s	Surface Concentration	7×10^{19}	cm^{-3}	Fig 2.18 in text
	D_p	Diffusivity pre-dep	6.5×10^{-15}	cm^2/s	Fig 6.4 in text
Field Ox	D_d	Diffusivity drive-in	1.2×10^{-13}	cm^2/s	Fig 6.4 in text
N^+src/drain	C_s	Surface Concentration	6.5×10^{20}	cm^{-3}	Fig 2.18 in text
	D_p	Diffusivity pre-dep	1.1×10^{-15}	cm^2/s	Fig 6.4 in text
Gate Ox	D_d	Diffusivity drive-in	1.2×10^{-14}	cm^2/s	Fig 6.4 in text

Table 3.2: **Constants used in the diffusion analysis.**

The other constant used in the following analysis is the background concentration of the wafer substrate (C_B). For silicon, C_B is typically 10^{14} - 10^{16}. We used 10^{15} for our calculations.

To distinguish between the predeposition and drive-in stages of the doping process, we express equation 3.3 using the constants above as

$$C\left(x,t_p\right)=C_s\ erfc\left(\frac{x}{2\sqrt{D_pt_p}}\right) \tag{3.10}$$

Similarly, we express equation 3.6 as

$$Q\left(t_p\right)=\frac{2}{\sqrt{\pi}}C_s\ \sqrt{D_pt_p}\ \cong\ 1.13\ C_s\ \sqrt{D_pt_p} \tag{3.11}$$

and the revised equation for the limited source drive-in model is

$$C\left(x,t_d\right)=\frac{Q(t_p)}{\sqrt{\pi D_dt_d}}\ e^{\left(-\frac{x^2}{4D_dt_d}\right)} \tag{3.12}$$

Since the total number of dopant atoms at drive-in time is the number of atoms that were deposited in the deposition stage, $S=Q(t_p)$. All calculations were performed with UN*X MapleV5, a sophisticated and programmable mathematic system, on a SunBlade 100. Numerical integration for both $N\left(x\right)$ and $\mu\left(N\left(x\right)\right)$ was performed in OpenOffice 1.1.0 on the same platform.

3.2.3 Results

Results are obtained by analytical model, computer simulation, and empirical tests. Comparison and analysis of each is presented in the following subsections. Possible causes for error are discussed in section 5.3.

Computer Model

In addition to the analytical work that follows, we performed a computer simulation using Silvaco's SSuprem3. The simulating machine was a single-processor Pentium3 running Windows 2000.

Short scripts that describe each stage of operation were run to generate models of the outcome. The scripts used are shown in tables in each stage of analysis. For every script, not shown are the "extract" statements used to have SSuprem3 report the junction depth and sheet resistivity. These statements are as follows:

print layers chemical concentration phosphorus boron net

extract name="xj" xj material="Silicon" mat.occno=1 x.val=0 junc.occno=1

extract name="rho" sheet.res material="Silicon" mat.occno=1 x.val=0 region.occno=1
tonyplot

P-well predeposition

An analytical evaluation of the resulting junction depth can be calculated by setting equation 3.10 equal to the background concentration.

$$\begin{aligned} C\left(x_j, t_p\right) = C_s \text{ erfc}\left(\frac{x_j}{2\sqrt{D_p t_p}}\right) &= C_B \\ 4.5 \times 10^{19} \cdot \text{erfc}\left(\frac{x}{2\sqrt{8 \times 10^{-16} \times 15 \times 60}}\right) &= 1 \times 10^{15} \\ x_j &= 5.0 \times 10^{-2}\ \mu m \end{aligned}$$

The total dopant introduced to the wafer is obtained by equation 3.11.

$$\begin{aligned} Q\left(t_p\right) &= \frac{2}{\sqrt{\pi}} C_s \sqrt{D_p t_p} \cong 1.13\ C_s \sqrt{D_p t_p} \\ Q\left(t_p\right) &\cong 1.13 \cdot 4.5 \times 10^{19} \times \sqrt{8 \times 10^{-16} \cdot 15 \times 60} \\ Q\left(t_p\right) &\cong 4.3 \times 10^{13}\ \frac{carriers}{cm^3} \end{aligned} \tag{3.13}$$

The SSuprem3 code fragment to simulate this stage is shown in the figure 3.3. The graphical output from SSuprem3 is shown in figure 3.4.

```
title P-well Predeposition
initialize <100> si phos conc=1e15 thick=7
comment B predep for p-well (inf src)
diffusion time=15 temp=850 boron solidsol
```

Figure 3.3: **SSuprem3 P-well predep code.**

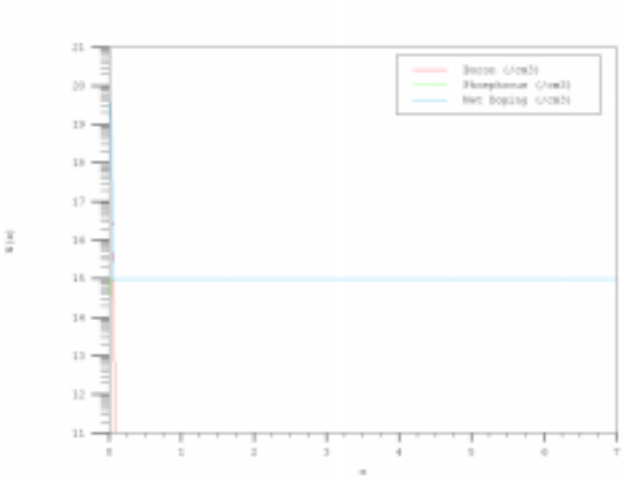

Figure 3.4: **SSuprem3 P-well predep result.**

The result from SSuprem3's `extract` statement indicated a junction depth x_j of 0.062 μm, compared to our analytical 0.050 μm. This represents an error in our calculations of approximately 23%.

Four-point probe readings of 2.46$k\Omega$, 2.52$k\Omega$, 4.40$k\Omega$, 3.90$k\Omega$, and 2.42$k\Omega$ were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 3107 Ω is obtained. SSuprem3 reported a sheet resistance of 1309 Ω, while our numerical integration resulted in 3050 Ω. Our error compared to the four-point probe is 2%, while our error compared to SSuprem3 is approximately 58%.

P-well drive-in

The drive-in junction depth is calculated by equation 3.12. The total dopant concentration S is the result from calculation of total carriers in the pre-dep stage.

$$\begin{aligned} C\left(x,t_d\right) &= \frac{S}{\sqrt{\pi D_d t_d}}\, e^{\left(-\frac{x^2}{4D_d t_d}\right)} \\ C\left(x,t_d\right) &= \frac{4.3\times10^{13}}{\sqrt{\pi\cdot 8\times10^{-13}\cdot 360\times 60}}\cdot e^{\left(-\frac{x^2}{4\cdot 8\times10^{-13}\cdot 360\times 60}\right)} \\ x_j &= 6.0\mu m \end{aligned} \tag{3.14}$$

The SSuprem3 code fragment to simulate this stage is shown in figure 3.5. The graphical output from SSuprem3 is shown in figure 3.6.

```
title P-well Predeposition
initialize <100> si phos conc=1e15 thick=7
comment B predep for p-well (inf src)
diffusion time=15 temp=850 boron solidsol
comment B drive-in (lim src)
diffusion time=360 temp=1175
```

Figure 3.5: **SSuprem3 P-well drivein code.**

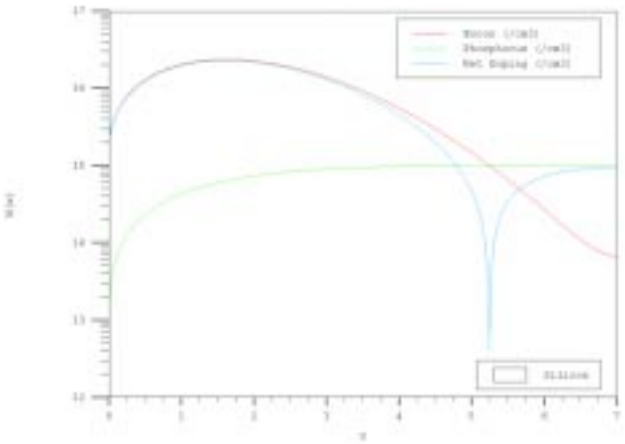

Figure 3.6: **SSuprem3 P-well drivein result.**

The result from SSuprem3's extraction statement indicated a junction depth x_j of 5.25 μm, versus our 6.0 μm. This represents (assuming SSuprem3 is accurate) an error in our calculations of approximately 12%.

Four-point probe readings of 1.11$k\Omega$, 899Ω, 1.30$k\Omega$, 1.94$k\Omega$, and 746Ω were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 1188 Ω is obtained. SSuprem3 reported a sheet resistance of 2313 Ω, while our numerical integration resulted in 749 Ω. Our error compared to the four-point probe is 59%, while our error compared to SSuprem3 is approximately 300%.

P source/drain

An analytical evaluation of the resulting junction depth can be calculated by setting equation 3.10 equal to the background concentration.

$$\begin{aligned} C\left(x,t_p\right) = C_s\, erfc\left(\frac{x}{2\sqrt{D_p t_p}}\right) &= C_{bkgnd} \\ 7\times10^{19}\cdot erfc(\frac{x}{2\sqrt{6.5\times10^{-15}\cdot 30\times 60}}) &= 1\times10^{15} \\ x_j &= 2.1\times10^{-1}\,\mu m \end{aligned}$$

The total dopant introduction to the wafer is obtained by equation 3.11.

$$\begin{aligned} Q(t_p) &= \frac{2}{\sqrt{\pi}} C_s \sqrt{D_p t_p} \cong 1.13\ C_s \sqrt{D_p t_p} \\ Q(t_p) &\cong 1.13 \cdot 7 \times 10^{19} \times \sqrt{6.5 \times 10^{-15} \cdot 30 \times 60} \\ Q(t_p) &\cong 2.7 \times 10^{14}\ \frac{carriers}{cm^3} \end{aligned} \tag{3.15}$$

The SSuprem3 code fragment to simulate this stage is shown in figure 3.7. The graphical output from SSuprem3 is shown in figure 3.8.

```
title P+ Src/Drn predep
initialize <100> si phos conc=1e15 thick=7
comment B predep (inf src)
diffusion time=30 temp=935 boron solidsol
```

Figure 3.7: **SSuprem3 P-src/drn predep code.**

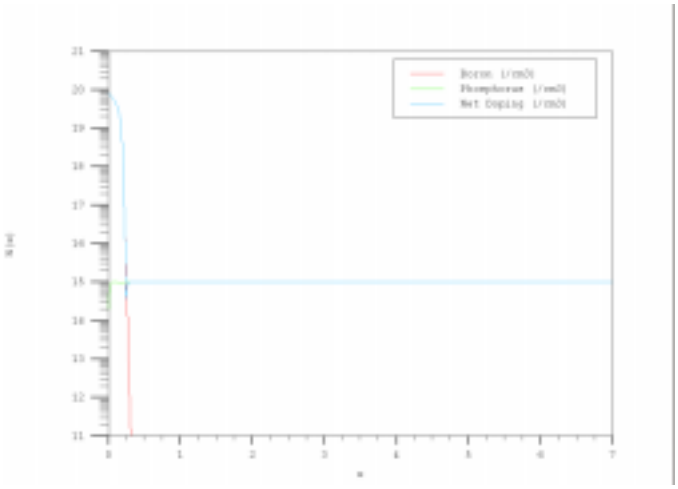

Figure 3.8: **SSuprem3 P-src/drn predep result.**

The result from SSuprem3's extraction statement indicated a junction depth x_j of 0.25 μm. This represents (assuming SSuprem3 is accurate) an error in our calculations of approximately 20%.

