devices Documentation

Release 1.0

Guofu Niu

December 06, 2011

CONTENTS

CHAPTER

ONE

OVERVIEW

1.1 Objectives

To provide an overview of the history and recent development in semiconductor devices, with emphasis on:

- Five imperfections underlying the usefulness of semiconductors
- The two letters, P and N
- Diffusion and drift of carriers
- Bipolar and field-effect transistors
- Scaled sub-100nm devices
- Bandgap engineering

1.2 Notes

1. `overview`

CHAPTER

TWO

SEMICONDUCTOR FUNDAMENTALS ILLUSTRATION WITH TCAD

2.1 Objectives

- Understand quantitatively how fundamental material parameters vary in semiconductor, e.g. band gap, electron affinity, dielectric constant
- Understand relationships between various potentials and energies important for understanding device operation
- Understand how biasing affects various potentials and energies at device terminals / ohmic contacts
- Introduce essential concepts of surface charge modulation

2.2 Required Programs

- tecplot_sv 2010 version (2008 version may work as I use latest features of Sentaurus, I did not check this. The .cmd files did not work with sdevice 2008). As of 8/21/2011, there is issue with running tecplot_sv on the engineering Linux workstations. I have reported the problem and hopefully the issue will be resolved soon. I will then supply you with the necessary files and give you a live tutorial on how to inspect TCAD simulation outputs. To keep it manageable, we will focus on using existing codes provided so that you can use TCAD to understand devices without having to spend a lot of time putting together working TCAD input decks or coding.

2.3 Overview

There have been many good books on semiconductors and devices you can refer to for theories. In this lecture, we will look at a concrete numerical example using TCAD, with a 1D cut done on a 2D MOSFET structure.

By design, we will simulate the polysilicon gate as polysilicon, rather than a piece of ideal metal gate with an artificially adjusted work function. This is a much better approximation to reality, and allows the important polysilicon gate depletion effect to be correctly accounted for.

This choice has a unique advantage of allowing easy visualization of effects of bias on quasi-Fermi potential energies and potentials at device terminals. The easiness is a result of zero current flowing vertical through the MOS structure. The quasi-Fermi potentials or energies are therefore spatially constant in the polysilicon gate, and thus easy to visualize. For similar reasons, we can see the impact of bias on potentials clearly this way.

Let us first look at structure and material parameters.

2.4 Device Structure

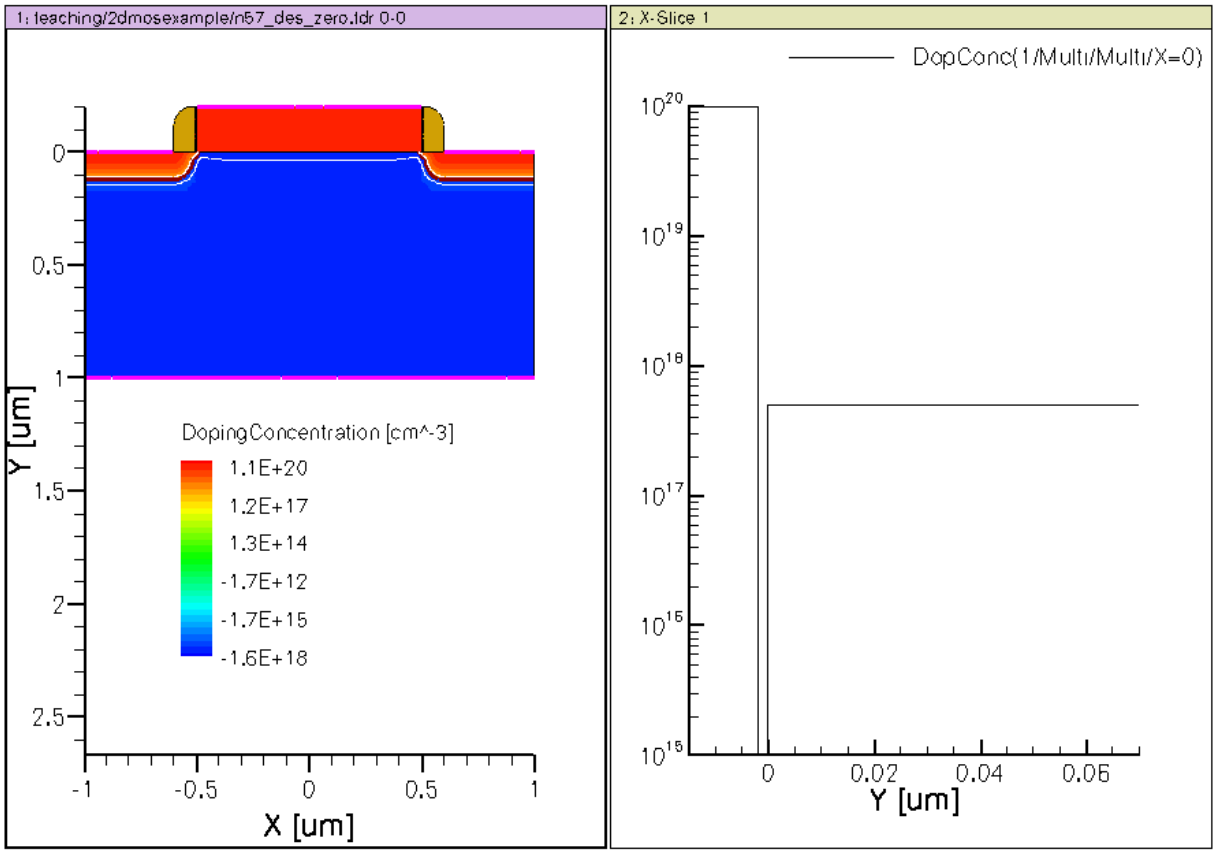


Figure 2.1: 2D MOSFET simulated and 1D cut at x=0 showing the N+ poly gate doping level and p-substrate doping level.

2.5 Material Parameters

2.5.1 band gap Narrowing (BGN)

2.5.2 Effective band gap

band gap in the gate oxide is much larger than in Si.

2.5.3 Electron affinity

The amount of energy it takes to free an electron at the bottom of the conduction band, i.e. the Ec level, is called electron affinity. You can also consider this to be the difference between the so-called *vacuum energy level* and Ec.

Electron affinity can be used to determine the Ec alignment or offset between two different material regions, e.g. between the gate oxide and N+ polysilicon gate, between the gate oxide and the silicon substrate.

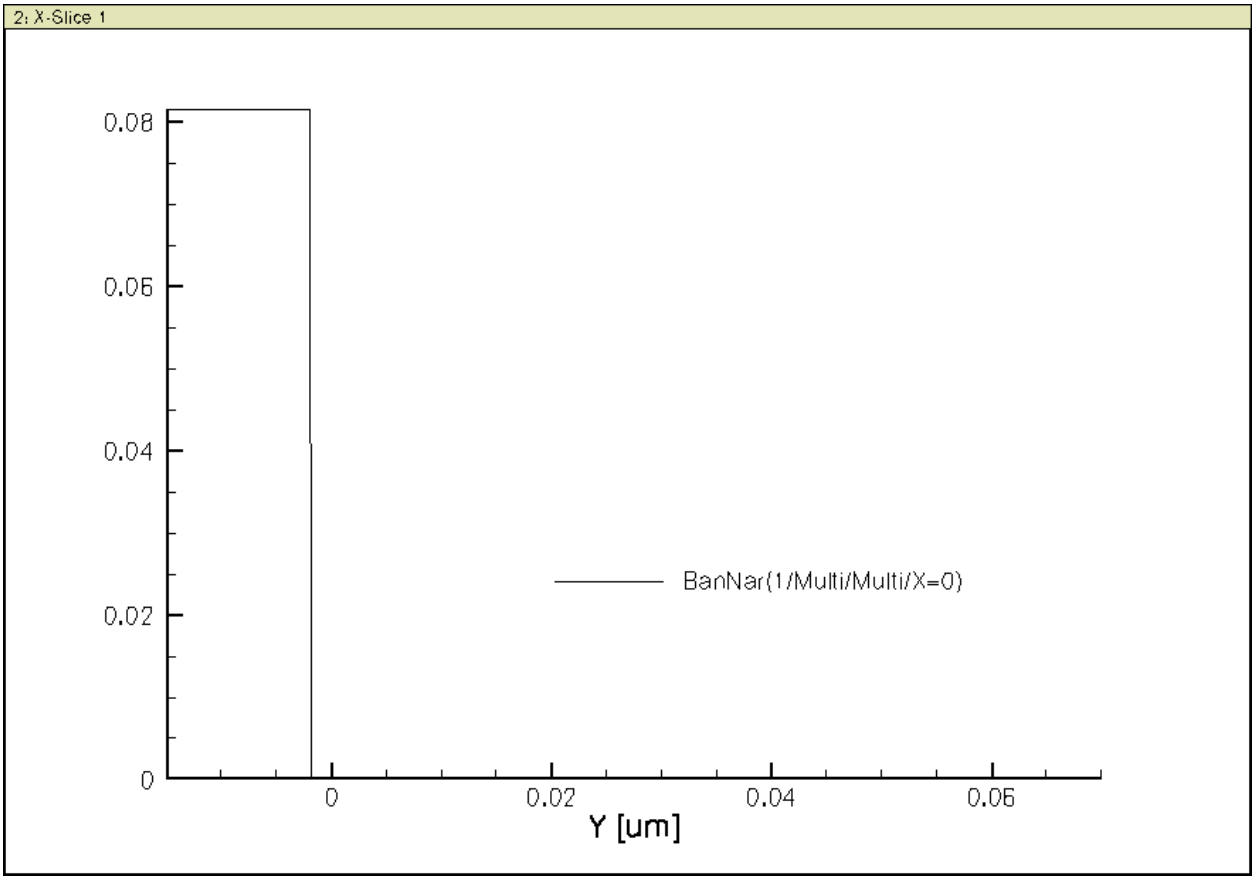


Figure 2.2: BGN is observed in N+ polysilicon gate because of heavy doping

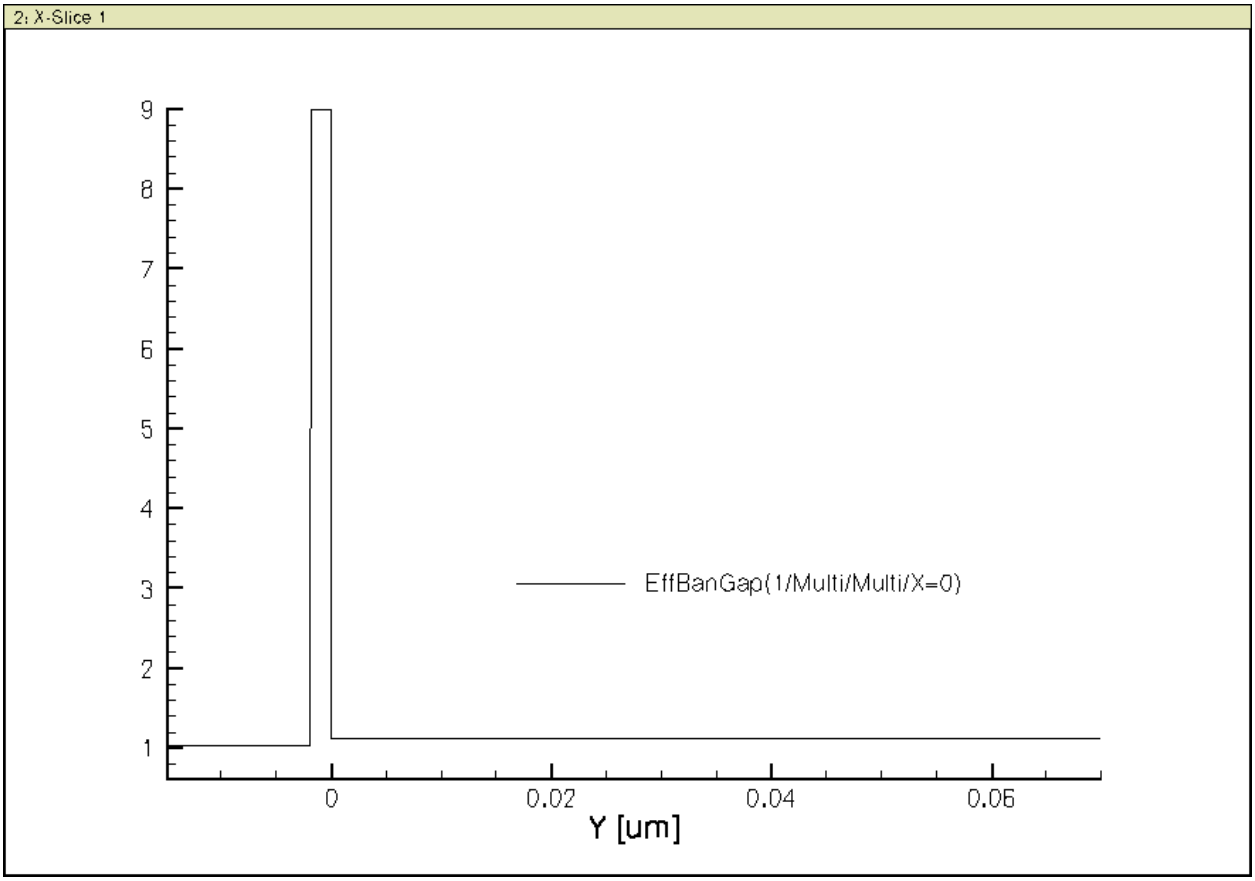


Figure 2.3: Effective band gap is smaller in N+ polysilicon gate because of heavy doping induced band gap Narrowing (BGN).