Four-point probe readings of 40.6Ω, 40.3Ω, 40.5Ω, 39.9Ω, and 49.9Ω were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 42 Ω is obtained. SSuprem3 reported a sheet resistance of 130 Ω, while our numerical integration resulted in 483 Ω. Our error compared to the four-point probe is 1000%, while our error compared to SSuprem3 is approximately 350%.

Field Oxidation

During the growth of a field oxide prior to N+ diffusion, the high temperatures caused a drive-in of the P source/drain stage. The drive-in junction depth is calculated by equation 3.12. The total dopant concentration S is the result from calculation of total carriers in the pre-dep stage for the source/drain phase.

$$\begin{aligned} C(x,t_d) &= \frac{S}{\sqrt{\pi D_d t_d}}\, e^{\left(-\frac{x^2}{4D_d t_d}\right)} \\ C(x,t_d) &= \frac{2.7 \times 10^{14}}{\sqrt{\pi \cdot 1.2 \times 10^{-13} \cdot 150 \times 60}} \cdot e^{\left(-\frac{x^2}{4 \cdot 1.2 \times 10^{-13} \cdot 150 \times 60}\right)} \\ x_j &= 1.9 \mu m \end{aligned}$$

The SSuprem3 code fragment to simulate this stage is shown in figure 3.9. The graphical output from SSuprem3 is shown in figure 3.10.

```
title P+ Src/Drn predep
initialize <100> si phos conc=1e15 thick=7
comment B predep (inf src)
diffusion time=30 temp=935 boron solidsol
comment B src/drn drivein (lim src)
diffusion time=150 temp=1100
```

Figure 3.9: **SSuprem3 P-src/drn drivein code.**

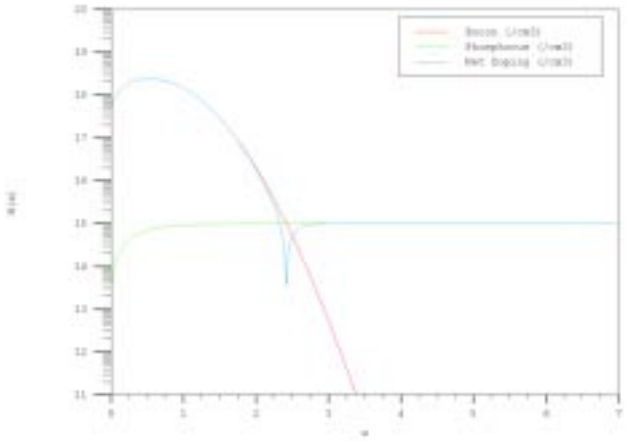

Figure 3.10: **SSuprem3 P-src/drn drivein result.**

The result from SSuprem3's extraction statement indicated a junction depth x_j of 2.41 μm. This represents (assuming SSuprem3 is accurate) an error in our calculations of approximately 27%.

Four-point probe readings of 180.3Ω, 196.1Ω, 250Ω, 166.1Ω, and 231Ω were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 202 Ω is obtained. SSuprem3 reported a sheet resistance of 228 Ω, while our numerical integration resulted in 354 Ω. Our error compared to the four-point probe is 57%, while our error compared to SSuprem3 is approximately 46%.

N source/drain

The N source/drain calculation is somewhat complicated compared to the others. Here, the background concentration is not the background concentration of the wafer. Rather, it is the background concentration of the P-well which we will be diffusing into. The P-well, however, has been diffused multiple times at multiple temperatures. The basic equations presented in section 3.1.3 are unable to handle this problem. In order to determine the proper background concentration, we used Maple, as described previously, to plot both the P-well concentration over multiple diffusions, as shown in Figure 3.11, and the N-source/drain predep, as shown in Figure 3.12. The junction of these two graphs is magnified and shown in Figure 3.13. From this, we can see both the analytical background concentration and the expected junction depth. Thus, the background concentration at the surface of the P-well is approximately 1.85×10^{17}, and $x_j = 0.051 \mu m$. (From analytical models, we observed a shift of less than 1% with the junction depth during multiple diffusion cycles.)

An analytical evaluation of the resulting junction depth can be calculated by setting equation 3.10 equal to the background concentration.

$$\begin{aligned} C\left(x, t_p\right) = C_s \; erfc\left(\frac{x}{2\sqrt{D_p t_p}}\right) &= C_{bkgnd} \\ 6.5 \times 10^{20} \cdot erfc\left(\frac{x}{2\sqrt{1.1 \times 10^{-15} \cdot 15 \times 60}}\right) &= 1.85 \times 10^{17} \\ x_j &= 5.1 \times 10^{-1} \; \mu m \end{aligned}$$

The total dopant introduction to the wafer is obtained by equation 3.11.

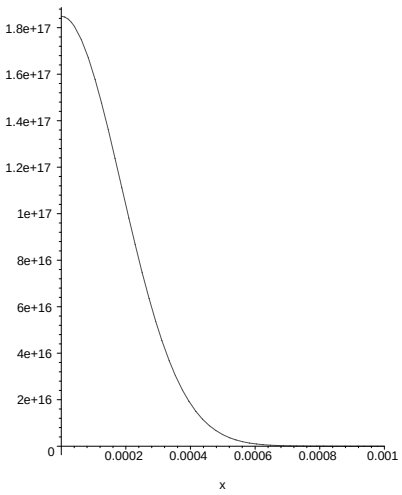

Figure 3.11: **Maple plot of x versus background concentration for multiple P-well drive-ins.**

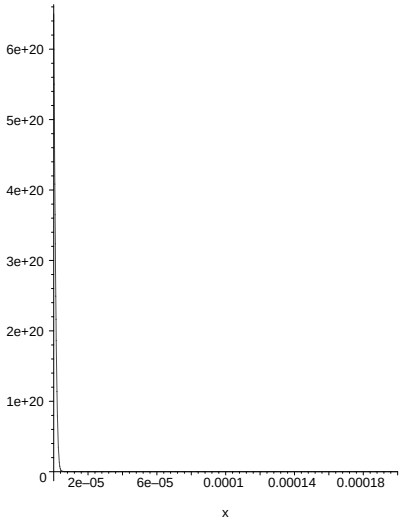

Figure 3.12: **Maple plot of x versus background concentration for N-source/drain predeposition.**

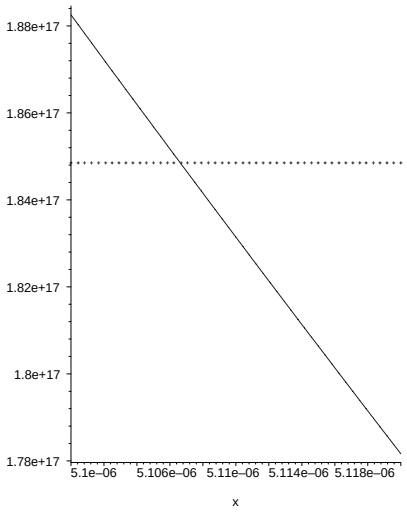

Figure 3.13: **Crossing of the N-source/drain pre-dep and the multiple P-well drive-ins from Maple.**

$$
\begin{aligned}
Q\left(t_p\right) &= \frac{2}{\sqrt{\pi}} C_s \sqrt{D_p t_p} \cong 1.13\ C_s \sqrt{D_p t_p} \\
Q\left(t_p\right) &\cong 1.13 \cdot 6.5 \times 10^{20} \times \sqrt{1.1 \times 10^{-15} \cdot 15 \times 60} \\
Q\left(t_p\right) &\cong 7.3 \times 10^{14}\ \frac{carriers}{cm^3}
\end{aligned}
\tag{3.16}
$$

The SSuprem3 code fragment to simulate this stage is shown in figure 3.14. The graphical output from SSuprem3 is shown in figure 3.15.

```
title N+ Src/Drn predep
initialize <100> si phos conc=1e15 thick=7
comment B predep p-well (inf src)
diffusion time=15 temp=850 boron solidsol
comment B drivein p-well (lim src)
diffusion time=350 temp=1175
comment P predep N+ src/drn (inf src)
diffusion time=15 temp=890 phos solidsol
```

Figure 3.14: **SSuprem3 N+ src/drn predep code.**

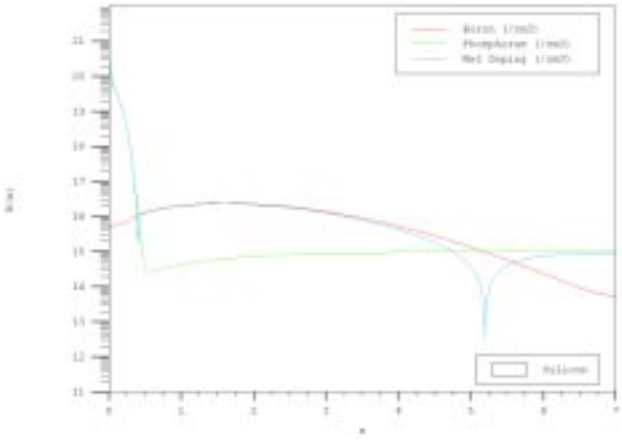

Figure 3.15: **SSuprem3 N+ src/drn predep result.**

The result from SSuprem3's extraction statement indicated a junction depth x_j of 0.34 μm. This represents (assuming SSuprem3 is accurate) an error in our calculations of approximately 33%.

Four-point probe readings of 22.8Ω, 21.2Ω, 20.6Ω, 21.8Ω, and 20.7Ω were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 21

Ω is obtained. SSuprem3 reported a sheet resistance of 67 Ω, while our numerical integration resulted in 98 Ω. Our error compared to the four-point probe is 400%, while our error compared to SSuprem3 is approximately 31%.