Similarly, we can determine Ev alignment between two material. We need to account for the band gap difference in addition to electron affinity difference.

Electron affinity and band gap differences will also be used to determine band alignments at *heterojunction* interfaces as well. So keep this in mind.

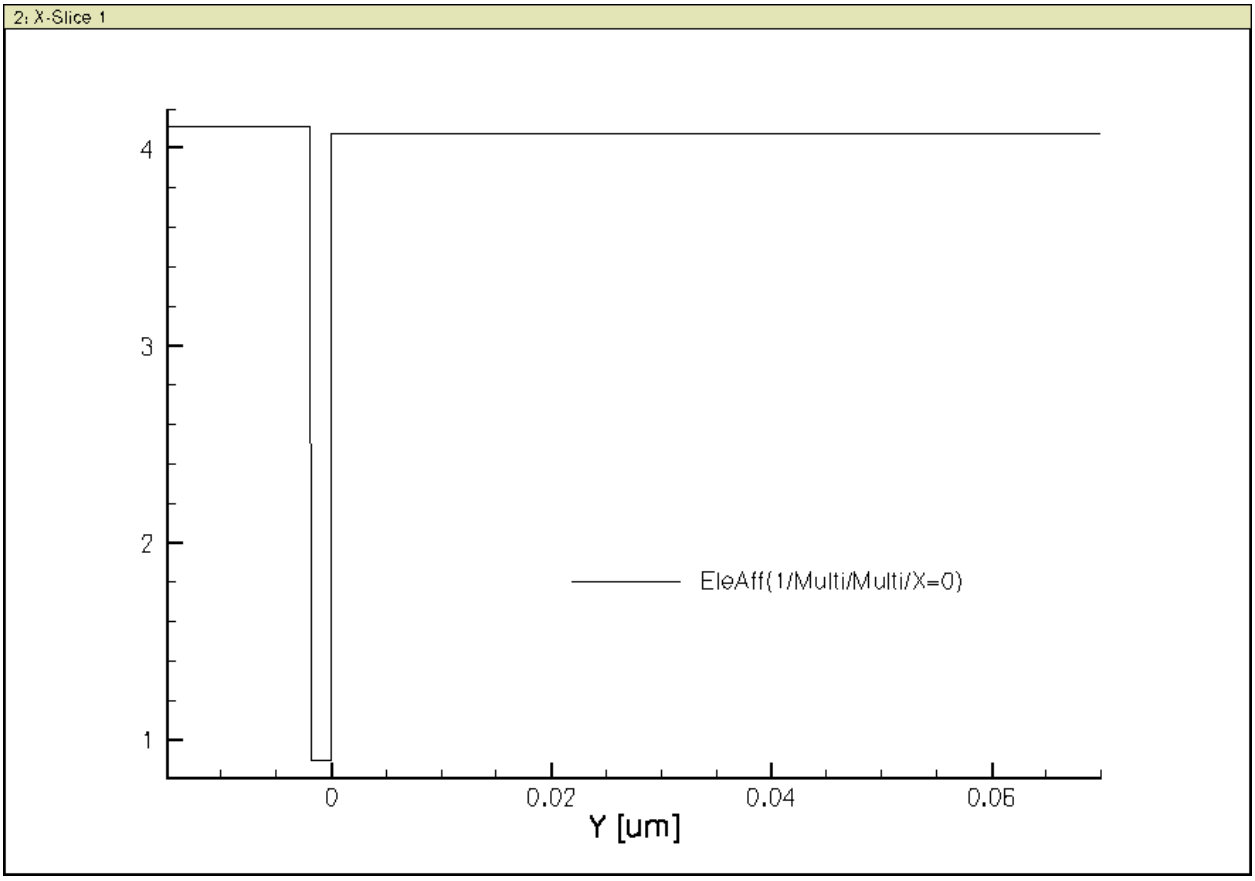


Figure 2.4: Electron affinity.

2.6 Zero Bias

2.6.1 Electrostatic Potential

Consider zero bias at all terminals. There is still *built-in* potential. The most convenient reference is a piece of intrinsic semiconductor in TCAD. Hand analysis in textbooks, however, uses different *reference*. This is very important to note so that you know precisely how to associate simulation results with first order theories in textbooks.

For example, in the p-body, potential is -0.45 eV, as its potential is lower than that of an intrinsic piece by 0.45 eV.

2.6.2 Band Diagrams with Vacuum Level

The quasi Fermi potentials at all contacts are simply equal to the applied voltages. Since we are applying zero voltages at both gate and body, the quasi Fermi potentials are zero.

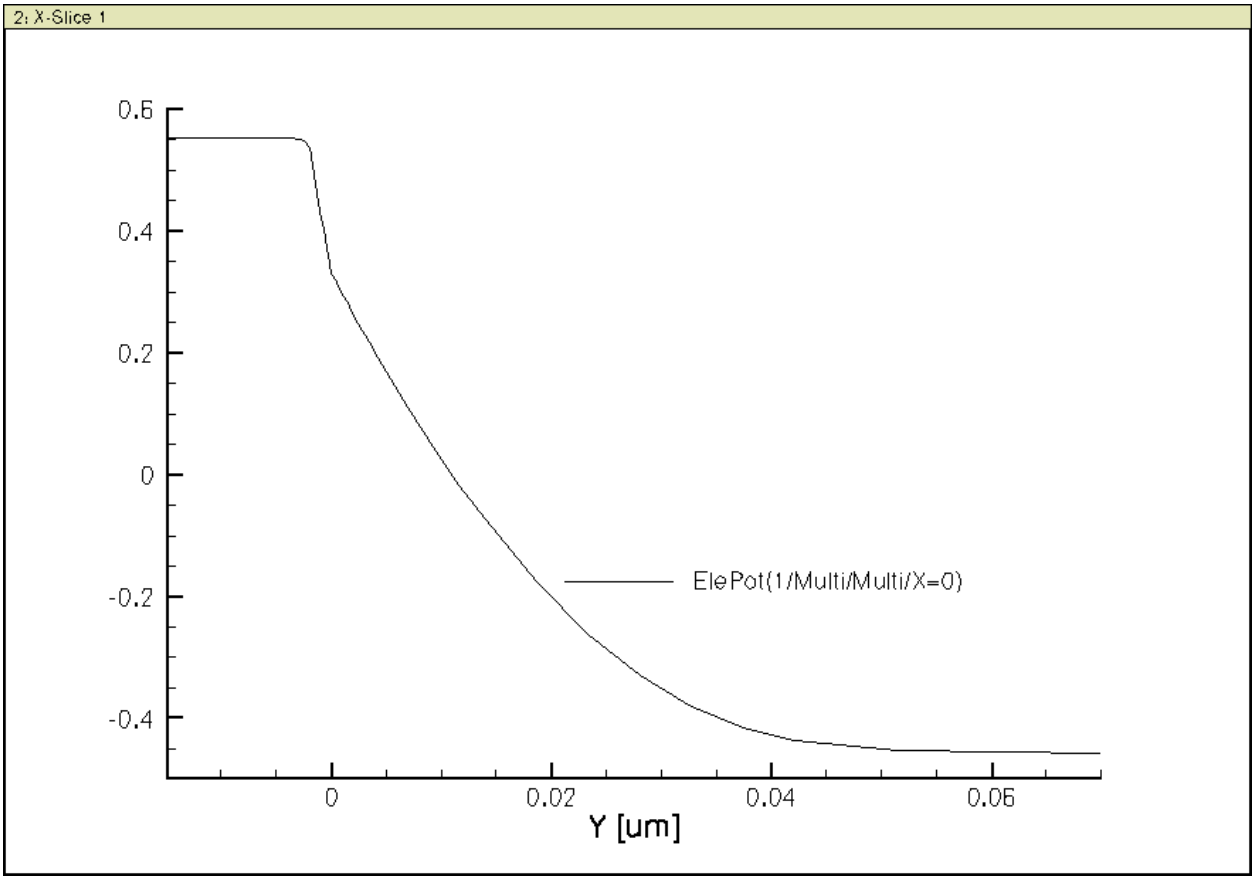


Figure 2.5: Electrostatic potential. Note reference is the potential of an intrinsic semiconductor under zero applied bias.

Quasi Fermi energy is simply the inverse of quasi Fermi potential, if you use eV for energy, and V for potential.

There is built-in potential between the N+ polysilicon gate and the p-body. Potential is higher in the N+ poly gate, lower in the p-body.

Therefore, the Ec at the gate is *lower* than the Ec deep in the body.

I have also calculated the vacuum energy level, which is simply equal to Ec + affinity. The vacuum energy level is no longer spatially constant, because of the built-in electric field. The gradients of Ec, Evac, Ev are all the same, determined by electric field.

Electron affinity and band gap differences will also be used to determine band alignments at *heterojunction* interfaces as well. So keep this in mind.

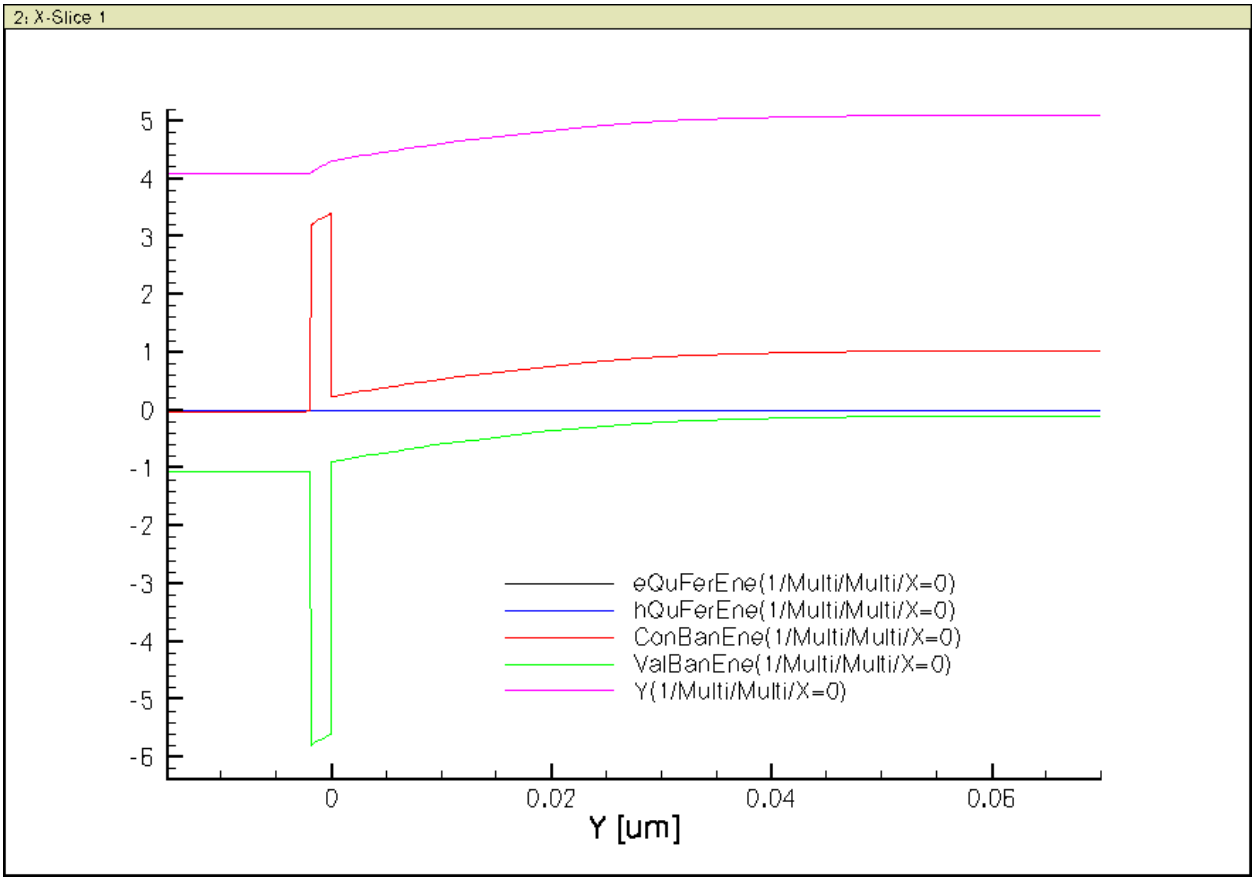


Figure 2.6: Band diagram with Evac, the vacuum energy level.

2.6.3 Band Diagrams with Intrinsic Fermi Level

Often we do not draw the vacuum energy level, despite its great importance described above. We draw only Ec, Ev, Efn, and Efp.

That is enough information as we know n, p, and electric field (from gradient of band edges).

However, often we also draw the *intrinsic* Fermi level, or Ei, particularly in hand analysis. The reason is simple. It allows easy visualization of whether a region is n-type or p-type.

In regions where Efn = Efp = Ef, if Ef > Ei, it is n-type. If Ef < Ei, it is p-type.