Gate Oxide

During the growth of the gate oxide after N+ diffusion, the high temperatures caused a drive-in of the N+ source/drain stage. The drive-in junction depth is calculated by equation 3.12. The total dopant concentration S is the result from calculation of total carriers in the pre-dep stage for the source/drain phase.

$$\begin{aligned} C(x,t_d) &= \frac{S}{\sqrt{\pi D_d t_d}}\, e^{\left(-\frac{x^2}{4D_d t_d}\right)} \\ C(x,t_d) &= \frac{7.3\times 10^{14}}{\sqrt{\pi\cdot 1.2\times 10^{-14}\cdot 55\times 60}}\cdot e^{\left(-\frac{x^2}{4\cdot 1.2\times 10^{-14}\cdot 55\times 60}\right)} \\ x_j &= 0.3\mu m \end{aligned}$$

The SSuprem3 code fragment to simulate this stage is shown in table 3.16. The graphical output from SSuprem3 is shown in figure 3.17.

```
title N+ Src/Drn predep
initialize <100> si phos conc=1e15 thick=7
comment B predep p-well (inf src)
diffusion time=15 temp=850 boron solidsol
comment B drivein p-well (lim src)
diffusion time=350 temp=1175
comment P predep N src/drn (inf src)
diffusion time=15 temp=890 phos solidsol
comment P drivein N src/drn (lim src)
diffusion time=55 temp=1000
```

Figure 3.16: **SSuprem3 N+ src/drn predep code.**

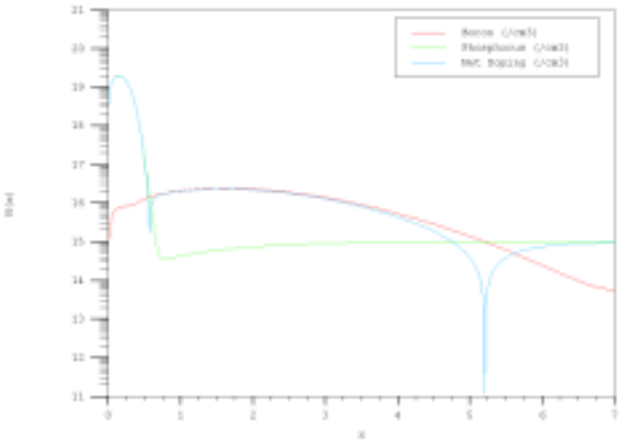

Figure 3.17: **SSuprem3 N+ src/drn predep result.**

The result from SSuprem3's extraction statement indicated a junction depth x_j of 0.573 μm. This represents (assuming SSuprem3 is accurate) an error in our calculations of approximately 33%.

Four-point probe readings of 9.54Ω, 9.30Ω, 8.34Ω, 7.01Ω, and 6.96Ω were measured on a control wafer. Using the average value from these readings, and multiplying by the correction factor 0.991, a value of approximately 8.23Ω is obtained. SSuprem3 reported a sheet resistance of 115 Ω, while our numerical integration resulted in 96 Ω. Our error compared to the four-point probe is 1100%, while our error compared to SSuprem3 is approximately 9%.

3.3 Conclusion

Our analytical models had errors in the range of 5% to 1000%. When evaluating the error based solely on our analytical model when compared to SUPREM, our error was in the range of 5-200%. The errors due to our usage of the Irvin's curve plots was in the range of 20-1000%.

There are a wide variety of sources for error in our calculations, mentioned below in brief format. More likely sources of error are simplifications in the models used, and our inherent lack of understanding of why the model works in the first place. More careful study of the models constructed could indicate that certain assumptions (such as concentration gradients) were incorrect, and the wrong equations were used.

Another question is whether we programmed SSuprem3 correctly. While SSuprem3 has a very rudimentary command set, we found unexpected behavior if we used the "obvious" commands. For example, in drive-in diffusion, by specifying "boron inert" on the diffusion command line, SSuprem3 treated this as a solid source diffusion. After much frustration and experimentation we discovered the problem. Similar subtle bugs between an intuitive interface language and the actual demands of SSuprem3 may be present.

Possible sources of error include:

1. Ionic contaminants: *Impure compounds or furnace contaminants*
2. Flow baffle: *Wafer placement baffles gas flow in furnace*
3. Uneven Sources: *Repeat-use sources may be breaking down*
4. Uneven Flow: *Carrier design reduces gas flow at bottom*
5. Reverse diffusion: *Diffusion of dopant to chamber during drive-in*
6. Hidden diffusion: *Other heating effects causing more diffusion*
7. Inaccurate Constants: *Reading graphs to approximate constants is inaccurate*

While we successfully used the numerical integration method to obtain acceptable sheet resistance results, we found that our model deviated from both the measured probe results and the SSuprem3 outputs. A graphical depiction of the error swing is shown in figure 3.18. Note how we are originally very close to the four-point probe and far from the SSuprem3 output. This inverts as the processes move on. We attribute such deviation primarily to differences between the three models: empirical current testing, detailed analytical model evaluation, and numerical approximation.

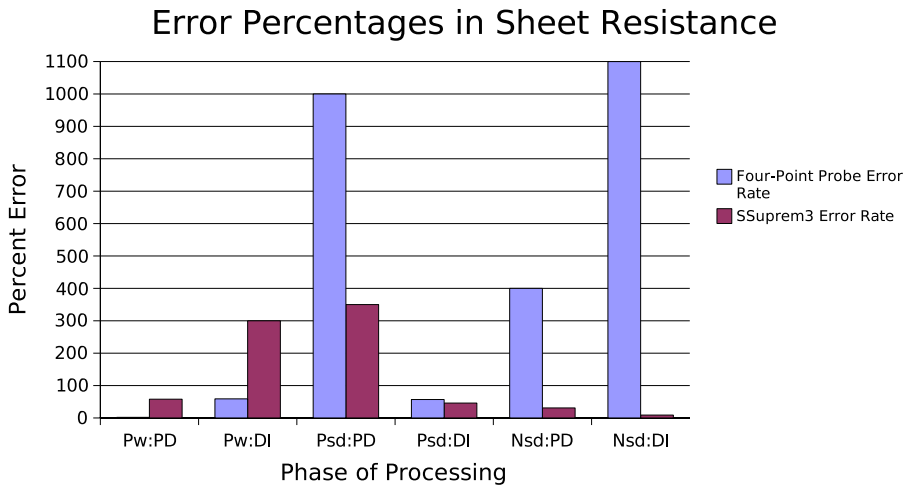

Figure 3.18: **Relative error measurements between SSuprem3, four-point probe, and numerical approximation techniques.**

This strongly suggests that the analytical models, while approximations for rule-of- thumb measurements, should not be trusted as "authoritative" and that only real measurements of process phases are meaningful. The corollary to this observation is that we need better analytical models.

Chapter 4

Metallization

4.1 Theory

The process of *metallization* generally refers to the transfer of a very thin metal film onto a wafer. The metal film is used as "wiring" between locations on the wafer. Typically, a contact hole is etched through the oxide into an N-, P-type region. The accumulating metal deposits during metallization fill in the hole, and cover the entire wafer with a solid sheet of conducting material. This sheet is then patterned with a photoresist (positive photoresist for our lab) over the regions of metal to be *kept*. After the photoresist is hard-baked, the open regions of metal film are etched away. After the metal etching is complete, the photoresist is removed, leaving only thin traces of metal running across the wafer. These traces are referred to as leads, metal lines, interconnects, or wires.

Multiple metal layers can be applied by a repeated process of metallization, oxidation, metallization, and so forth. A cross-sectional view of how the layers are organized is shown in Figure 4.1.

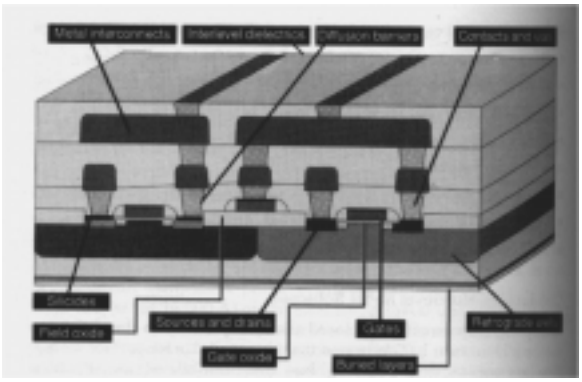

Photo from Microchip Fabrication, van Zant, 4th ed.

Figure 4.1: **Cross-section view of multiple layers during wafer processing**

There are multiple methods for metallization. The most common early mechanism was vacuum deposition, which used aluminum or gold. Multi-metal layer designs required better control and coverage, which resulted in the development of a sputtering system (PVD). Refractory metal use has led to a third method, chemical vapor deposition (CVD). In our laboratory, we only have an old-style evaporation system with select modificaitons.

4.1.1 Procedure

The sequence of operational steps for the metallization process are as follows, following contact etch and photoresist removal:

Cleaning	Light cleaning (*See Lab 1*)
Metallization	Metal film deposition
Photolith Mask 6	Protect wires to remain (*See Lab 2*)
PAN etch	Remove unwanted metal
Photoresist Removal	Strip protective PR (*See Lab 2*)
Plasma Etch	Removal any remaining PR
Sinter/Anneal	Bond metal to silicon

Metallization

In the metallization process, Aluminum is used since it has a low contact resistance with Silicon, and it doesn't tend to "wreak havoc" on the devices if it gets into unwanted areas (van Zant, Microchip Fabrication, 4th ed.). Aluminum also has a low resistivity (approx. $2.7\mu\Omega$ cm), and adheres well to Silicon.

The common setup is to have a quartz or stainless steel bell jar under high vacuum. The wafers are mounted at the top of the jar, upside down, and directly over a heating filament and boat of deposition metal. This boat, commonly made of Titanium, and with either a Titantium or Tungsten heating filament, is positioned in the center of the chamber. By carefully controlling the current to the filament, the temperature of the chamber, and the vacuum pressure, a semi-controlled rate of metal film deposition is achieved. However, the vacuum technique is rarely used in industry because it is inconsistent in how it forms metal films. Part of the problem with this technique is the inability to control a constant temperature across the heating filament, and dealing with impurities in the boat, metal source, or on the wafer surface.

PAN etch

The PAN (Phosphoric, Acetic, and Nitric acids) etch is meant to remove the unprotected Aluminum regions from the wafer. During the PAN etch, the Aluminum is "eaten" off the wafer by the acid mix. The wafer must be inspected closely to ensure that no metal is left except for what is protected by photoresist. Any stray metal could cause a short that could render an entire chip useless.

Plasma Etch

The plasma etch process involves introducing a gas into a low pressure electrically charged environment to transition the gas into a plasma form. The ionized particles of the plasma are then "aimed" at the wafer, and then react with the the target substance on the surface of the wafer, whether it is to burn it off or destructively bombard it. The Plasma etch is performed to remove any remaining hardened photoresist or minute Aluminum deposits not discernible during QC inspections.

Sinter/Anneal

The final stage of the process is to bond the metal film directly onto the wafer such that it can sustain the contact necessary for probing and device testing. Under high temperatures, the Silicon will diffuse into the Aluminum. Obviously, this process must be done exactly; if the annealing process goes on for too long, the integrity of the Aluminum wires might be compromised by too much Silicon diffusion, and if the annealing temperature is too high, the Aluminum will melt. The latter problem is compounded because the Al-Si interface exhibits *eutectic* properties, meaning that the melting point of the diffused Al and Si is lower than the melting point if either individually.

4.2 Procedure and Observations

4.2.1 Process steps

HF Dip

A brief oxide stripping is performed by HF dip. This is not a full RCA cleaning procedure, which would be overkill. Since the wafer is about to have molten metal deposited on it, the only real problem with performance would be a thin oxide layer preventing the metal from contacting those regions we wish it to contact. The HF dip solution we used was a ratio of 1:50 for 20 seconds. A thorough DI rinse and N_2 drying was performed prior to loading into the metallization machine.

Metallization

A vacuum evaporation system is used in our laboratory. This system has been modified from the original design in several important ways. First, it no longer uses an oil-based pump for reducing pressure. Due to the problems with *backstreaming*, or the injection of oil into the primary chamber, the system was switch to cryogenic pumps. Liquid N_2 is also used to speed the vacuum processing. Also, the mechanism for loading wafers was replaced with an assembly from an old sputtering machine.

A Denton 502-A machine was used. The cryogenic pumps were monitored via a Kurt J. Lesker Co. 194400 Ionization gauge controller. The rate of metal evaporation and deposition was monitored via an INFICON Deposition Controller XTC/2. Instead of using 100% Aluminum, we used 99% Aluminum mixed with 1% Silicon. Pure Aluminum will leave a "milky" metal layer, indicating multiple crystal sites and poor bonding by melting *into* the Silicon layer due to eutectic melting points by the combination of Aluminum and Silicon. By using the diluted Aluminum, we obtain better transfer of metal and better bonding without melting into the processed wafer layers.

The Aluminum used was from Angstrom Sciences, with a 2-inch cut piece of 1mm diameter. The material lot number was CC Verbal/99-11361, with a mix marked as $Al_{99}Si_1$. To ensure the Aluminum is clean prior to use, it is packed under Argon at the factory, and on arrival dipped in an HF bath for 20+ seconds. After the HF bath, it is long-term stored in ethanol. This prevents oxidation, dirt or particle buildup, etc.

A distance of approximately 12-inches existed between our wafers and the metal source. A Titanium boat and filament were used, with the boat and filament being replaced after each metallization process. A target rate of 3-4 Å/sec was desired, using a vacuum of 3×10^{-6}torr. Our target thickness was 3500 Å.

During metallization, a control wafer was inserted that had a partial covering of a slide. After metallization, this slide was removed on the control wafer and a Tencor Alpha-Step was used to measure the actual thickness of the metal film deposited. The two-piece Alpha-Step used a measuring component (Model 10-00020) with a 45-X visual magnification for placement of the probe. The logging component (Model 10-00030) was set to record on graph paper in units of 10kÅ.

Two measurement runs were performed. In each case, the probe tip was placed on the metal-free region, and then automatically dragged across the metal film. The height change was recorded on the graph paper, which can be seen for each run in Figures 4.2 and 4.3. The final metal thickness was approximately 3700 Å.

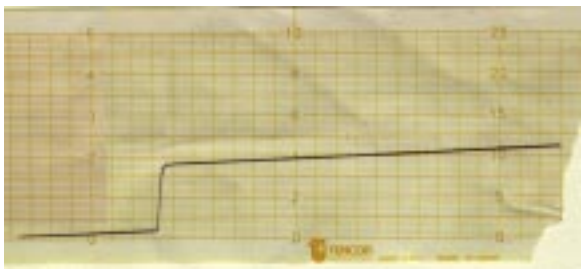

Figure 4.2: **First reading by Tencor Alpha-Step**

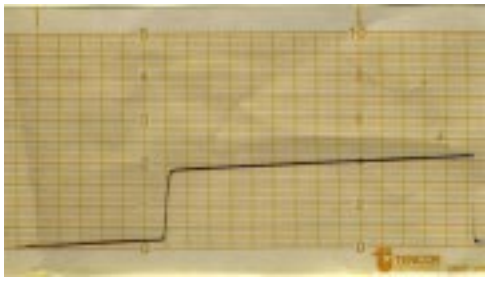

Figure 4.3: **Second reading by Tencor Alpha-Step**

Photolith Mask 6: Lightfield

Following the metallization, a final photolithography step takes place. Given that we only want to keep a few small traces of metal on each die, this photolith mask is mostly dark and extra care must be taken with alignment. Unlike prior steps, dehydration occurred at 120°C to avoid over-cooking. The primer and photoresist were applied with standard values – 10 sec dwell of Shipley HMDS primer and a 30-second spin of Shipley 1813 positive photoresist. Soft baking was for 5 minutes at 95°C , followed by exposure with a 200W bulb power for 35 seconds for an exposure energy of 6mW. After exposure, the wafers were placed in Shipley MF319 developer for 2 minutes, DI rinsed, dried, and then QC inspected.

Note: This QC phase should show target regions as yellow, with aluminum on the outside.

PAN Etch

The PAN etch was applied for 1 minute and 20 seconds at 50-55 °C . The process is a syrup-like compound due to the presence of phosphoric acid, and during DI rinse, the rinse should continue until all signs of the "syrupy" nature are gone. This is typically 20+ seconds.

Note: The PAN etch should "sizzle" like Alka-Seltzer, which indicates a good metal layer. If the aluminum lifts off in a sheet like aluminum foil, a bad metal deposition took place and the entire metallization process must be done over.

Following the PAN etch and rinse, another drying cyle is performed so that the wafers can be QC checked again. During this phase, it is imperative to look for shorts in the aluminum deposited regions. It is possible that the PAN etch did not completely eat away the excess aluminum, due to variable thickness of the film, and additional PAN etching may be required.

PR removal

The protective photoresist was then removed with Shipley 1165 stripper at 70 °C for 2 minutes. A 30+ second DI rinse and drying was performed so that the wafers could be loaded for plasma etching.

Plasma Etch

The plasma etch was carried out in Technics Micro-PD Series 95 machine. High- purity O_2 gas was used, with 100 W of power at 150 mtorr pressure for 10 minutes of etch time.

Sinter/Anneal

The final stage was annealing the remaining metal onto the wafer. This step enables the metal to undergo the forces necessary during device testing without damaging the traces. The furnace used was a Lindberg Model 55342-4, with an Electromantle (no model discernible) remote controlled gas valve. The temperature was 400 °C with an N_2 gas flow rate of 1000 sccm. The total time of the stage was 30 minutes.

4.3 Conclusion

We had a desired thickness of 3500 $\mathring{A}$. Our measured thickness was approximately 3700 $\mathring{A}$. Our error was approximately 5%. We attribute this error to the very imprecise nature of vacuum evaporation.

Chapter 5

Device Characterization

5.1 Theory

After completing the many stages in wafer processing, the ultimate test of the correctness is careful examination and characterization of the devices manufactured. Not only is characterization necessary to verify that the development process worked, it is also essential to understand how every parameter impacts the overall operation. Ultimately we need high characteristic repeatability (precision) and consistently close results to a desired performance ideal (accuracy).

In our case, many devices were made but only three types of devices are examined: resistors, transistors, and inverters. Multiple types and methods are required to fully evaluate these units. While many tests are destructive and require destroying a test wafer, this section focuses on non-destructive end-of-process characterization.

5.1.1 Resistors

Measurement of a resistor is accomplished by connecting leads to each end of the target resistor. Applying a voltage across these leads will cause a current flow based on the resistivity of the layer that is conducting. As the voltage is swept from 0V to some nV, a curve tracer shows the resultant graph. The voltage is typically on the x-axis while the current flow in on the y-axis. The sweep of several voltages is performed to verify that a real linear graph is the depiction. High contact resistance or leaking junctions could cause non-linear behavior, and would indicate a failure at some stage of the processing.

By applying some voltage V and obtaining a current flow I, Ohm's law provides the resistance, as in equation 5.1. While this provides an indicator as to the resistance of a given resistor, a more general form for expected resistance is based on the physical parameters due to processing. By considered the physical characteristics of the thickness of conducting layer, t, the length of the resistor line, l, and the width of the resistor line, w, we can determine the expected resistance according to equations 5.3 and 5.4 using the intrinsic resistance ρ.

$$V = IR \tag{5.1}$$

$$R = \frac{V}{I} \tag{5.2}$$

$$R_{line} = \frac{\rho \cdot l}{w \cdot t} \tag{5.3}$$

$$R_{line} = R_s \cdot \frac{l}{w} = R_s \cdot \#squares \tag{5.4}$$

5.1.2 Transistors

Measurement of transistors is accomplished by using four contact probes. Two probes are used to the *source* and *drain* connections. One probe is used to apply a gate voltage which sweeps from some voltage V_A to some voltage

V_B. The fourth probe monitors the output of the transistor as the gate voltage swings. The output of the transistor has three basic regions: non-conducting *(cutoff)*, non-saturated *(linear or transcode)*, and saturated. Each of these will be discussed in turn. A fourth characterization point is the *threshold voltage*, which is the critical voltage where the non-conducting FET begins to conduct, leading to the saturated region.

Essentially, when the gate voltage V_G is below some threshold voltage V_T, no conductance occurs. As V_G exceeds V_T, but is still close to V_T, non-saturation current flows from the source to the drain, as shown in equation 5.5. At some critical voltage above V_T, saturation is reached. This saturation voltage V_{sat} results in current from from source to drain as shown in equation 5.6. Special components of these equations are the β-ratio and the constant ε_{ox}, shown in equations 5.7 and 5.8 respectively.

$$I_D = \left(\frac{\beta}{2}\right)\left[2\cdot\left(V_{GS}-V_T\right)\cdot V_{DS}-V_{DS}^2\right] \quad (non-saturation) \tag{5.5}$$

$$I_D = \left(\frac{\beta}{2}\right)\left(V_{GS}-V_T\right)^2 \quad (saturation) \tag{5.6}$$

$$\beta = \mu\left(\frac{W}{L}\right)\left(\frac{\varepsilon_{ox}}{t_{ox}}\right) \tag{5.7}$$

$$\varepsilon_{ox} = 3.9\cdot\varepsilon_0 = 3.9\cdot 8.854\times 10^{-14} \tag{5.8}$$

To find the saturation voltage V_{sat}, we take the partial derivative of equation 5.5 and setting it equal to zero. This results in the solution shown in equation 5.9.

$$\begin{aligned} \frac{\partial I_D}{\partial V_{DS}} &= 0 \\ 2\cdot\left(V_{GS}-V_T\right)-2\cdot V_{DS_n} &= 0 \\ V_{sat} &= V_{DS_n}|_{peak\ current} \\ V_{sat} &= V_{GS}-V_T \end{aligned} \tag{5.9}$$

The only substantial difference between analysis of n-FETs and p-FETS is that for p-FETs we substitute V_{SG} for V_{GS}, V_{SD} for V_{DS}, and $|V_T|$ for V_T. Note that there will be two different β values – β_n and β_p – and correspondingly different μ values – μ_n and μ_p.

A complete analysis would show that I_D continues to increase even in saturation, as represented by equation 5.10. The parameter λ is an empirical quantity called the *channel-length modulation (Introduction to VLSI Systems, Uyemura)*.

$$I_D = \left(\frac{\beta_n}{2}\right)\left(V_{GS}-V_T\right)^2\left[1+\lambda\left(V_{DS}-V_{sat}\right)\right] \tag{5.10}$$

For our purposes, we assume $\lambda = 0$, and we ignore this additional effect. A further characterization is the device transconductance g_m, which can be solved for by equation 5.11. The transconductance is generally a measure of the gain obtained from a transistor with respect to the intrinsic resistance to that gain from the device construction.

$$g_m = 2\cdot\left(\frac{W}{L}\right)\cdot\left(\frac{\mu\varepsilon_{ox}}{t_{ox}}\right)\cdot\left(V_G-V_T\right) \tag{5.11}$$

5.1.3 Inverters

Application of *DC analysis* will provide basic understanding between the relationship of inverter output voltage, V_{out}, to input voltage, V_{in}. For basic DC analysis, V_{in} changes very slowly while V_{out} stabilized before a reading is made (for example, by a multi-meter). The analysis will indicate what the ranges are for input and output voltages, thereby permitting the isolation of logic 0 and 1 values. The resulting relationship is typically graphed into the *voltage transfer characteristic* (VTC) curve, which plots V_{out} as a function of V_{in}. The mid-point voltage V_M of the VTC graph corresponds to $I_{D_n} = I_{D_p}$. Analytically, this results in equation 5.12.

$$V_M = \frac{V_{DD} - \left|V_{T_p}\right| + \sqrt{\frac{\beta_n}{\beta_p}} \cdot V_{T_n}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}} \tag{5.12}$$

This indicates that the mid-point voltage is really a function of the $\frac{W}{L}$ ratios for the n- and p-FETs involved in the inverter design. Symmetric inverters have equal "0" and "1" voltage ranges – that is, $V_M = \frac{1}{2}V_{DD}$.