Below is a band diagram showing Ei:

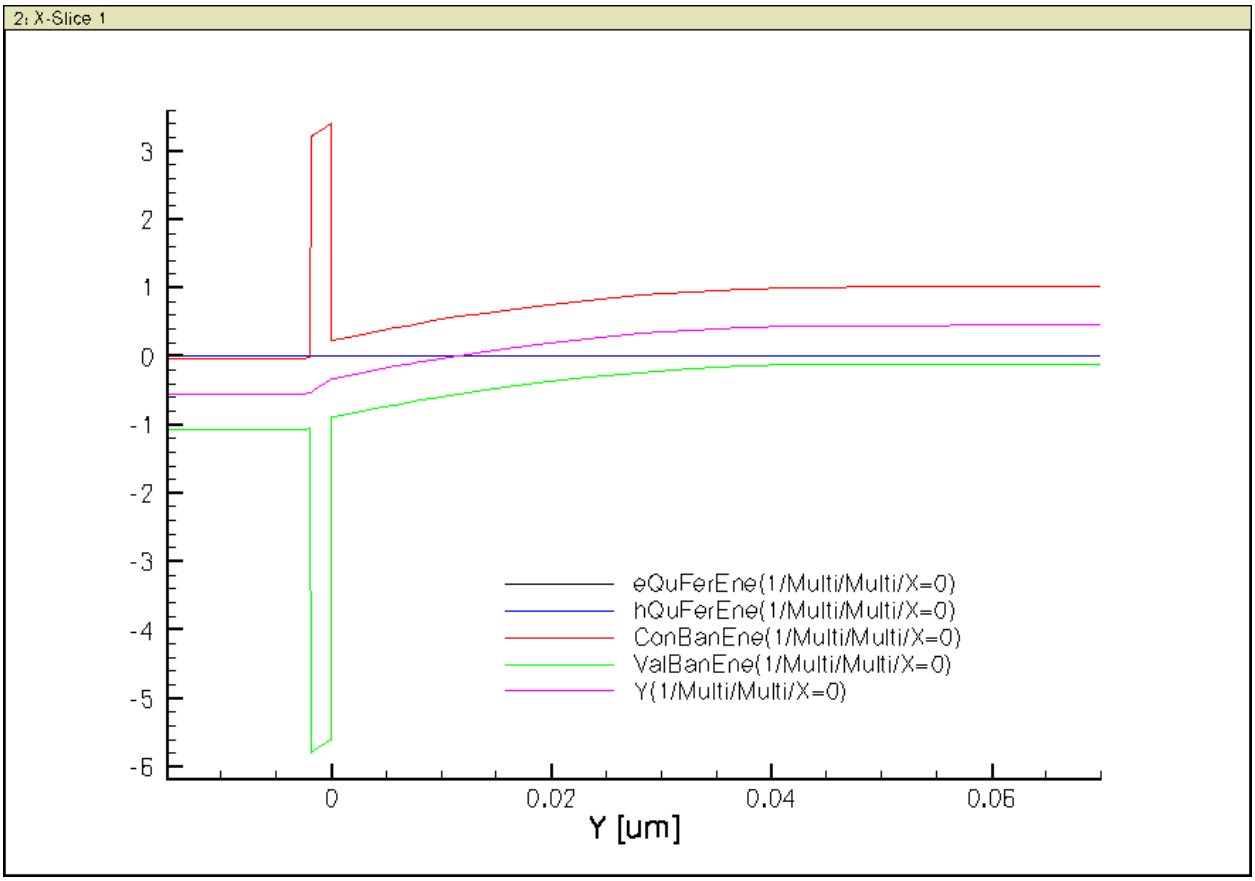


Figure 2.7: Band diagram with Ei, the intrinsic Fermi level.

Because of the doping difference between N+ poly gate and p-body, there is band bending at the silicon surface. You can easily picture how electron concentration n and hole concentration p vary from this band diagram.

Sketch that yourself, then compare with the following plots:

Observe how n increases towards the gate oxide/Si interface and p decreases in the process. The pn product is kept constant.

Can you explain why the pn product in the N+ poly gate is larger than in the Si body?

2.7 Gate Bias Dependences

Recall from Elec6700 that at ohmic contacts, no deviation from equilibrium is allowed. This means that both n and p values are kept at their equilibrium values, determined by doping at the contacts.

Also recall that Efn and Efp must then be equal at the contact points.

In the MOS structure, there is no gate current for ideal gate oxide. Therefore, Efp and Efn are spatially constant in the N+ poly gate.

The Ef difference between the gate and the body contact (the contact points only) is set by the applied bias voltage difference.

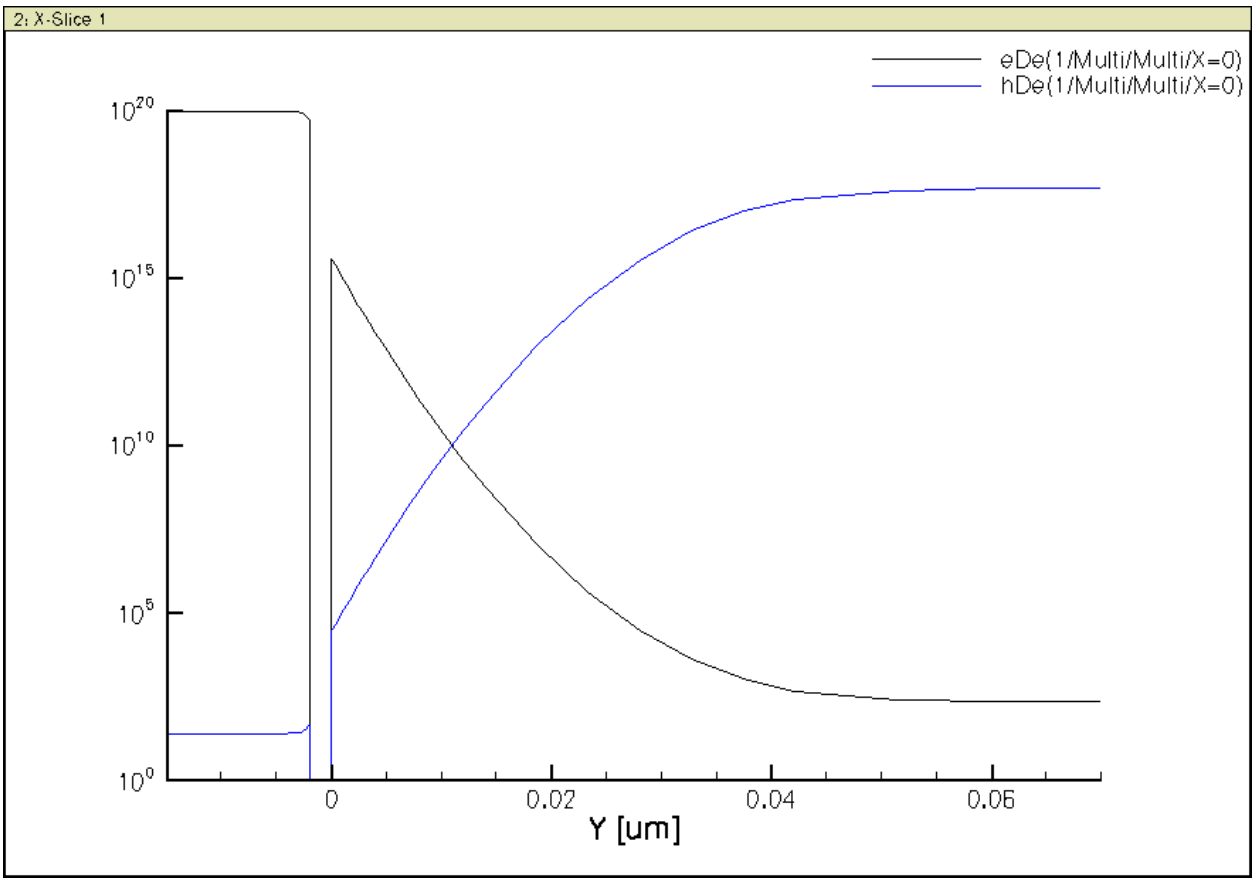


Figure 2.8: n and p at zero bias.

If we simply set the Ef at the contacts to the applied bias voltage, we can prove that the electrostatic potential at a contact must then be equal to the applied voltage + the built-in potential of that contact with respect to an intrinsic reference. I'll leave the proof to you. Stop by if you have trouble doing so.

The (quasi) Fermi potentials at the contacts are therefore simply equal to their applied voltage values. Now let us keep the body grounded, i.e. at zero bias, and vary the gate voltage from 0 to 1.2 V in 0.2 V step.

Analytical treatment of the 1D MOS structure will be made in a few lectures. Below we show the internal distribution of physical quantities as a function of gate biases from TCAD. A small Vds of 0.05V is applied while the gate bias increases.

2.7.1 Electrostatic Potential

The electrostatic potential distributions for all gate biases are shown below in figure 2.9:

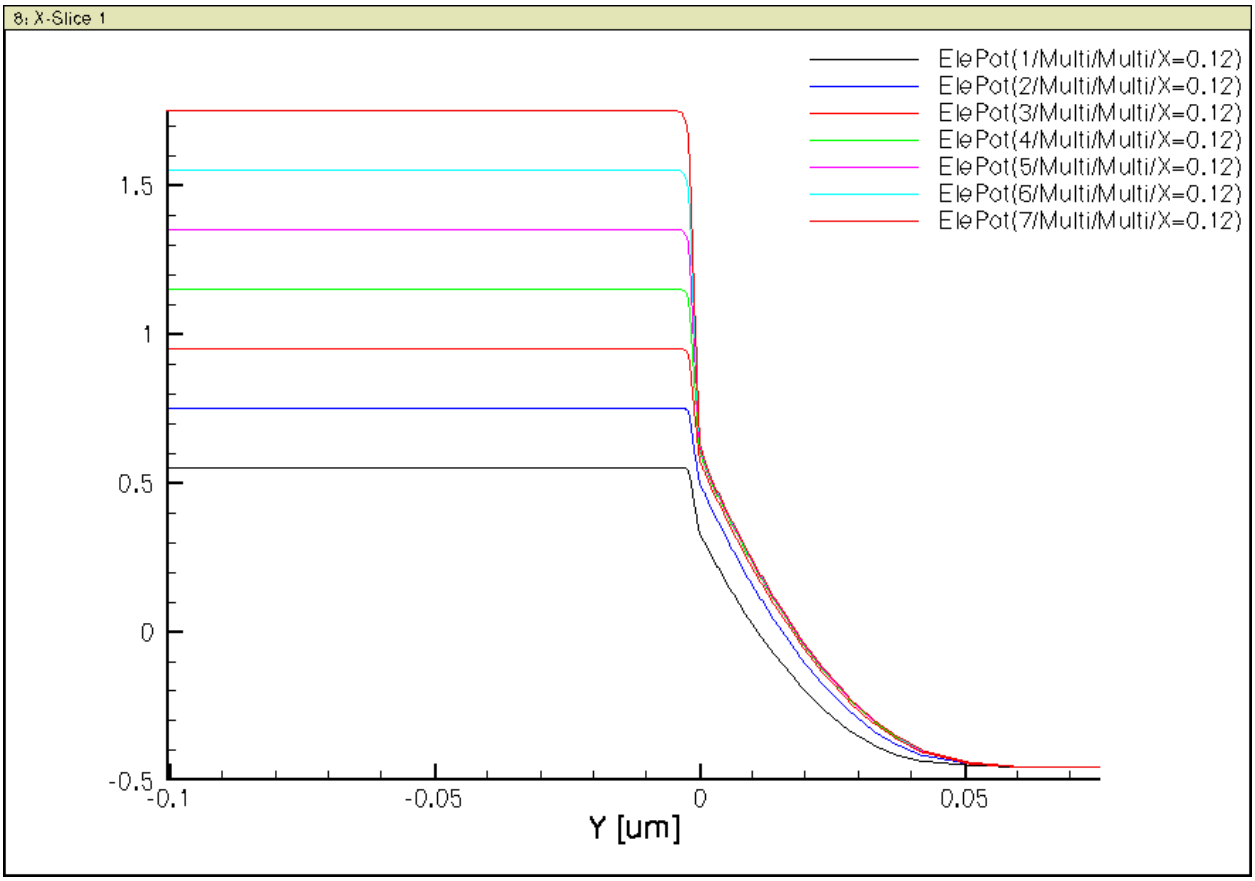


Figure 2.9: Electrostatic potential for all gate biases.

Observe that for Vgs>=0.4V, the potential in the Si surface region changes very little. Much of the further gate voltage increase is absorbed by the very thin gate oxide, there is also a tiny but noticeable amount of potential drop in the heavily doped N+ poly gate, near the gate/oxide interface.

Why this is the case will become clear once we examine the band diagrams.