Given these trends, the actual VTC graph can be broken down into four subregions: output high voltage V_{OH}, output low voltage V_{OL}, the input high voltage V_{IH}, and the input low voltage V_{IH}. Based on these values, the voltage noise margins for high V_{NM_H} and low V_{NM_L} characteristics can be determined. The corresponding equations that determine these values follow.

$$V_{OH} = V_{DD} \tag{5.13}$$
$$V_{OL} = 0V \tag{5.14}$$
$$V_{IH} = VTC\ slope \cong -1 \tag{5.15}$$
$$V_{IL} = VTC\ slope \cong -1 \tag{5.16}$$
$$V_{NM_H} = V_{OH} - V_{IH} \tag{5.17}$$
$$V_{NM_L} = V_{IL} - V_{OL} \tag{5.18}$$

Note that the V_{IH} and V_{IL} are defined as the points where the VTC curve has a slope equal to -1. The VTC plot will have two such points, and with the V_{in} moving from 0 to V_{DD}, the left-most point will correspond to V_{IH} while the right-most point will correspond to V_{IL}.

5.2 Procedure and Observations

The apparatus used in evaluating our wafer devices was uniformly quaint. A Tektronix 571 Curve tracer was used for most data collection, with a GPIB interface connected to a Canon K10060 BubbleJet printer. The probing station was a Signatone H150 base with four Signatone S-725-SLM probes. The probing station had a mounted Leica StereoZoom-6 photo microscope, which was also connected to a standard 12" CRT monitor.

To control the power supplies for V_{DD}, Gnd, and V_G outside of the Tektronix 571 Curve Tracer, a Tenma 72-2005 Variable DC Power Supply was used in conjunction with an HP 6255A Dual DC Power Supply. For recording of voltages outside of the curve tracer, both of a Fluke 23 III and a Fluke 23 II digital multimeter (DMM) were used.

Two wafers were used for testing, with random selection of resistors, FETs, and inverter from both wafers. The wafer IDs were #24 and #28. For readability, forced page breaks separate the analysis of resistors, transistors, and inverters.

5.2.1 Resistors

We manufactured four different types of resistors. Each resistor had the same dimensions and pads, the only variation was the material of the resistor itself. The four types were: metal, p-well, p-source/drain, and n- source/drain.

The resistance for each resistor is determined from Ohm's law, as given in alternate form (below) by equation 5.2.

$$R = \frac{V}{I} = \frac{\Delta V}{\Delta I} \tag{5.19}$$

Verification of Curve Tracer

In the following sections, results will be reported that vary from the expected. In order to reduce some sources of error, two tests were performed with the curve tracer: testing a known resistor, and testing the probe-contact-probe resistance path. The first test is to verify that the curve tracer generates an output that is believable according with varying internal R_{load} settings. The second test is to ensure that the resistance of the probes and contacts are not significant compared to the resistance of other components. The results of these tests can be seen in figures 5.1 and 5.2.

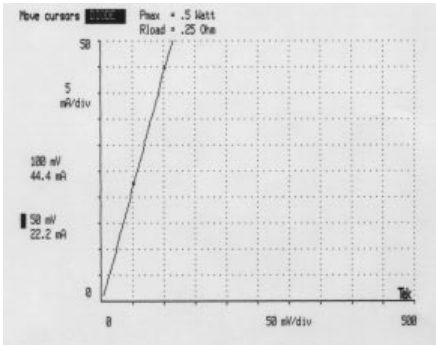

Figure 5.1: **Testing a known resistance value.**

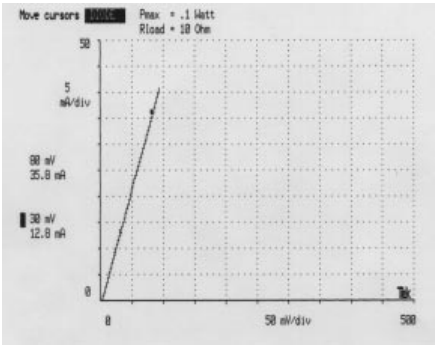

Figure 5.2: **Probe-contact-probe resistance.**

The resistor used was manufactured as 2.2 Ω with a 5% tolerance. Measurement by a BK Precision 875B high-precision analyzer indicated an actual resistance of 2.36 Ω. As can be directly calculated from equation 5.19, the resistance according to the curve tracer is 2.25 Ω. This represents an error of approximately 4.9%. As the R_{load} varies, the result varies by an additional 1-2%. This suggests a reading error in the range of (2.9, 6.9)%.

The probe-contact-probe test indicates, by equation 5.19, a resistance of 2.17 Ω. Given the error reading as above, this suggests the *real* resistance of any measure value R_X would be in the range R_X - (2.02, 2.11) Ω. For this report, we use a subtraction value of 2.05 Ω to correct measured values from the curve tracer.

Metal

Three random metal resistors from different dies on one wafer were chosen. The curve tracer plots for these three resistors can be seen in Figures 5.3, 5.4, and 5.5. These graphs all show nice linear characteristics, indicating proper processing steps and no problems. For each of the metal resistors, we calculate the resistor value as shown in table 5.1. Note that the first two resistors are in excellent agreement, with only 1% difference. These two resistors were from nearby die locations while the third metal resistor was quite far away. We attribute the error of approximately 4% to the variation in process techniques, particularly during metallization, resulting in non-uniform results.

Briefly solving this backwards, we know that each resistor is drawn with 20 μm width in five winding maze segments. Each segment has a 20 μm separation. The total length of the resistor is 10200 μm, or approximately 510 squares. By equation 5.4, and by correcting for the curve tracer error, we can determine that R_s for the metal resistor

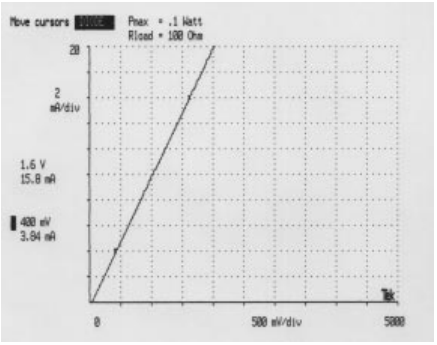

Figure 5.3: **First metal resistor.**

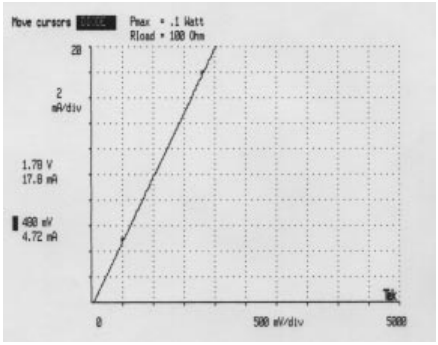

Figure 5.4: **Second metal resistor.**

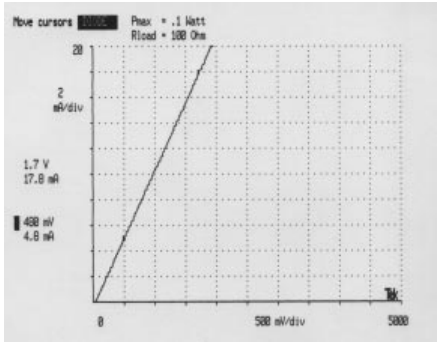

Figure 5.5: **Third metal resistor.**

ΔV_1	$=$	$1.6V$
ΔI_1	$=$	$15.86 \times 10^{-3} A$
R_1	$\cong$	$101 - 2.05 \cong 98.95\ \Omega$
ΔV_2	$=$	$1.78V$
ΔI_2	$=$	$17.8 \times 10^{-3} A$
R_2	$\cong$	$100 - 2.05 \cong 97.95\ \Omega$
ΔV_3	$=$	$1.7V$
ΔI_3	$=$	$17.8 \times 10^{-3} A$
R_3	$\cong$	$96 - 2.05 \cong 93.95\ \Omega$

Table 5.1: **Resistance calculations for three random metal resistors.**

should be in the range (0.18, 0.19) Ω. Using the standard bulk resistivity (ρ) of 100% Aluminium at 25°C to be 2.65×10^{-6} $[\Omega \cdot cm]$, we determine based on our 3700 $\mathring{A}$ (3700×10^{-8} cm) thickness an expected sheet resistance of 0.072 Ω.

However, our metal film deposit used 99% Aluminium and 1% Si, not 100% Al. The bulk resistivity (ρ) for $Si_{99}Al_1$ is in the range (2.6, 3.5) $\times 10^{-6}$ $[\Omega \cdot cm]$[1]. We also note that even though we have used a doped Al for metallization, the Al bonding will still draw some Si from the underlying surface. The advantage of using doped Al was to minimize the reduction of Si underlying the metal film, but it is not possible to *eliminate* it outright. This range of ρ corresponds to a range of expected sheet resistances between (0.070, 0.095) Ω. This represents an error of approximately 2x - 3x (200 - 300 %). While this error is large, we can account for contributions to it in many ways.

For example, we did *not* reduce the exposure time or energy on the final PR layer. The highly reflective metal surface may have over- exposed by reflection the PR layer, causing a longer etch than normal. By considering the thickness to be approximately 3700 $\mathring{A}$, we estimate the over-etching on each side to be in "units" of $0.5 \cdot 3700\mathring{A}$. Since we need to consider both sides, we simply use over-etching in units of 3700 $\mathring{A}$. This could amount to 1-5 "units" of over etching, where each unit corresponds to 0.37 μm. The reduction in width will increase the square count, such that at 2-unit error, the total square count elevates from 510 to 530. This reduces the expected R_S by 0.02, to the range (0.16, 0.18) Ω, which places our results closer to the expected values.

Additional sources of error may include imprecise PR feature transfer (such as sub-20μm resistor lines), impurities in the metal film, crystal barriers formed during metal cooling, poor pad-to-resistor attachment, etc.

P-well

Three random p-well resistors from different dies on one wafer were chosen. The curve tracer plots for these three resistors can be seen in Figures 5.6, 5.7, and 5.8. These graphs all show nice linear characteristics, indicating proper

[1] GoodFellow Industries, www.goodfellow.com

processing steps and no problems.

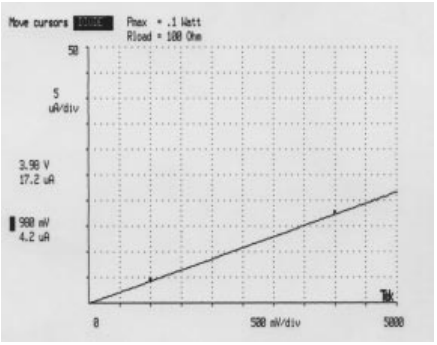

Figure 5.6: **First p-diffused resistor.**

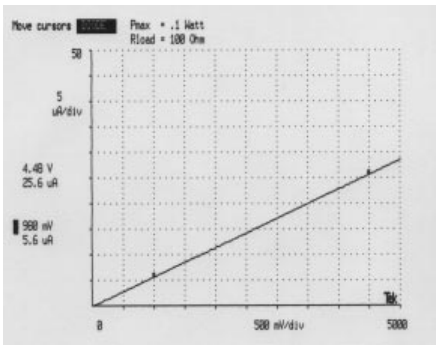

Figure 5.7: **Second p-diffused resistor.**

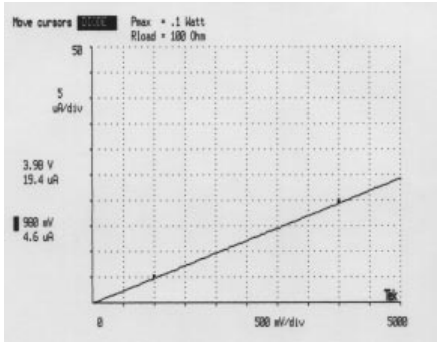

Figure 5.8: **Third p-diffused resistor.**

For each of the p-well resistors, we calculate the resistor value as shown in table 5.2.

ΔV_1	$=$	$3.98V$
ΔI_1	$=$	$17.2 \times 10^{-6} A$
R_1	$\cong$	$231 - 2.05 \cong 229\,k\Omega$
ΔV_2	$=$	$4.48V$
ΔI_2	$=$	$25.6 \times 10^{-6} A$
R_2	$\cong$	$175 - 2.05 \cong 173\,k\Omega$
ΔV_3	$=$	$3.98V$
ΔI_3	$=$	$19.4 \times 10^{-6} A$
R_3	$\cong$	$205 - 2.05 \cong 203\,k\Omega$

Table 5.2: **Resistance calculations for three random P-well resistors.**

Briefly solving this backwards, we know that each resistor is drawn with 20 μm width in five winding maze segments. Each segment has a 20 μm separation. The total length of the resistor is 10200 μm, or approximately 510 squares. By equation 5.4, and by correcting for the curve tracer, our p-diffused layer should be in the range (449, 339) Ω. Our measured values were 1.2 $k\Omega$ (4-point probe), 2.3 $k\Omega$ (SSuprem3), and 0.8 $k\Omega$ (numerical integration). Only our numerical was close to the target range, with overall results 2x to 4x of the expected. Sources for error would be similar to the metal, with the addition of extra diffusion during repeated high-temperature heatings. These extra diffusions would result in shifts in sheet resistance values for the target layer.

P-source/drain

Three random p-source/drain well resistors from different dies on one wafer were chosen. The curve tracer plots for these three resistors can be seen in Figures 5.9, 5.10, and 5.11. These graphs all show nice linear characteristics, indicating proper processing steps and no problems. For each of the p-source/drain resistors, we calculate the resistor value as shown in table 5.3.

Briefly solving this backwards, we know that each resistor is drawn with 20 μm width in five winding maze segments. Each segment has a 20 μm separation. The total length of the resistor is 10200 μm, or approximately 510 squares. By equation 5.4, and by correcting for the curve tracer, our p-source/drain layer should be in the range (191, 176) Ω. Our measured values were 202 Ω (4-point probe), 228 Ω (SSuprem3), and 384 Ω (numerical integration). The four-point probe reading was almost in range, with the other values within a factor of 2x of the target. Error sources would be as discussed previously.

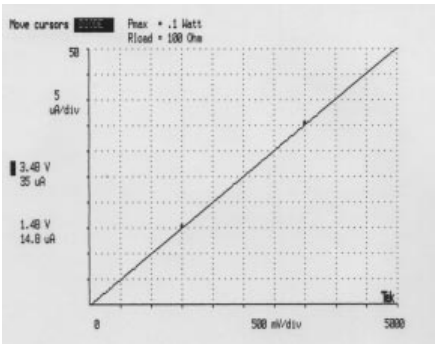

Figure 5.9: **First p-sd resistor.**

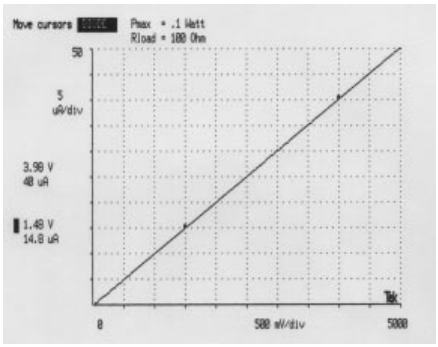

Figure 5.10: **Second p-sd resistor.**

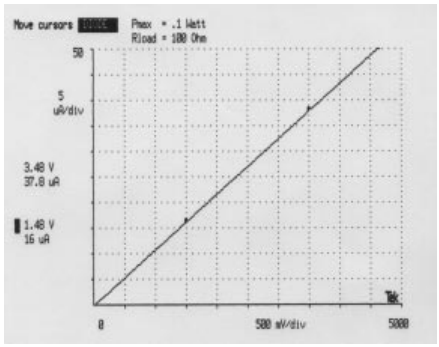

Figure 5.11: **Third p-sd resistor.**

ΔV_1	$=$	$3.48V$
ΔI_1	$=$	$35 \times 10^{-6}A$
R_1	$\cong$	$99.4 - 2.05 \cong 97.4\,k\Omega$
ΔV_2	$=$	$3.98V$
ΔI_2	$=$	$40 \times 10^{-6}A$
R_2	$\cong$	$99.5 - 2.05 \cong 97.5\,k\Omega$
ΔV_3	$=$	$3.48V$
ΔI_3	$=$	$37.8 \times 10^{-6}A$
R_3	$\cong$	$92.0 - 2.05 \cong 90.0\,k\Omega$

Table 5.3: **Resistance calculations for three random P-source/drain resistors.**

N-source/drain

Three random n-source/drain resistors from different dies on one wafer were chosen. The curve tracer plots for these three resistors can be seen in Figures 5.12, 5.13, and 5.14. These graphs all show nice linear characteristics, indicating proper processing steps and no problems.

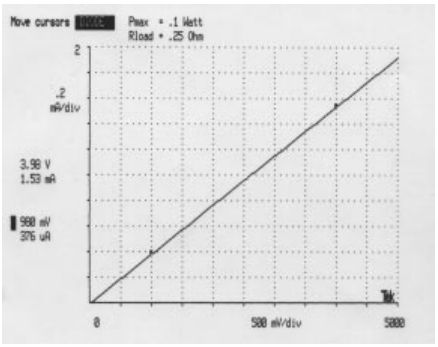

Figure 5.12: **First n-sd resistor.**

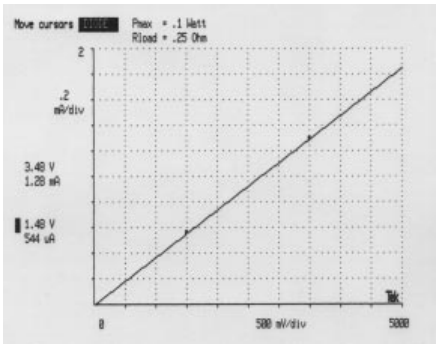

Figure 5.13: **Second n-sd resistor.**

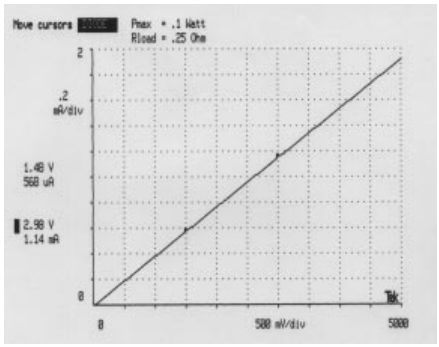

Figure 5.14: **Third n-sd resistor.**

For each of the n-source/drain resistors, we calculate the resistor value as shown in table 5.4.

Briefly solving this backwards, we know that each resistor is drawn with 20 μm width in five winding maze segments. Each segment has a 20 μm separation. The total length of the resistor is 10200 μm, or approximately 510 squares. By equation 5.4, and by correcting for the curve tracer, our n-source/drain layer should be in the range (5.05, 5.29) Ω. Our measured values were 8.23 Ω (4-point probe), 115 Ω (SSuprem3), and 96 Ω (numerical integration). Again the four-point probe was very close to the expected output, but the other two values were within an order of magnitude. This represents the worst error between the analytical models and the measured results.

ΔV_1	$=$	$1.48V$
ΔI_1	$=$	$568 \times 10^{-6} A$
R_1	$\cong$	$2.61 - 2.05 \cong 2.59\, k\Omega$
ΔV_2	$=$	$3.48V$
ΔI_2	$=$	$1280 \times 10^{-6} A$
R_2	$\cong$	$2.72 - 2.05 \cong 2.70\, k\Omega$
ΔV_3	$=$	$3.98V$
ΔI_3	$=$	$1530 \times 10^{-6} A$
R_3	$\cong$	$2.60 - 2.05 \cong 2.58\, k\Omega$

Table 5.4: **Resistance calculations for three random N-type resistors.**

Commentary

For each of these computed regions, it is interesting to note that the four-point probe and the curve tracer seldom agree on the sheet resistance. While this may suggest calibration errors between the two, it also hints at how measurements during a process are not indicative of measurements after a process. That is, measuring the resistance of the p-well after oxidation will yield a result, but that result **unconditionally will be different** as additional diffusions, oxidations, etc. occur. As we discussed briefly in our diffusion section (see section 3.2.3), further diffusions will cause a gradual shift in the characteristics. However, most of the change occurs during the initial drive-in period.

5.2.2 Transistors

Each of the transistors will be characterized by determining the threshold voltage V_T and the transconductance g_m in saturation. Each type of FET occurs in 4 varieties: 10 μm channel, 20 μm channel, 20 μm channel disconnected, and 40 μm channel sizes. Each will be considered separately. Calculation of the threshold voltage V_T is done by analysis of the graph output from the curve tracer. Rewriting equation 5.9 slightly, we observe that V_T is easily found in this manner.

$$V_T = V_G - V_{sat} \tag{5.20}$$

n-FET 10μm

The curve-tracer output for the characterization of this transistor is shown in Figure 5.15. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.5. The average of these values indicates an approximate V_T of 1.69V. The far right of figure 5.15 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

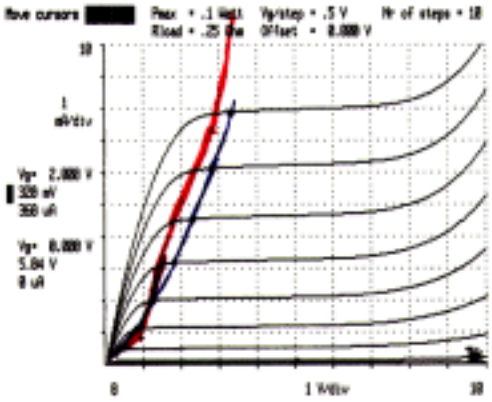

Figure 5.15: **n-FET with 10 μm channel.**

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This n-FET has a length L of 10 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.21. In figure 5.15, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{10}\right) \cdot \left(\frac{1380 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 1.69V) \end{aligned} \tag{5.21}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.6. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. We attribute the error to two sources: (1) our lack of understanding exactly where on the curve we should estimate the V_{sat} occuring; and (2) imprecise determination of V_T and other factors from graphs. There was also a wide tolerance for reading different values on the curve tracer without a corresponding value change. (For example, moving a cursor four steps up or down would change I_D without changing V_{in}. This amounted to approximately a $\pm$ 8 mA variation possible on the same V_{in}. The I_D tables are truncated at V_G of 3.0V due to the very unreliable reading of curve tracer output graphs.