2.7.2 Quasi Fermi Potentials

electrons

The electron quasi Fermi potential distribution for all gate biases are shown in figure 2.10:

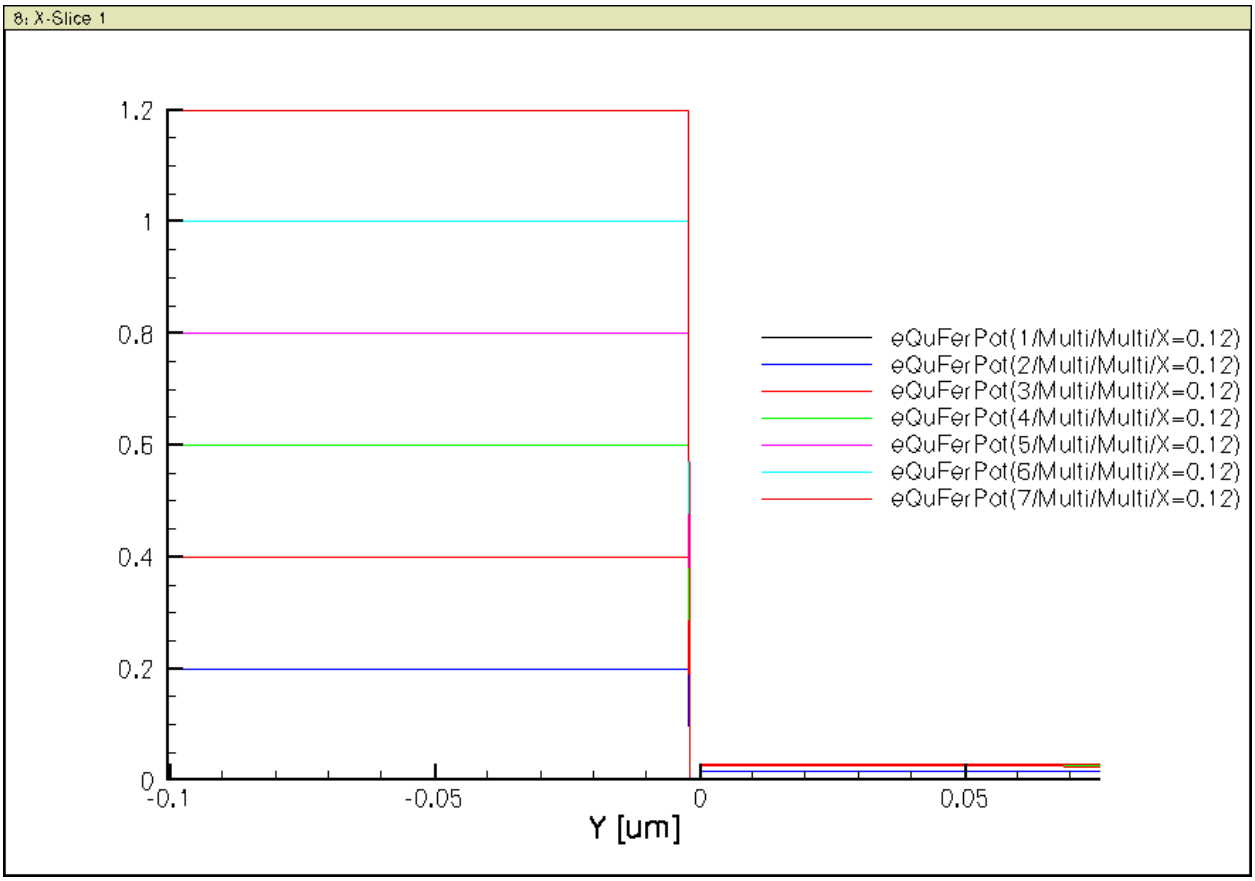


Figure 2.10: Electron quasi Fermi potential distribution for all gate biases.

Observe that the quasi Fermi potential is equal to zero in the silicon substrate (body), and equal to the gate voltage in the N+ polysilicon gate.

holes

The hole quasi Fermi potential distribution for all gate biases are shown in figure 2.11:

Observe that the quasi Fermi potential is equal to zero in the silicon substrate (body), and equal to the gate voltage in the N+ polysilicon gate.

2.7.3 Quasi Fermi Energies

electrons

The electron quasi Fermi energy distribution for all gate biases are shown in figure 2.12:

Observe that inverting quasi Fermi potential expressed in V gives quasi Fermi energy in eV.

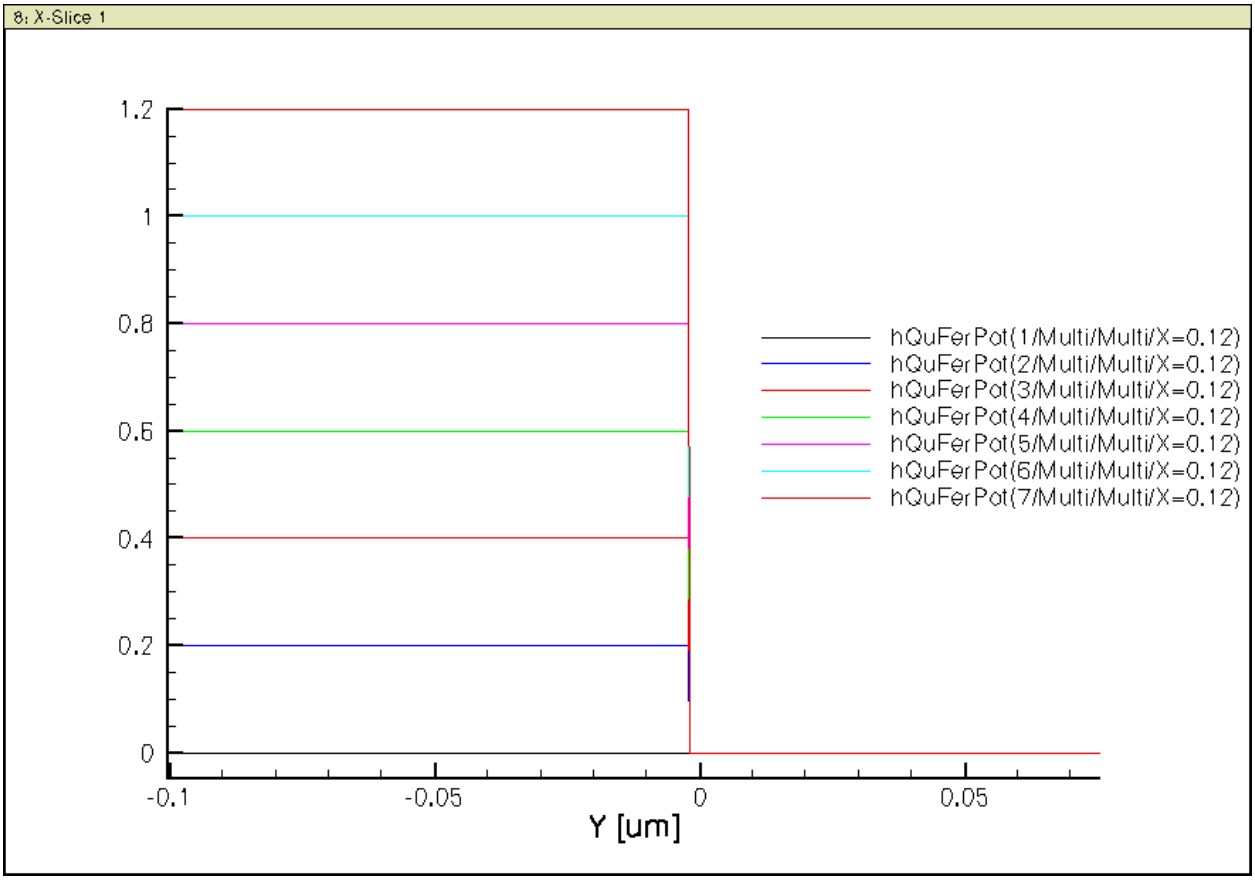


Figure 2.11: Hole quasi Fermi potential distribution for all gate biases.

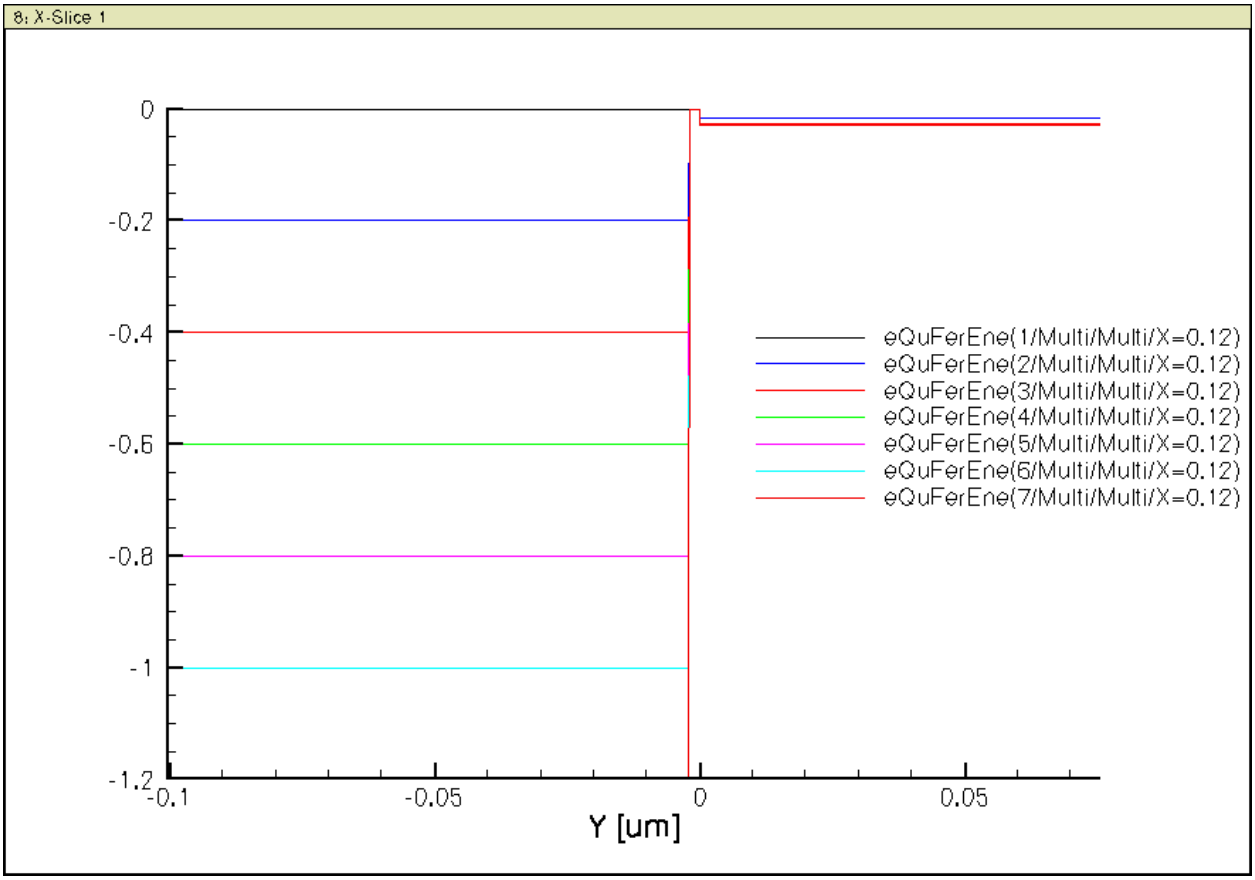


Figure 2.12: Electron quasi Fermi energy distribution for all gate biases.

holes

The hole quasi Fermi energy distribution for all gate biases are shown in figure 2.13:

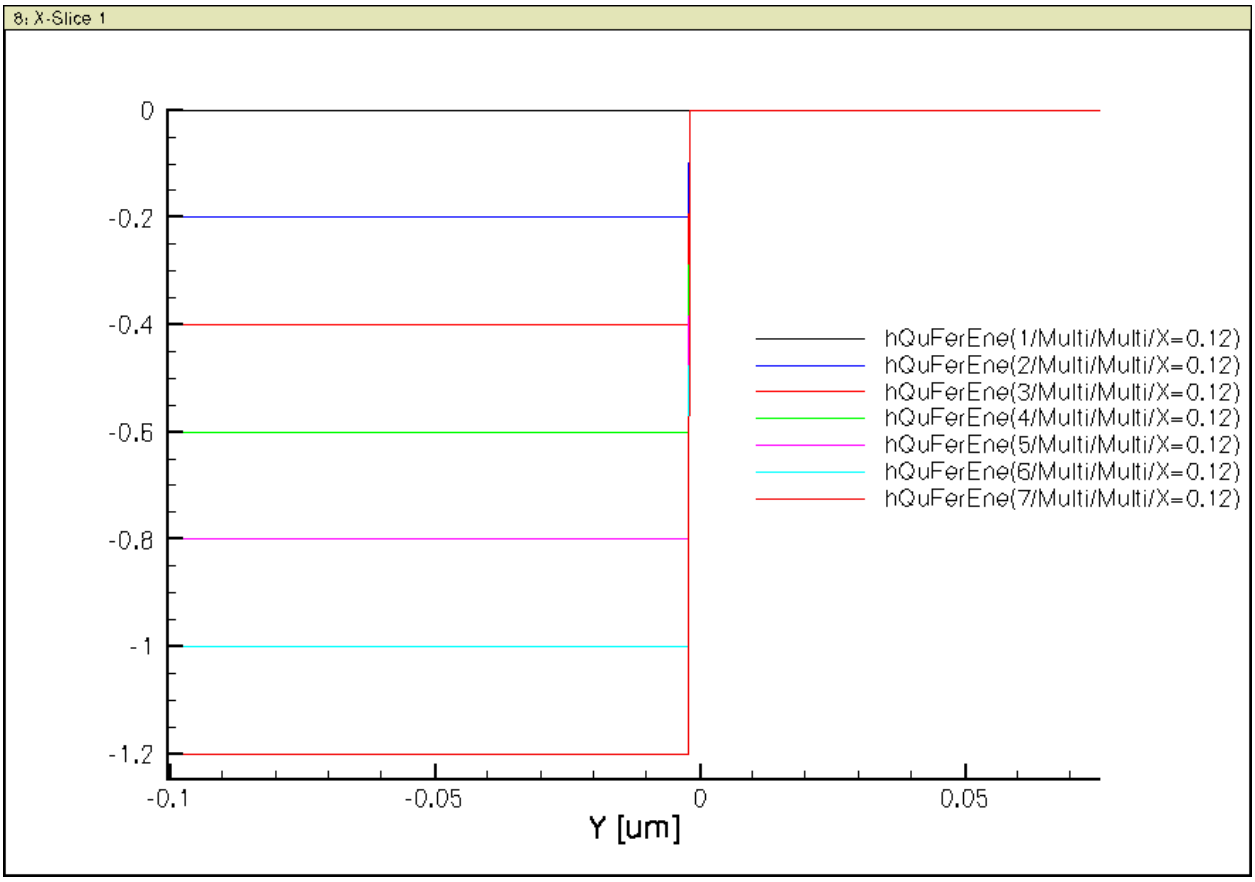


Figure 2.13: Hole quasi Fermi energy distribution for all gate biases.

Observe that inverting quasi Fermi potential expressed in V gives quasi Fermi energy in eV.

2.7.4 Band Diagrams

Now let us look at the band diagrams for all biases shown below in figure 2.14:

The nearly flat curves in the silicon substrate are the quasi Fermi energies of electrons and holes, with a very small separation due to the small 0.05V Vds. For all practical purposes, just view them as spatially constant and equal to 0 eV, the substrate terminal's applied voltage, or its inverse strictly speaking.

Observe that for Vgs>=0.4V, just a little bit more band bending occurs with further increase of gate voltage. This is because the surface electron density (n) is already very high, any small increase produces enough negative electron charge to shield the vertical gate field. Again, recall the letter *n* means negative. Much of the further voltage drop occurs now in the gate oxide, as indicated by the increasing slope of the band bending inside the thin gate oxide.

A plot of Ec alone is shown in figure 2.15:

You might better see the difference in band bending amount between the region inside the gate oxide and the region near the Si surface from the Ec plots.

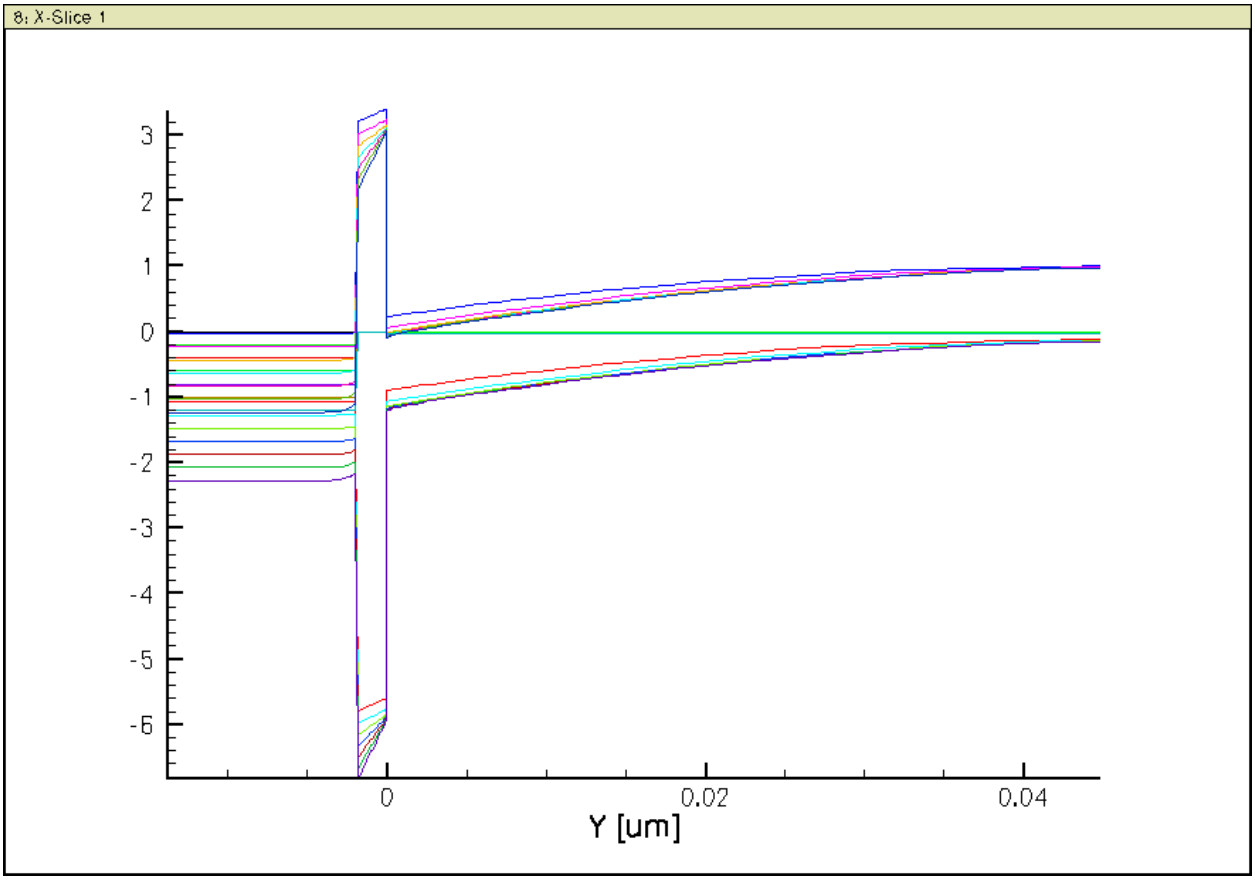


Figure 2.14: Band diagrams- all bands

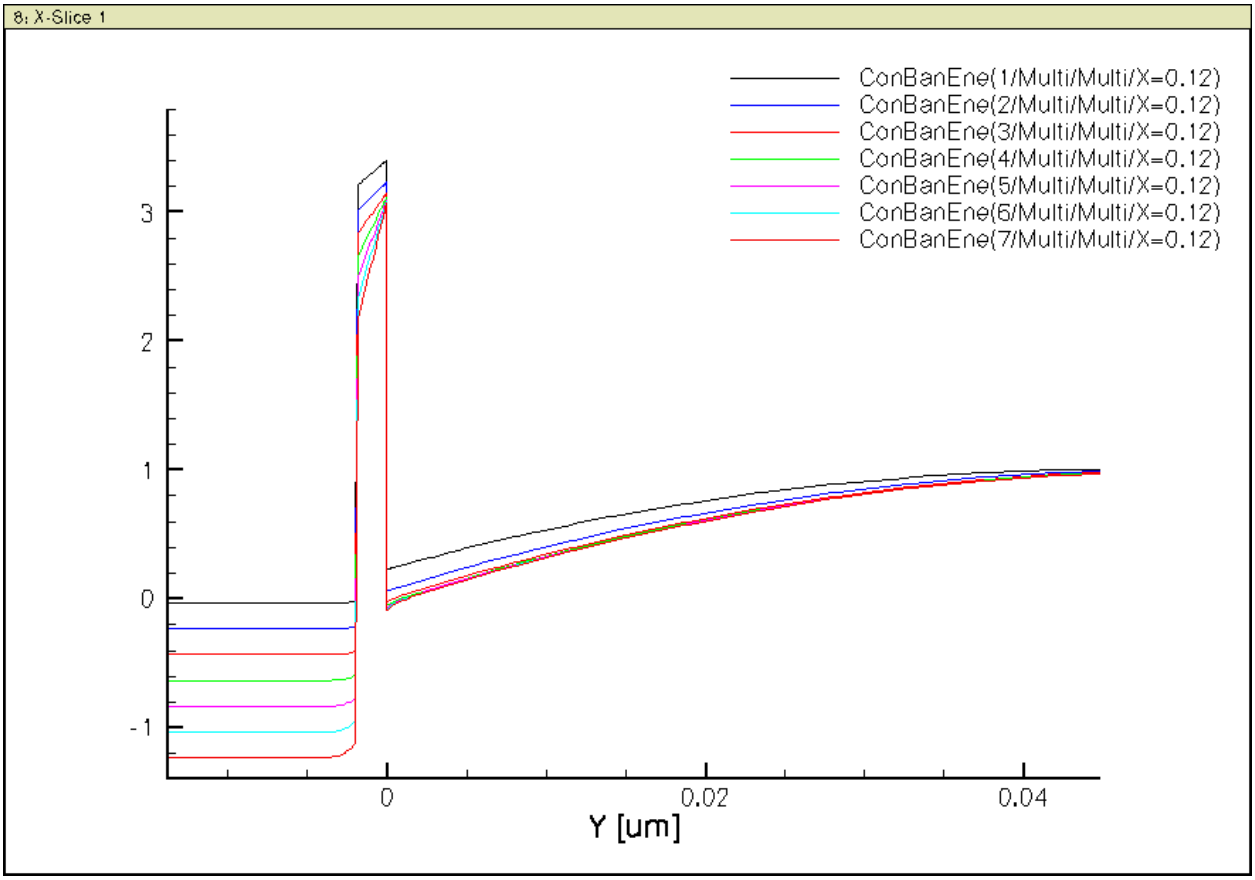


Figure 2.15: Band diagrams - Ec

Also observe that there is indeed a noticeable potential drop inside the heavily doped N+ polysilicon gate, which is responsible for the so-called *poly depletion* effect. That is one of the reasons why in most advanced CMOS, metal gate replaced polysilicon gate.

Just for completeness, let us look at the Ev diagrams shown figure 2.16: too:

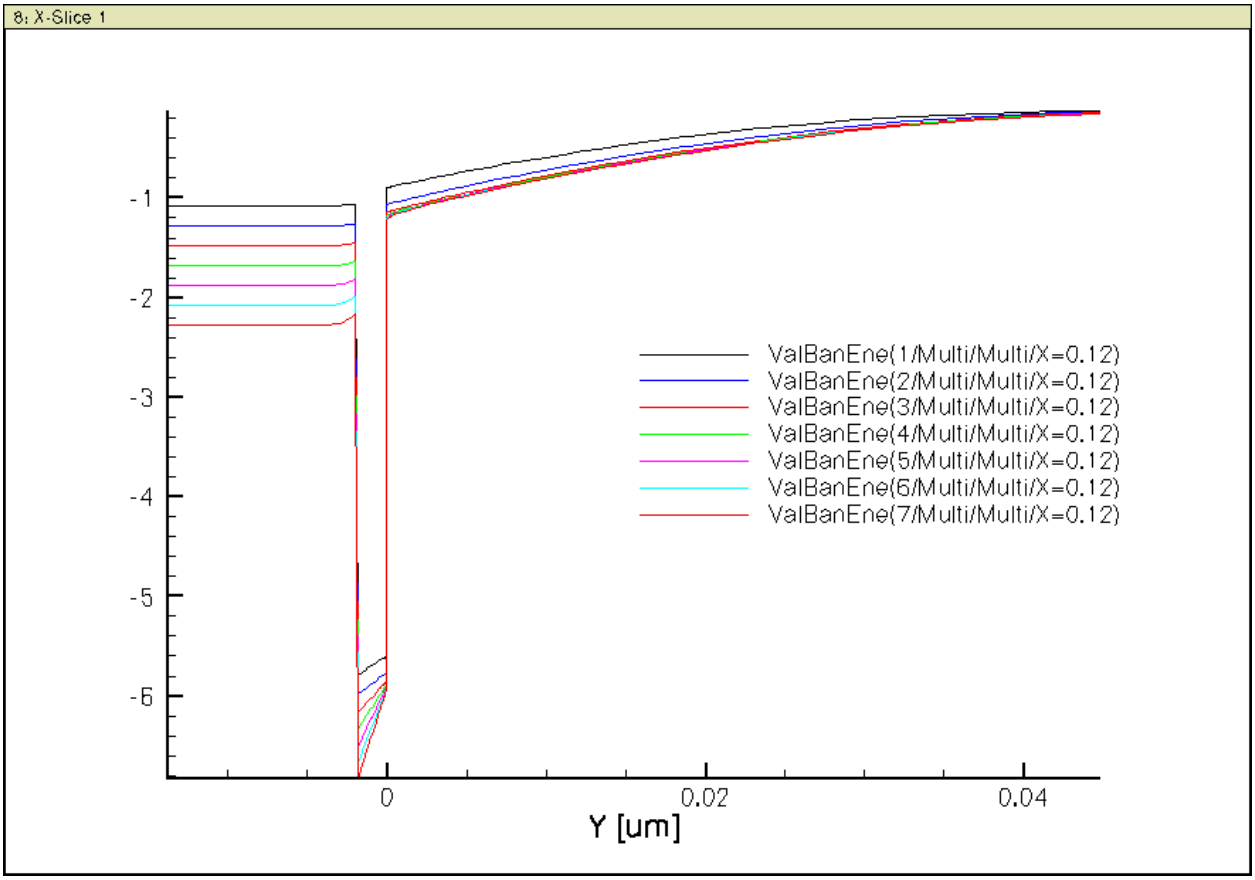


Figure 2.16: Band diagrams - Ev

2.7.5 Inversion Charge and Space Charge

From the band diagrams, you should have expected the following inversion carrier density distributions, in this case, n vs depth plots of figure 2.17:

The p vs depth plots look like that shown in figure 2.18:

The net space charge density, defined by p-n+NdPlus-NaMinus, is shown in figure 2.19:

A linear y-axis scaled is used by design here to make the point that a high spike of positive charge density exists near the gate/oxide interface. These positive charges are due to depletion of the N+ dopants in the N+ poly gate. Even with an extremely heavy 1e20 gate doping, we are still seeing quite a bit poly depletion.

As we will learn later, and some of you might intuitively expect just by inspecting the above plot, the thickness of poly depletion is comparable to the very thin gate oxide thickness. This effectively degrades the gate capacitance, as the depletion layer is in series with the gate oxide.