V_G	V_{sat}	V_T
5.0 V	3.4 V	5.0 - 3.4 = 1.6 V
4.5 V	2.7 V	4.5 - 2.2 = 1.8 V
4.0 V	2.2 V	4.0 - 2.2 = 1.8 V
3.5 V	1.8 V	3.5 - 1.8 = 1.7 V
3.0 V	1.4 V	3.0 - 1.4 = 1.6 V
2.5 V	0.8 V	2.5 - 0.8 = 1.7 V
2.0 V	0.4 V	2.0 - 0.4 = 1.6 V

Table 5.5: V_T **for** n**-FET with 10**μm **channel.**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
5.0 V	15.0 mA	7.8 mA	48
4.5 V	10.8 mA	6.2 mA	42
4.0 V	7.30 mA	4.6 mA	37
3.5 V	4.48 mA	3.2 mA	28
3.0 V	2.35 mA	2.0 mA	15

Table 5.6: g_m **for** n**-FET with 10** μm **channel.**

n-FET 20μm

The curve-tracer output for the characterization of this transistor is shown in Figure 5.16. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.7. The average of these values indicates an approximate V_T of 2.31V. The far right of figure 5.16 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

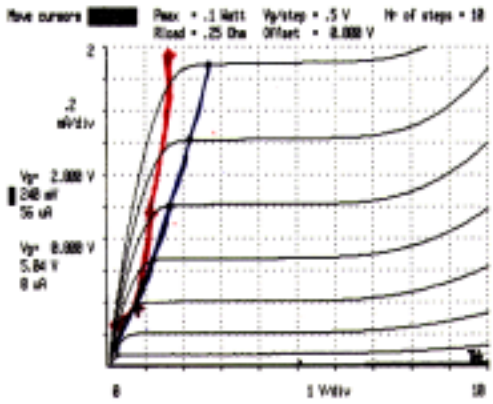

Figure 5.16: n**-FET with 20** μm **channel.**

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This n-FET has a length L of 20 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.22. In figure 5.16, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{20}\right) \cdot \left(\frac{1380 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 2.31V) \end{aligned} \tag{5.22}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.8. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

V_G	V_{sat}	V_T
5.0 V	2.4 V	5.0 - 2.4 = 2.6 V
4.5 V	2.0 V	4.5 - 2.0 = 2.5 V
4.0 V	1.5 V	4.0 - 1.5 = 2.5 V
3.5 V	1.1 V	3.5 - 1.1 = 2.4 V
3.0 V	0.8 V	3.0 - 0.8 = 2.2 V
2.5 V	0.4 V	2.5 - 0.4 = 2.1 V
2.0 V	0.1 V	2.0 - 0.1 = 1.9 V

Table 5.7: V_T **for** n**-FET with 20**μm **channel.**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
5.0 V	4.95 mA	1.88 mA	62
4.5 V	3.28 mA	1.40 mA	57
4.0 V	1.96 mA	1.00 mA	48
3.5 V	0.969 mA	0.70 mA	27
3.0 V	0.325 mA	0.36 mA	8

Table 5.8: g_m **for** n**-FET with 20** μm **channel.**

n-FET 20μm disconnect

The curve-tracer output for the characterization of this transistor is shown in Figure 5.17. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.9. The average of these values indicates an approximate V_T of 2.58V. The far right of figure 5.17 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

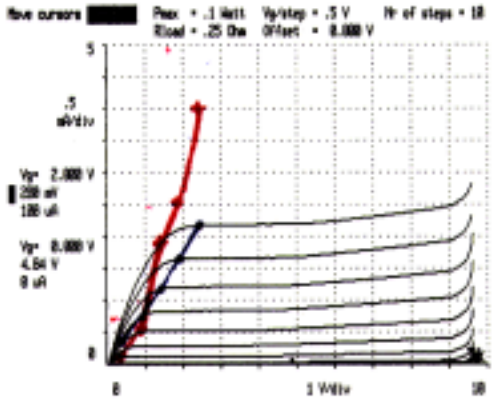

Figure 5.17: n**-FET with 20** μm **channel (disconnect).**

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This n-FET has a length L of 20 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.23. In figure 5.17, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{20}\right) \cdot \left(\frac{1380 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 2.58V) \end{aligned} \tag{5.23}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.10. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

V_G	V_{sat}	V_T
5.0 V	2.2 V	5.0 - 2.2 = 2.8 V
4.5 V	1.8 V	4.5 - 1.8 = 2.7 V
4.0 V	1.5 V	4.0 - 1.5 = 2.5 V
3.5 V	1.0 V	3.5 - 1.0 = 2.5 V
3.0 V	0.6 V	3.0 - 0.6 = 2.4 V

Table 5.9: V_T **for** n**-FET with 20**μm **channel (disconnected).**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
5.0 V	3.31 mA	2.25 mA	43
4.5 V	2.63 mA	1.6 mA	36
4.0 V	1.94 mA	1.3 mA	6
3.5 V	1.25 mA	0.75 mA	29
3.0 V	0.57 mA	0.2 mA	65

Table 5.10: g_m **for** n**-FET with 20** μm **channel (disconnected).**

n-FET 40μm

The curve-tracer output for the characterization of this transistor is shown in Figure 5.18. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.11. The average of these values indicates an approximate V_T of 2.18V. The far right of figure 5.18 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

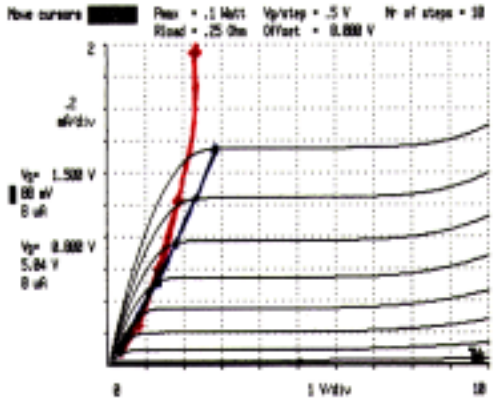

Figure 5.18: n**-FET with 40** μm **channel.**

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This n-FET has a length L of 40 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.24. In figure 5.18, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{40}\right) \cdot \left(\frac{1380 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 2.18V) \end{aligned} \tag{5.24}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.12. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

p-FET 10μm

This device, which has never worked for students, almost worked on our wafer #28. The curve tracer output is shown in figure 5.19. Due to the non-functional status of this transistor, no data was collected.

V_G	V_{sat}	V_T
5.0 V	2.6 V	5.0 - 2.6 = 2.4 V
4.5 V	2.1 V	4.5 - 2.1 = 2.4 V
4.0 V	1.7 V	4.0 - 1.7 = 2.3 V
3.5 V	1.4 V	3.5 - 1.4 = 2.1 V
3.0 V	1.0 V	3.0 - 1.0 = 2.0 V
2.5 V	0.6 V	2.5 - 0.6 = 1.9 V

Table 5.11: V_T **for** n**-FET with 40**μm **channel.**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
5.0 V	1.93 mA	1.3 mA	50
4.5 V	1.58 mA	1.0 mA	43
4.0 V	1.24 mA	0.75 mA	33
3.5 V	0.90 mA	0.5 mA	17
3.0 V	0.56 mA	0.25 mA	9

Table 5.12: g_m **for** n**-FET with 40** μm **channel.**

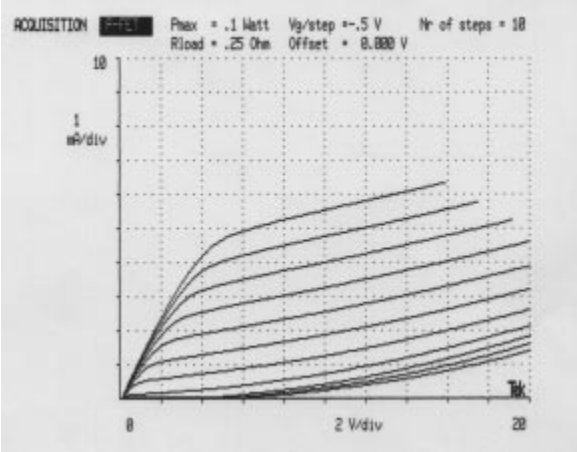

Figure 5.19: p**-FET with 10** μm **channel.**

p-FET 20μm

The curve-tracer output for the characterization of this transistor is shown in Figure 5.20. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.13. The average of these values indicates an approximate V_T of -1.41V. The far right of figure 5.20 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This p-FET has a length L of 20 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.25. In figure 5.20, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{20}\right) \cdot \left(\frac{480 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 1.41V) \end{aligned} \tag{5.25}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.14. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

p-FET 20μm disconnect

The curve-tracer output for the characterization of this transistor is shown in Figure 5.21. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.15. The average of these values indicates

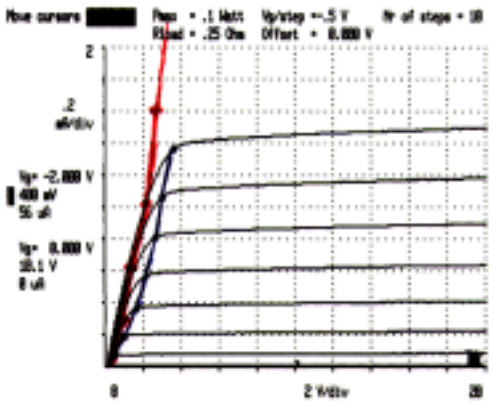

Figure 5.20: p-**FET with 20** μm **channel.**

V_G	V_{sat}	V_T
-5.0 V	-3.7 V	-5.0 - -3.7 = 1.3 V
-4.5 V	-3.1 V	-4.5 - -3.1 = 1.4 V
-4.0 V	-2.5 V	-4.0 - -2.5 = 1.5 V
-3.5 V	-2.0 V	-3.5 - -2.0 = 1.5 V
-3.0 V	-1.7 V	-3.0 - -1.7 = 1.3 V
-2.5 V	-1.0 V	-2.5 - -1.0 = 1.5 V

Table 5.13: V_T **for** p**-FET with 20**μm **channel.**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
-5.0 V	3.07 mA	1.3 mA	57
-4.5 V	2.27 mA	1.1 mA	51
-4.0 V	1.59 mA	0.8 mA	49
-3.5 V	1.05 mA	0.7 mA	25
-3.0 V	0.60 mA	0.4 mA	33

Table 5.14: g_m **for** p**-FET with 20** μm **channel.**

an approximate V_T of -1.10V. The far right of figure 5.21 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This p-FET has a length L of 20 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.23. In figure 5.21, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu \varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{20}\right) \cdot \left(\frac{480 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 1.10V) \end{aligned} \tag{5.26}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.16. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

p-FET 40μm

The curve-tracer output for the characterization of this transistor is shown in Figure 5.22. By utilizing equation 5.20, we estimate the threshold voltage V_T for each V_G as shown in table 5.17. The average of these values indicates an approximate V_T of -1.50V. The far right of figure 5.22 indicates a return to asymptotic behavior. This occurs as the voltage input begins to reach avalanche breakdown, where the input shorts through the underlying structure and current is directly passed from both source and gate.