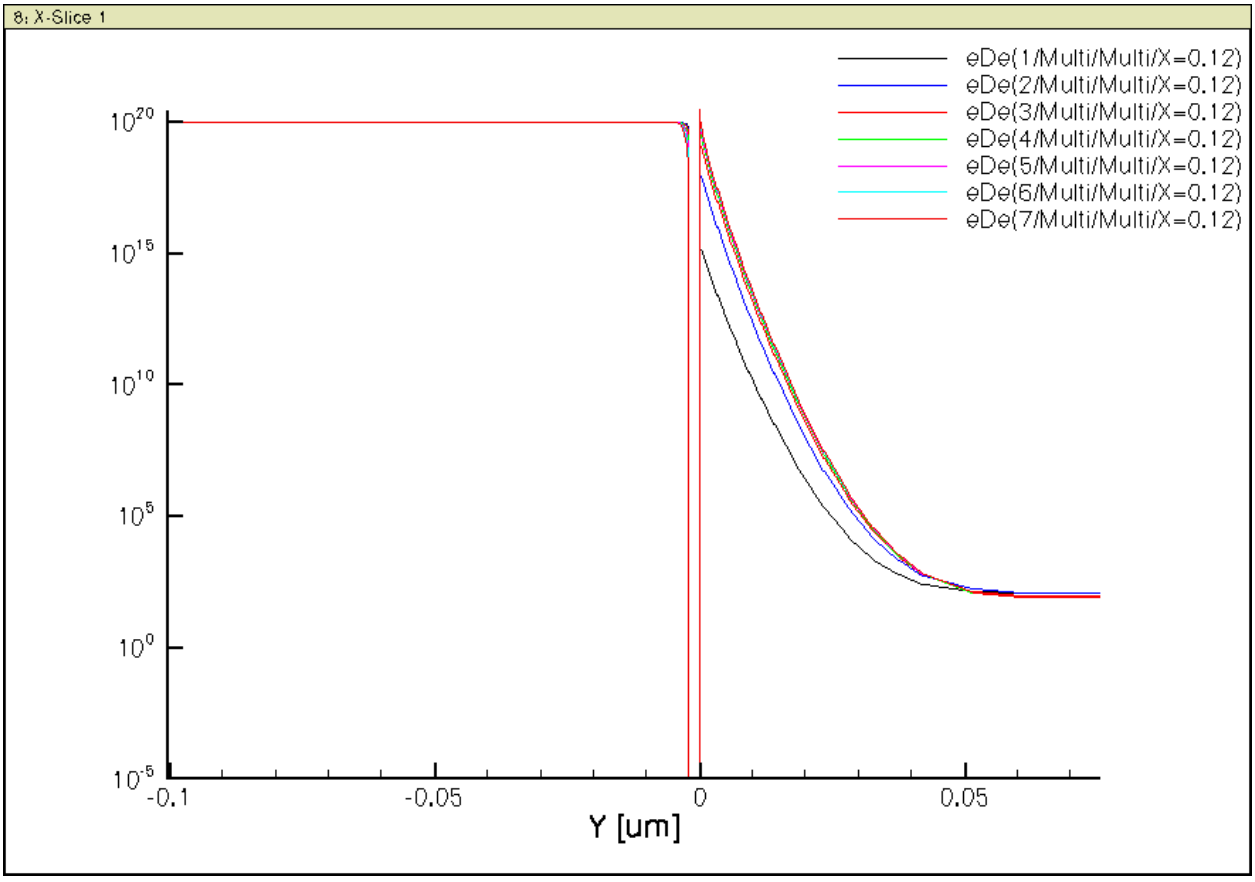


Figure 2.17: Electron density distribution for all gate biases

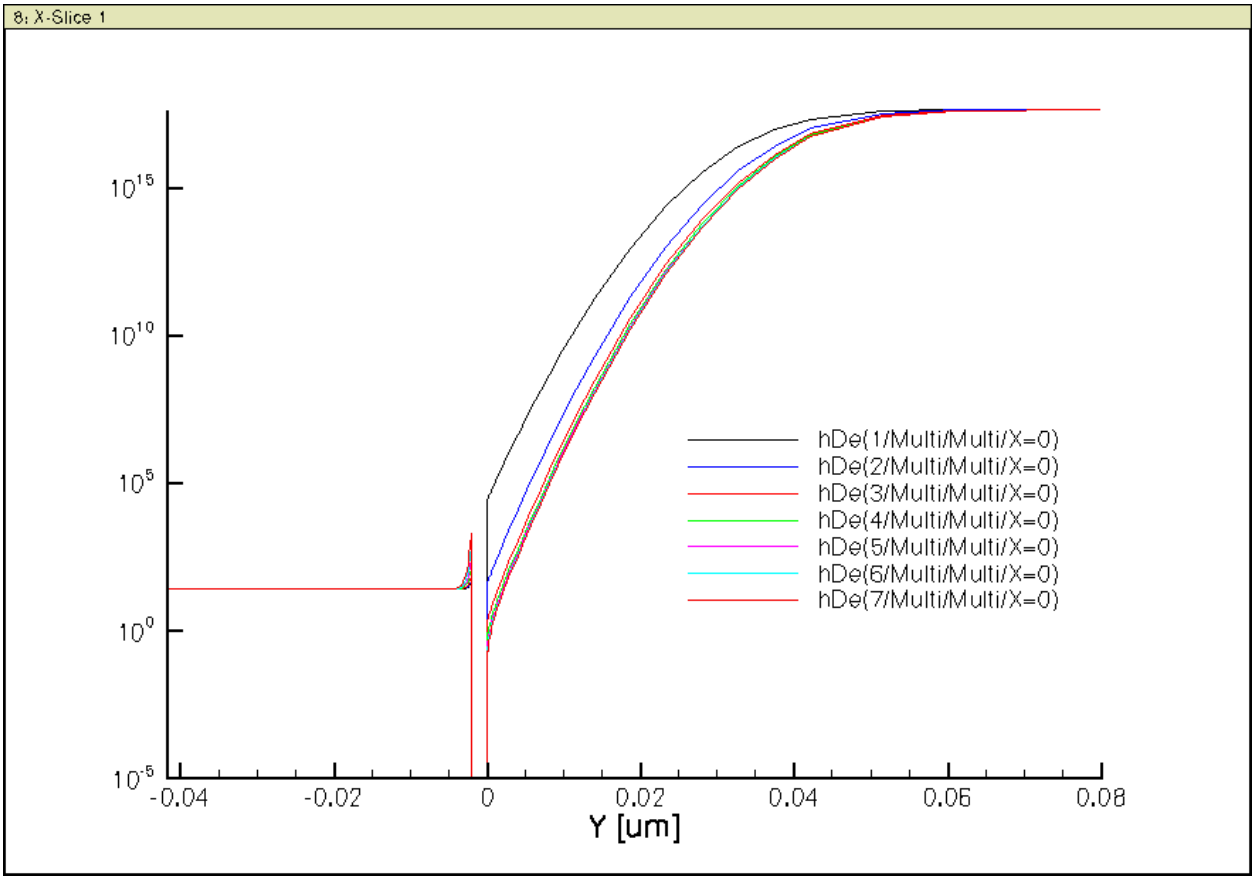


Figure 2.18: Hole density distribution for all gate biases

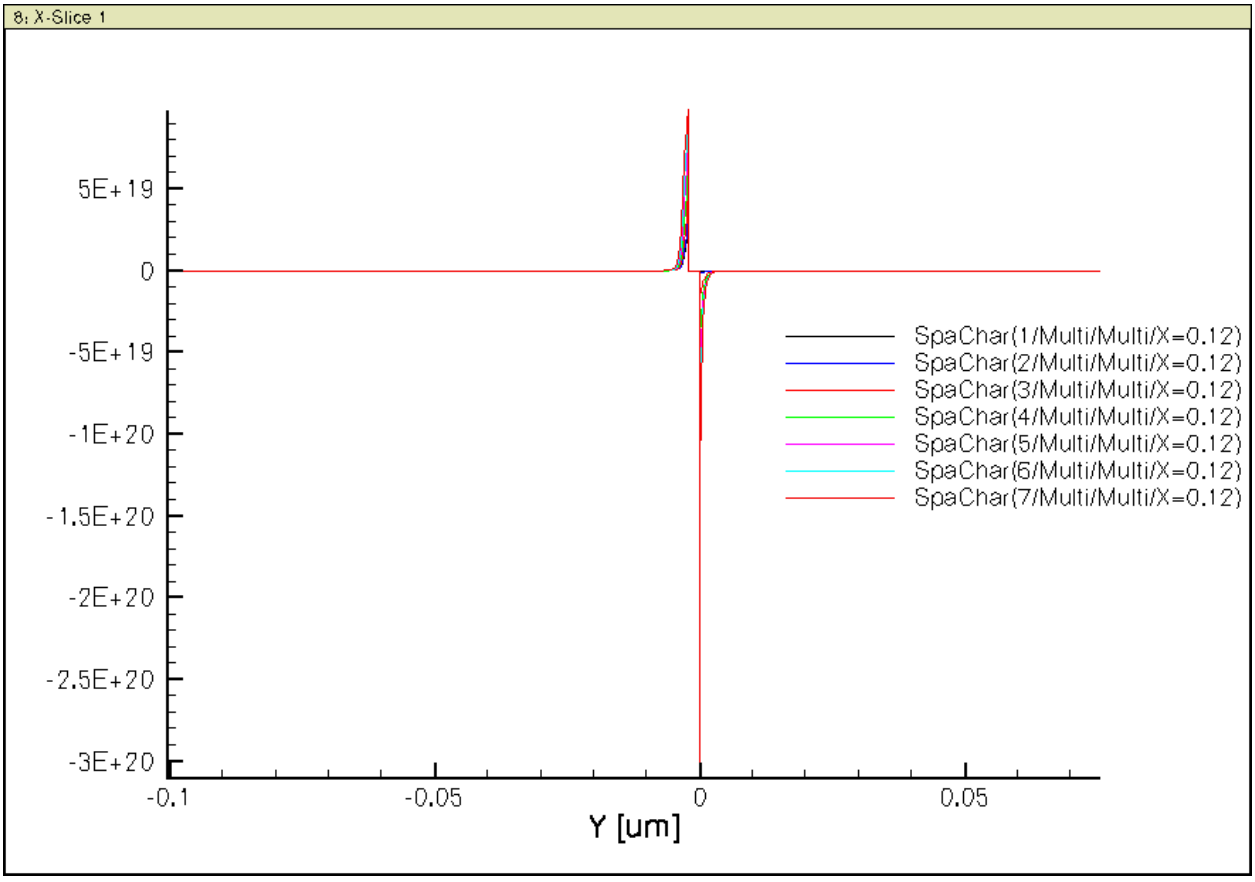


Figure 2.19: Space charge density distribution for all gate biases

2.7.6 Electric Field

The e-field distribution should not be surprising. Once surface is heavily inverted, the surface e-field increases strongly, however, near the surface only, so that there is large surface e-field increase, but little surface potential increase, as shown below in figure 2.20:

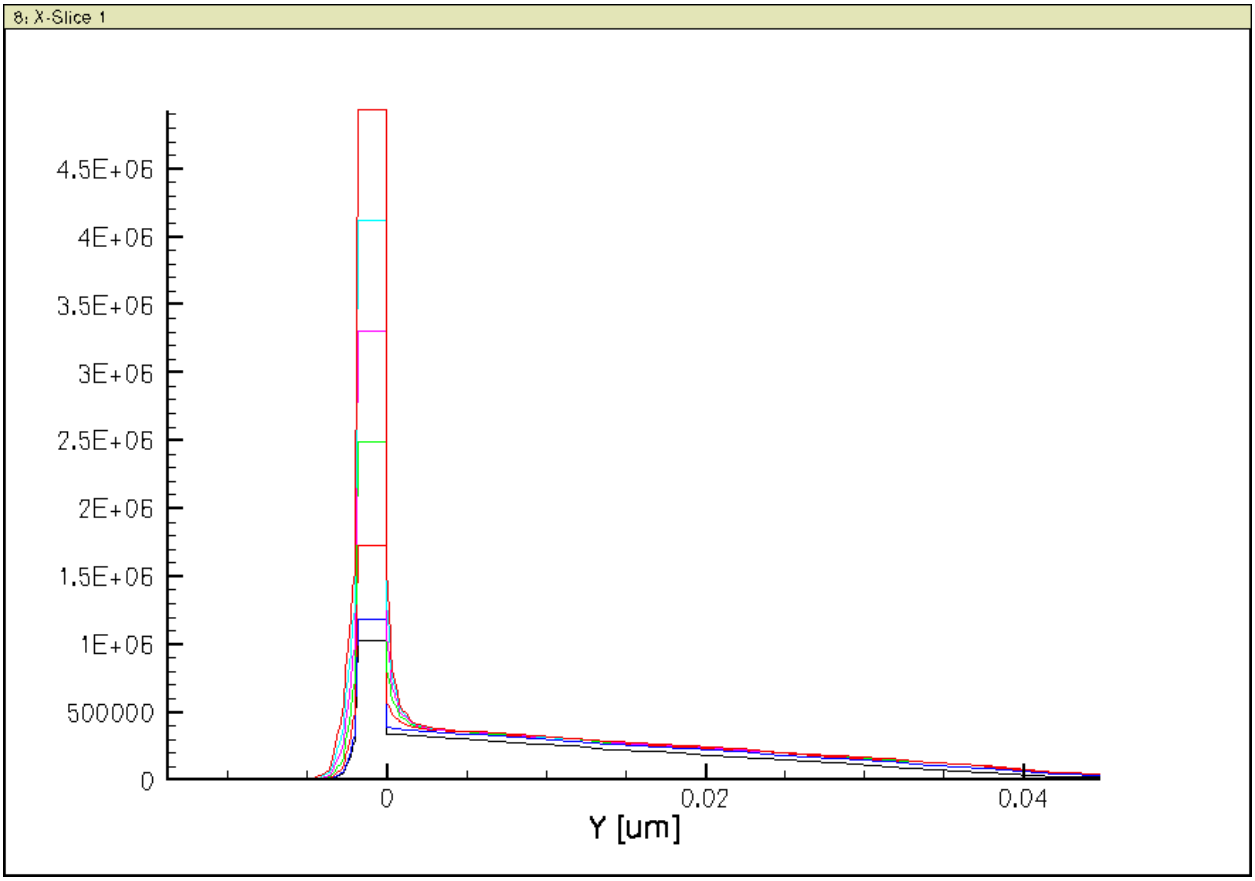
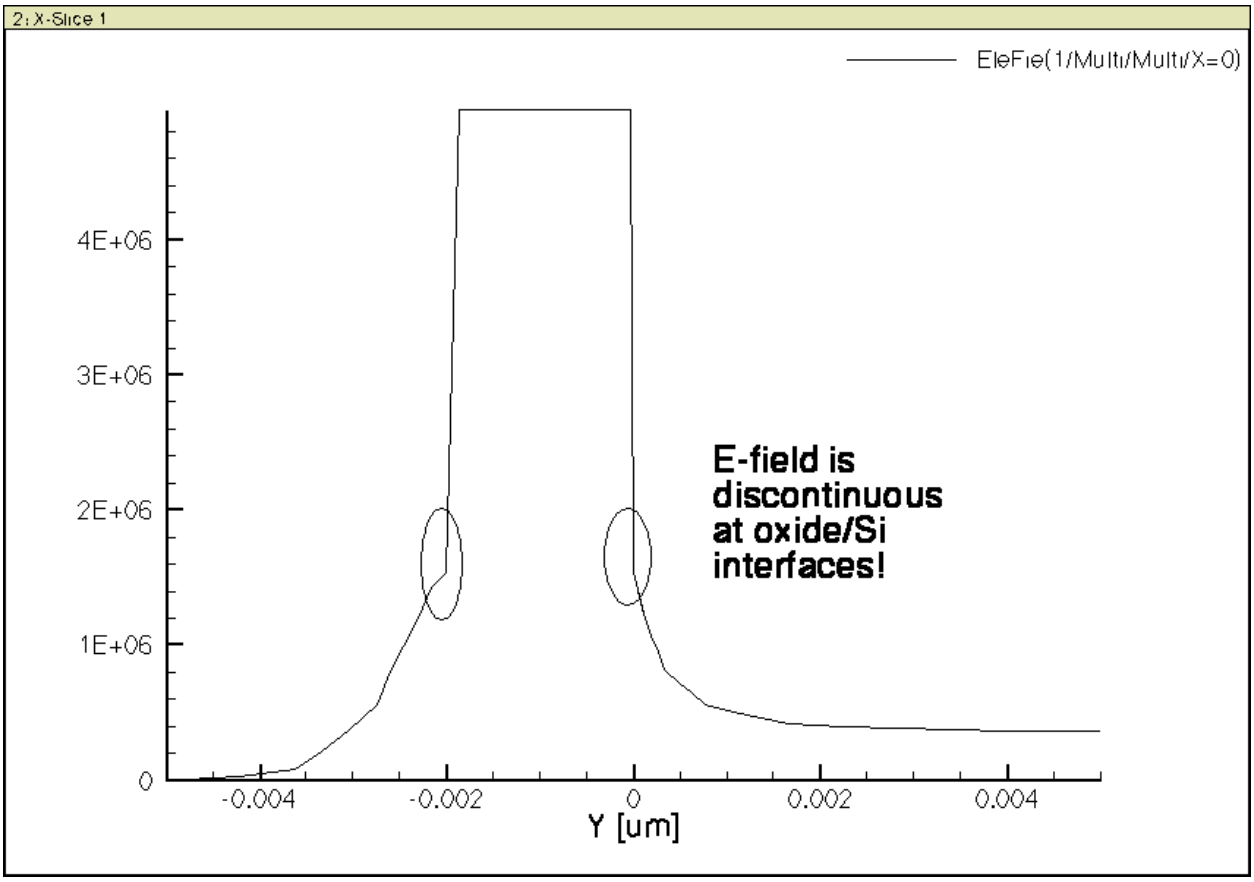


Figure 2.20: Vertical electric field distribution for all gate biases

At this point, I would like to remind you of the boundary condition for electric field at the gate oxide/Si interfaces. Assuming there is no interface charge, i.e. the interface is clean and free from interface charges, fixed or variable, the normal component of electric field should be continuous across the interface.

As dielectric constant of oxide is 3.9, while that of Si is 11.7, there is approximately a 3 times difference between the oxide field and the Si surface or interface field. To show this, I have made a zoomed in plot of the electric field in the oxide and near the two interfaces in figure 2.7.6:


2: X-Slice 1
EleFie(1/Multi/Multi/X=0)
4E+06
3E+06
2E+06
1E+06
0
-0.004
-0.002
0
0.002
0.004
Y [um]
E-field is
discontinuous
at oxide/Si
interfaces!

CHAPTER

THREE

MOSFET SPLIT CV SIMULATION

3.1 Objectives

- learn the basic flows of running device TCAD
- perform split CV simulation
- vary parameters of MOSFET structure to see impact on CV
- using split CV to understand poly depletion effect
- get a feel of computing via command lines
- get to know various scripts

3.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect

3.3 Overview

See lecture notes on mos capacitor and mosfet capacitors.

3.4 Files You Need

- the structure specification and mesh generation command file. There are two versions. You have used the version with graphical user interface to input structural parameters, which can also be downloaded here `mosfet.scm`

  The other version does not use the GUI. Instead, you will use a text editor to change the parameters. See the tutorial below. That file can be downloaded here `mosfet_fixed.scm`.

- In the mosfet_fixed.scm file, an inspection of the program will quickly show that the variables have the following meanings
  - w1: "Substrate Width"

  - w2: “Gate Width”
  - w3 “Spacer Width”
  - h1 “Substrate Height”
  - h2 “Oxide Thickness”
  - h3 “Spacer Height”
  - Nsub “Substrate Doping”
  - xj “S/D Junction Depth”

- the device simulation command file split_cv_des.cmd, which can also be downloaded here `split_cv_des.cmd`. It needs to be placed in the same directory as your other files, presumably, ~/mosfetcv, which was created earlier.
- the inspect command file for plotting AC simulated CV curves, cv_ins.cmd, which can also be downloaded `cv_ins.cmd`.

3.5 Video Tutorial

First download all the files above, place them in the mosfetcv folder.

Watch a 15 minute tutorial I prepared tonight. Do not have my good mic with me. So bare with the default laptop mic. It is audible.

Follow the tutorial. You should see CV curves pop up.

3.6 Steps

3.6.1 Meshing

To generate the mesh, first edit structural parameters in the mosfet_fixed.scm file. If you are on command line, working inside the mosfetcv folder already, just run

```
gedit mosfet_fixed.scm&
```

After editing, save the file. You do not have to close the editor as it is running in background.

Next, run

```
sde -l mosfet_fixed.scm
```

You will see sde working to generate the MOSFET mesh, and tecplot displaying the end result. You can examine the structure or doping etc inside tecplot. No need to save anything when asked. It has all been saved with commands inside the command file.

You can then close tecplot and sde programs.

3.6.2 Simulating

Examine split_cv_des.cmd file first. Set all voltages to desired values. For vp.dc, make it stop at 3V if you are using thin oxide just a few nm thick. You could modify the start gate voltage to say -3V too.

Then run

```
sdevice split_cv_des.cmd
```

You should see reporting of solution process. It should take about 1 minute cpu time to finish.

3.6.3 Plotting

Examine cv_ins.cmd file first. You do not have to completely understand it, but just a quick look will give you some useful idea about what it does. You may modify it for your purpose. You are not required to write such programs. See our notes about how and why we need to take the absolution values of certain capacitances. We talked about this in class too.

For thin gate oxide, set the vdd parameter to a value consistent with the vp.dc set above in the split_cv_des.cmd file. Then, run

```
inspect -f cv_ins.cmd
```

You should see some nice split CV plots.

CHAPTER

FOUR

MOSFET ID-VD OUTPUT CURVES SIMULATION AND PROBING OF INTERNAL PHYSICS

4.1 Objectives

- perform Id-Vd simulation
- learn how to save and load solutions for better convergence
- examine internal physics, particularly at the surface
- better understand what happens beyond "saturation"

4.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect

4.3 Overview

See lecture notes on 3-terminal and 4-terminal MOS devices.

4.4 Files You Need

Download the files from our class sites. Use the large file which has more up to date contents. If you home directory is small, download it to a PC first and pick files you need. You can take out the *.cmd files and *.par files, and run them to generate all other files you need.

Once you download the file, go to file browser, then right click `Extract here`.

Typically you should put it under your home folder for ease of access, then do extraction.

4.5 Meshing

To generate the mesh, first edit structural parameters in the halomosfet_sde.cmd file. In a terminal, cd to the source file folder, just run

```
gedit halomosfet_sde.cmd&
```

For now, you do not have to change anything. You should find the script is written such that you can modify structural parameters easily. After editing, save the file. You do not have to close the editor as it is running in background.

Next, run

```
sde -l halomosfet_sde.cmd
```

You will see sde working to generate the MOSFET mesh. You can then close the sde program. At this point, you can also use `tecplot_sv` to inspect the structure produced.

4.5.1 Halo Doping

One important feature to notice is the placement of halo doping, a pocket of relatively high concentration p-type dopants near the source and drain. This is important to limit the depletion thickness of the lateral n+ source/drain to p-channel junction, to reduce electrostatic coupling between the source and drain.

Inside tecplot_sv, do a 1D cut of the y-axis, specify a y location at the surface, then observe the plot. Due to numerical issues with tecplot_sv, you likely will run into trouble with not getting the entire distance along the x-axis plotted.

So I suggest that you use either `cut-along-boundary`, or simply just give a point right below surface, e.g. `y = 0.001`. The doping variation along the cut line should look like:

4.6 Simulating

Examine the idvd_des.cmd file first. As we explained in class, for better convergence, we do a Vg sweep at zero Vd first, save them. Then we load each Vg with zero Vd solution, and ramp up Vd. In this example, the source and body are both at zero, so you do not see body bias effect at the source end. This can be easily changed by modifying the body bias.

Then run

```
sdevice idvd_des.cmd
```

You should see reporting of solution process.

4.7 I-V Plotting and Unit of Current

Examine `idvd_ins.cmd` file first. You do not have to completely understand it, but just a quick look will give you some useful idea about what it does. You may modify it for your purpose. You are not required to write such programs.

Then, run

```
inspect -f idvd_ins.cmd
```

You should see a nice Id-Vd plot as follows:

You might wonder why the unit of current is A/um, not A. Good question indeed. This is a 2D simulation, we never gave the simulator any information on the dimension along the 3rd dimension. Real objects are 3D of course.

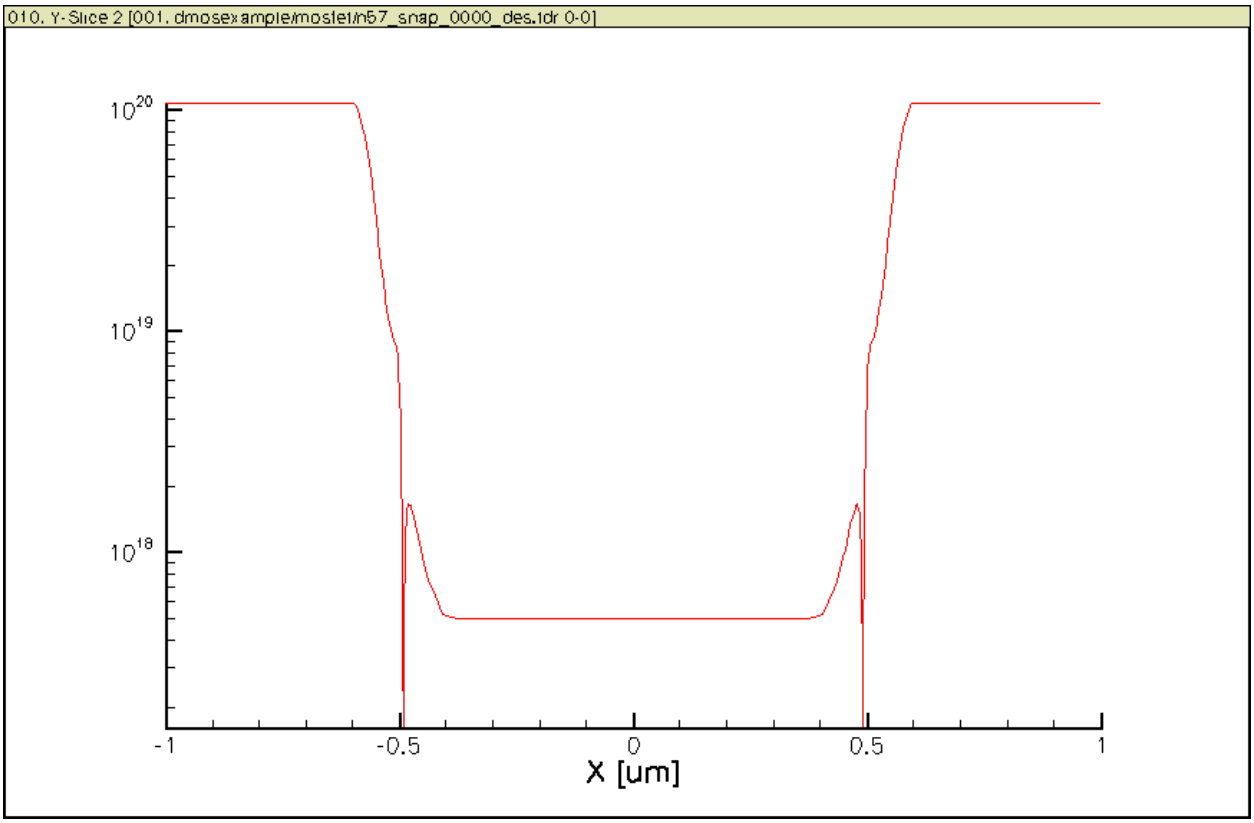


Figure 4.1: Doping variation along MOSFET channel at the surface showing halo doping.