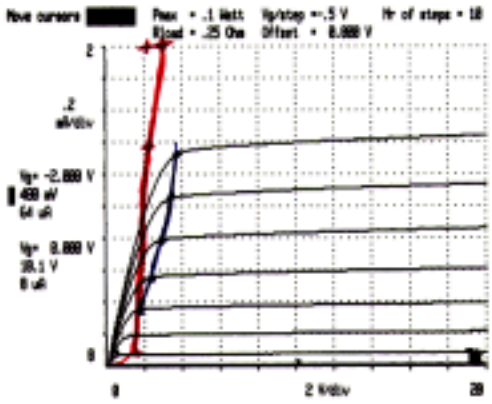

Figure 5.21: p-**FET with 20** μm **channel (disconnect).**

V_G	V_{sat}	V_T
-5.0 V	-4.0 V	-5.0 - -4.0 = -1.0 V
-4.5 V	-3.7 V	-4.5 - -3.7 = -1.2 V
-4.0 V	-3.0 V	-4.0 - -3.0 = -1.0 V
-3.5 V	-2.9 V	-3.5 - -2.3 = -1.2 V
-3.0 V	-2.3 V	-3.0 - -1.9 = -1.1 V

Table 5.15: V_T **for** p**-FET with 20**μm **channel (disconnected).**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
-5.0 V	1.85 mA	1.3 mA	64
-4.5 V	1.61 mA	1.1 mA	60
-4.0 V	1.38 mA	0.90 mA	55
-3.5 V	1.14 mA	0.55 mA	59
-3.0 V	0.90 mA	0.35 mA	58

Table 5.16: g_m **for** p**-FET with 20** μm **channel (disconnected).**

The theoretical transconductance can be calculated as shown in equation 5.11. As per the handouts from lab, each transistor has a width W of 200 μm. This n-FET has a length L of 40 μm. With a global gate oxide thickness t_{ox} of 696 $\mathring{A}$, we can find g_m. Therefore, the transconductance can be written as equation 5.27. In figure 5.22, the red line corresponds to the theoretical values, while the blue line corresponds to our measured/used values.

$$\begin{aligned} g_m &= 2 \cdot \left(\frac{W}{L}\right) \cdot \left(\frac{\mu\varepsilon_{ox}}{t_{ox}}\right) \cdot (V_G - V_T) \\ g_m &= 2 \cdot \left(\frac{200}{40}\right) \cdot \left(\frac{480 \cdot 3.9 \cdot \varepsilon_0}{696\mathring{A}}\right) \cdot (V_G - 1.50V) \end{aligned} \tag{5.27}$$

The corresponding theoretical transconductance values are shown recalculated to I_D in table 5.18. Our measured I_D for transconductance and the difference between the calculated and measured are also shown in this table. Sources of error are as suggested earlier.

V_G	V_{sat}	V_T
-5.0 V	-3.5 V	-5.0 - -3.5 = -1.5 V
-4.5 V	-3.0 V	-4.5 - -3.0 = -1.5 V
-4.0 V	-2.6 V	-4.0 - -2.6 = -1.4 V
-3.5 V	-1.9 V	-3.5 - -1.9 = -1.6 V
-3.0 V	-1.5 V	-3.0 - -1.5 = -1.5 V
-2.5 V	-1.1 V	-2.5 - -1.1 = -1.4 V

Table 5.17: V_T **for** p**-FET with 40**μm **channel.**

V_G	$i_{d_{calc}}$	$i_{d_{meas}}$	% diff (C-M)/C
-5.0 V	1.45 mA	0.61 mA	58
-4.5 V	1.07 mA	0.48 mA	55
-4.0 V	0.74 mA	0.35 mA	52
-3.5 V	0.47 mA	0.24 mA	49
-3.0 V	0.27 mA	0.16 mA	40

Table 5.18: g_m **for** p**-FET with 40** μm **channel.**

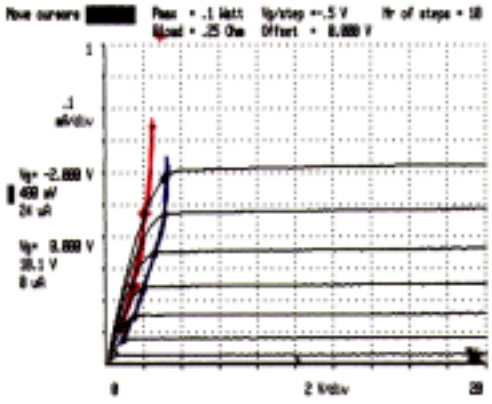

Figure 5.22: p-**FET with 40** μm **channel.**

5.2.3 Inverters

Selecting one each of an inverter with matched 20 μm features and one with matched 40 μm features, we carefully varied the input voltage V_{in} while recording the output voltage V_{out}. The resultant *voltage-transfer curve* (VTC) is shown for the 20 μm inverter in figure 5.23, and the VTC for the 40 μm inverter is shown in figure 5.24.

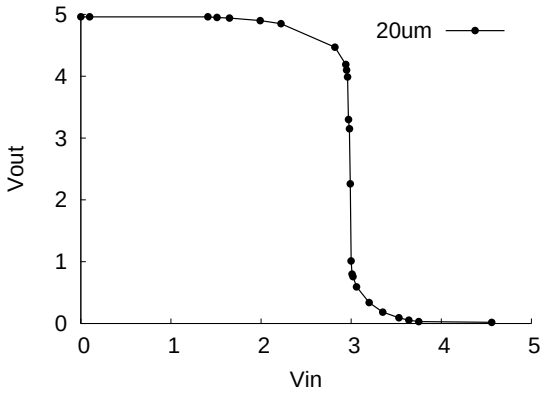

Figure 5.23: **VTC sketch for 20** μm **channel inverter.**

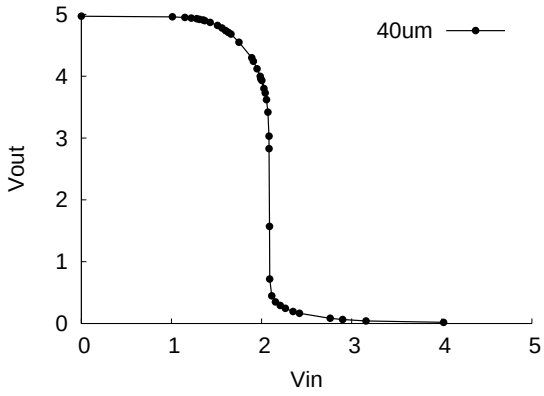

Figure 5.24: **VTC sketch for 40** μm **channel inverter.**

Considering the 20 μm inverter first, we determine that the V_{OH} is 5V, and V_{OL} is 0V. Moreover, the first tangent with slope of -1 occurs near the V_{in} value of 2.5V. The second tangent with slope of -1 occurs near the V_{in} value of 3.25V. Therefore, our V_{IH} is 2.5V and our V_{IL} is 3.25V. Our noise margins are $V_{NM_H} = 2.5V$ and $V_{NM_L} = 3.25V$.

Examining the 40 μm inverter first, we determine that the V_{OH} is 5V, and V_{OL} is 0V. The first tangent with slope of -1 occurs near the V_{in} value of 1.75V. The second tangent with slope of -1 occurs near the V_{in} value of 2.25V. Therefore, our V_{IH} is 1.75V and our V_{IL} is 2.25V. Our noise margins are $V_{NM_H} = 4.75V$ and $V_{NM_L} = 2.25V$.

Neither device is symmetrical.

5.3 Conclusion

All of our devices, with the exception of the 10 μm channel p-FET, worked correctly. We experienced some variation in different situations, with disagreements notable between the four-point probe, the curve tracer, and analytical evaluations. We also had varying threshold voltages per device, with some variations substantial. Many of these seemingly random variations can be explained by our random selection of devices to test across two wafers, not just one wafer.

Additional errors, as suggested in the various subsections prior, can be attributed to imprecise processing, over- or under-development of layers, damage from tweezer contact, and so forth.

On the whole, the most all-encompassing conclusions can be briefly summarized as follows:

1. *Models Are Untrustworthy*
   It is apparent that the models we used are not particularly accurate when dealing with the real world situation. Deviation between programs like SSuprem and the empirical measurements suggest better models not only are possible, but may be essential as process technology continues to improve.

2. *Graphs Are Not Sufficient*
   Reading critical data values off of graphs, without a proper underlying model to use, leads to highly imprecise results. Whether the graph is from a curve tracer, Irvin's Curves, or solid solubilities, it is not sufficient to obtain critical data in this manner. The variance in reading is too great.

3. *Controlling Each Parameter Is Imperative*
   During the various stages of processing, whether it be mask alignment, developing, etc., absolute monitoring and control of the stage is critical. Overdevelopment, as shown in this report, will lead to substantial errors in the final product. Control over agitation, flow rates, alignment of wafers in ovens, etc., can all make the difference between one-off prototypes every time compared to consistent results that always work.

4. *Be Paranoid*
   Trust no one and nothing with your wafer. You can't see if there are minute particles that could interfere with your processing, so be sure to clean the wafer every time. Always check that the solutions, temperatures, and machine settings are correct and never trust they are set the way you left them. Continuously check the calibration of sensitive equipment. Be paranoid, and be sure you do everything as correctly as possible. However, this must be balanced against the high cost of time in a real clean room facility. The balance to be struck is between high accuracy and fast in-and-out times on the ledger.

Chapter 6

About this document

This document was typeset in LaTeXon Linux and Solaris machines. SSuprem3 images were captured via Windows *Ctrl-PrntScrn* and pasted into Microsoft Paint. These BMPs were then saved and exported to industry standard EPS format. Standard packages were used to obtain SI units, picture handling, and content rendering.

The non-text scans (i.e., curve tracer outputs, etc.) were performed on a Linux machine with an attached HP ScanJet IV with Auto-Document Feeder (ADF). This greatly simplified the process of scanning each page and rendering a cropped, re-sized output file. The machine runs RedHat Linux 8.0.

Josh Fryman and Chad Huneycutt are PhD students in the College of Computing, working in the field of Computer Architecture. Josh emphasizes low-power and embedded processor design issues, while Chad is oriented toward compiler-architecture interactions and dynamic translation of binaries.

A hyper-linked PDF version of this document is available upon request.