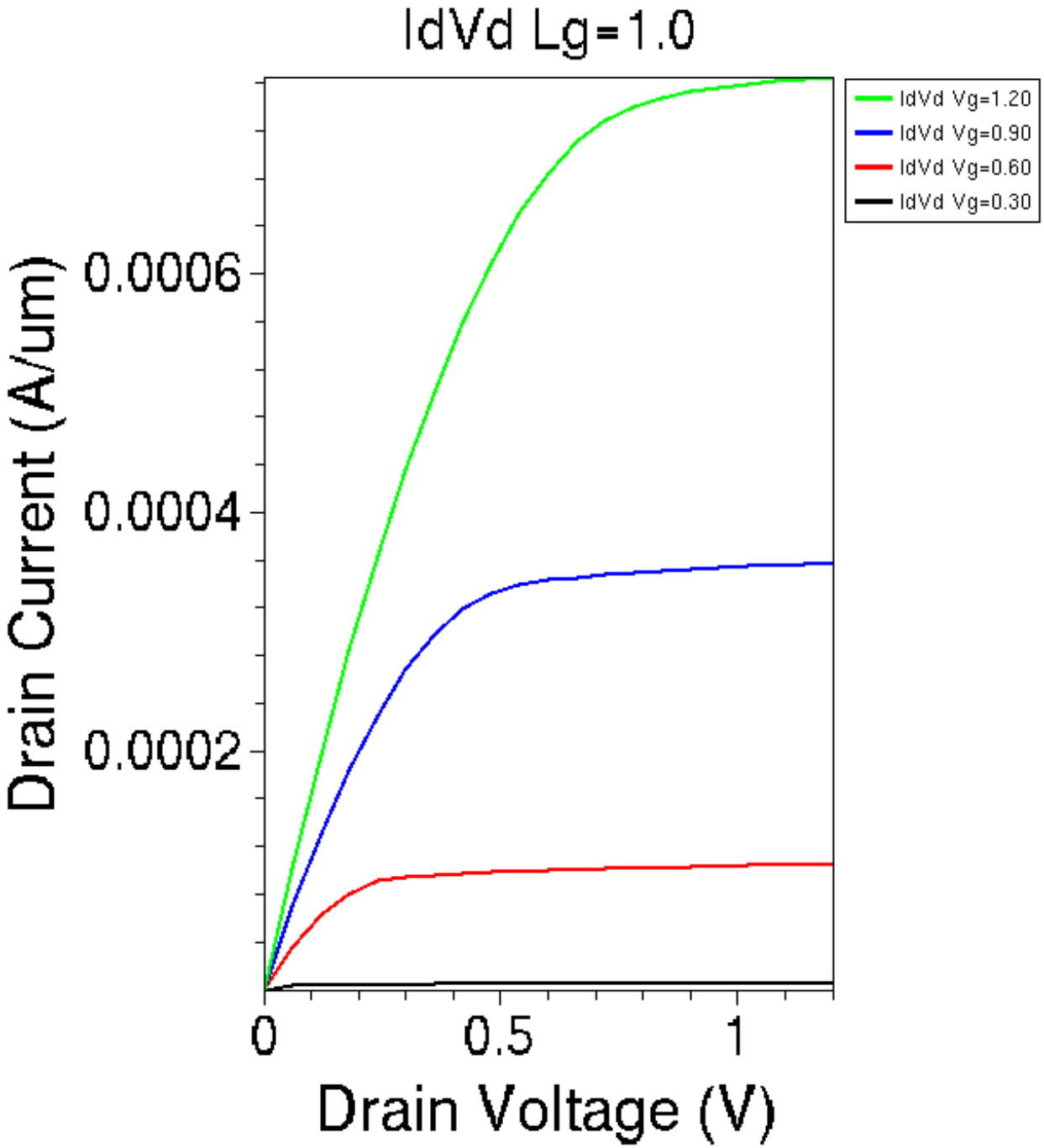


Figure 4.2: Simulated Id-Vd output family of curves.

By default, the simulator assumes a 1um width in the 3rd dimension. In our MOSFET case, it means the default channel width is 1um.

This of course can be changed. We will get to that later, or you can look up the manually to quickly find out how to change this.

Note: Think about what the unit of terminal resistance should be for a 2-D simulation.

4.8 Probing Internal Physics for Understanding

To probe internal physics, in a terminal, run

```
tecplot_sv
```

From reading the *des.cmd files, you should know what all the .tdr files are.

Make a y-cut at y=0.001 (0 strictly speaking, see above for why), the surface of Si. We can plot out Ec, Ev, electrostatic potential, Efn, Efp, phi_fn, phi_fp etc. along the cut.

4.8.1 High Vg

Recall the names of the highest Vg .tdr files. Load them in tecplot_sv as follows:

Remember to `Add` the files selected. Then `OK`.

Then go to `slicer -> Orthogonal Cut`. Select `Y` as normal direction. Deselect `Cut at mouse position`. Set `first cut at` to desired value, e.g. 0.001 here. Then `create cut`.

Click on the `rearrange frames` icon in the tool bar in the top left area of your tecplot window to focus on the cut plot. Disable the legends by clicking the `legends on /off` icon. Too busy a plot with them.

By default you see the doping. Now you can select other quantities of interest. Let us look at a few of them:

Potential and Field

As Vg is high, well over threshold voltage, with a Vds, current flows, potential drop varies along the channel. The amount of variation is equal to Vds (Vd here as Vs is zero).

At Vd=0.2V, 0.4V, 0.6V and 0.8V, all the Vds drops over the channel region. We can see an increase of lateral field with increasing Vd, as expected.

At Vd=1.0V and 1.2V, the potential in the channel barely changes, and stays about the same as at Vd=0.8V.

Hot Electrons and Reliability

The extra Vd drops in a vary narrow region near the drain, a sharp change of potential in a narrow region means high field. So keep this in mind, *lateral field is strongest near the drain after the so-called pinch off* occurs. This also leads to *hot electrons*, which can hit the oxide/Si interface, causing damages over time. This is one of the important considerations for long term reliability of transistors and CMOS ICs in general.

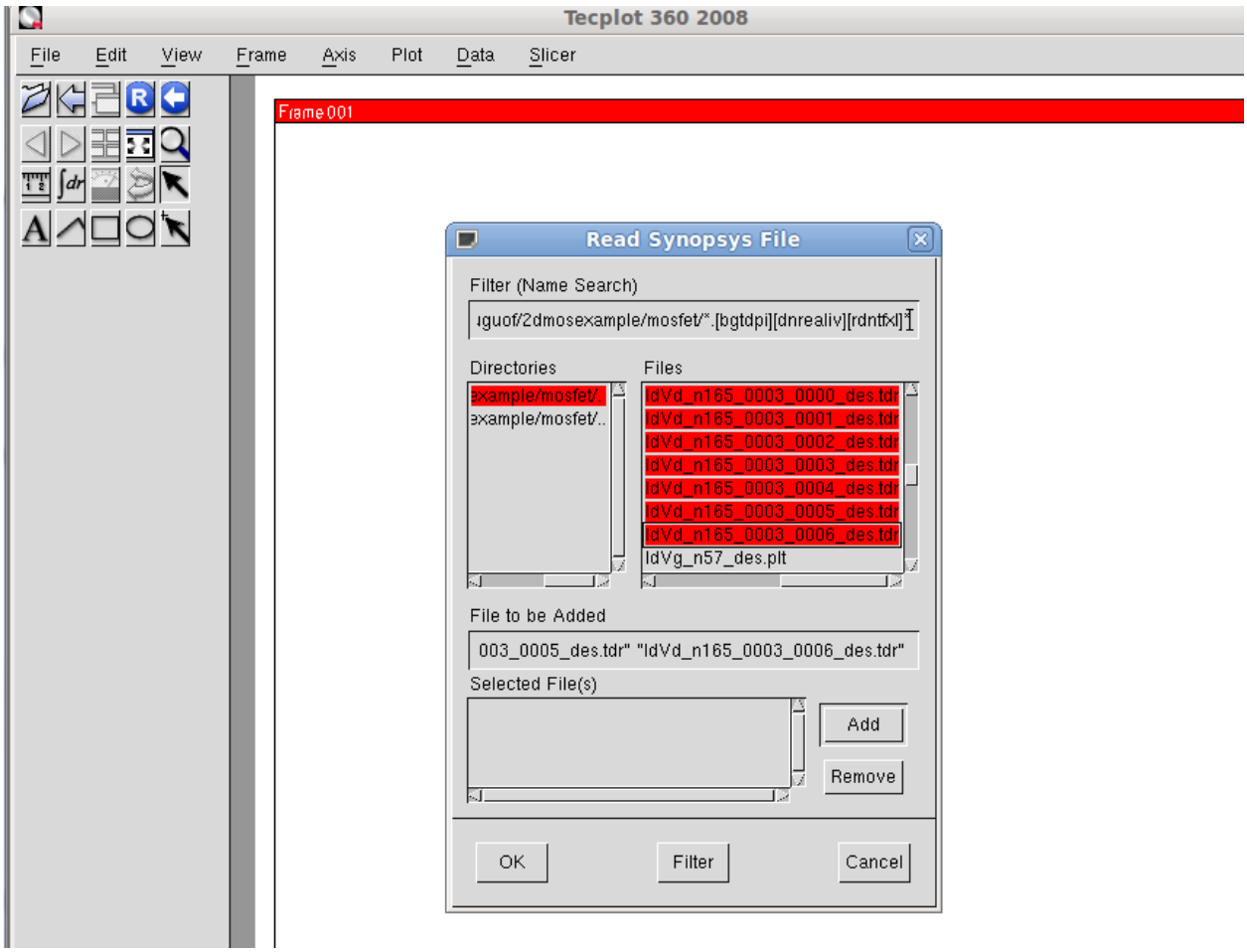


Figure 4.3: Loading high Vg (1.2V) tdr files for different Vd.

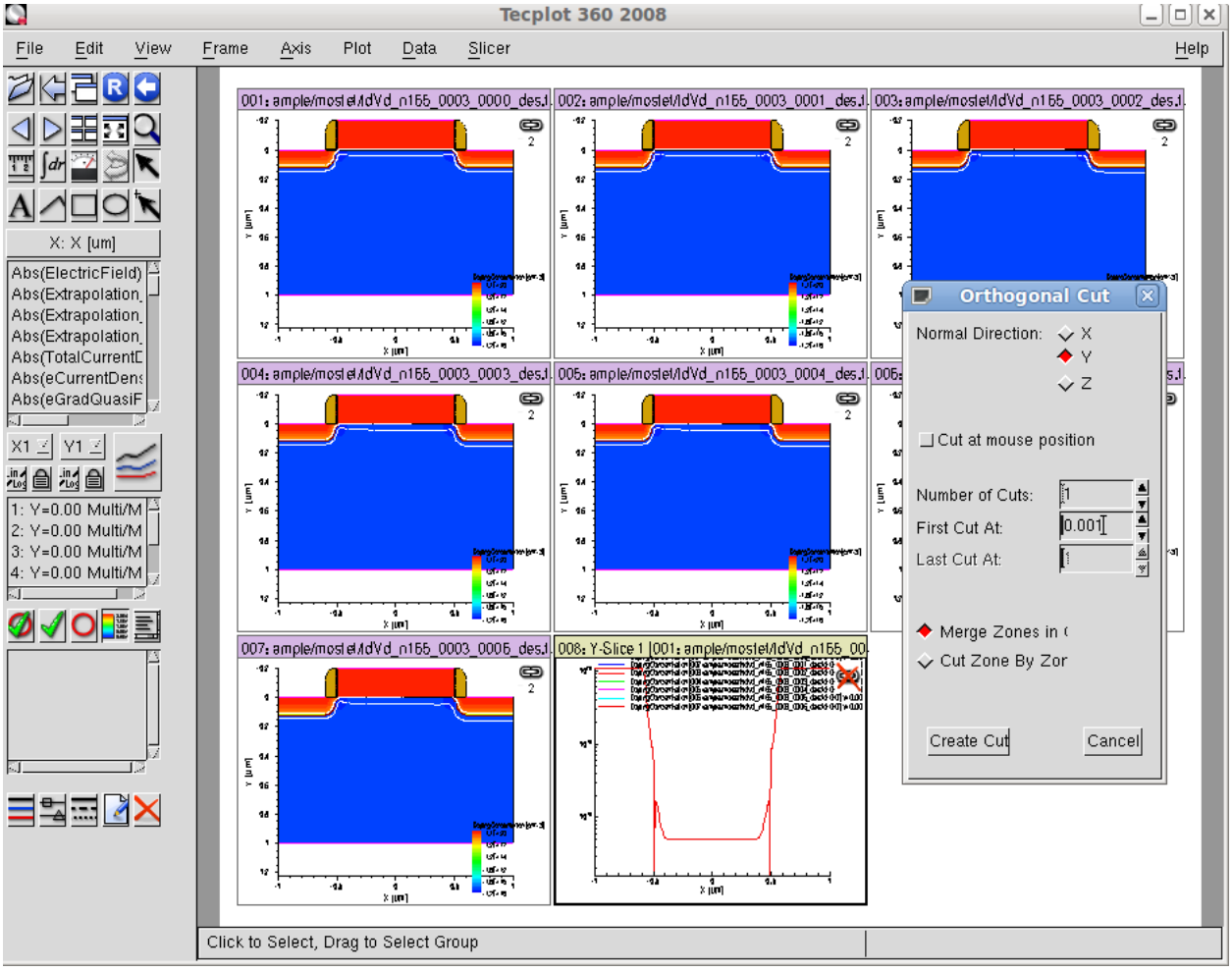


Figure 4.4: Making y-cut at desired location.

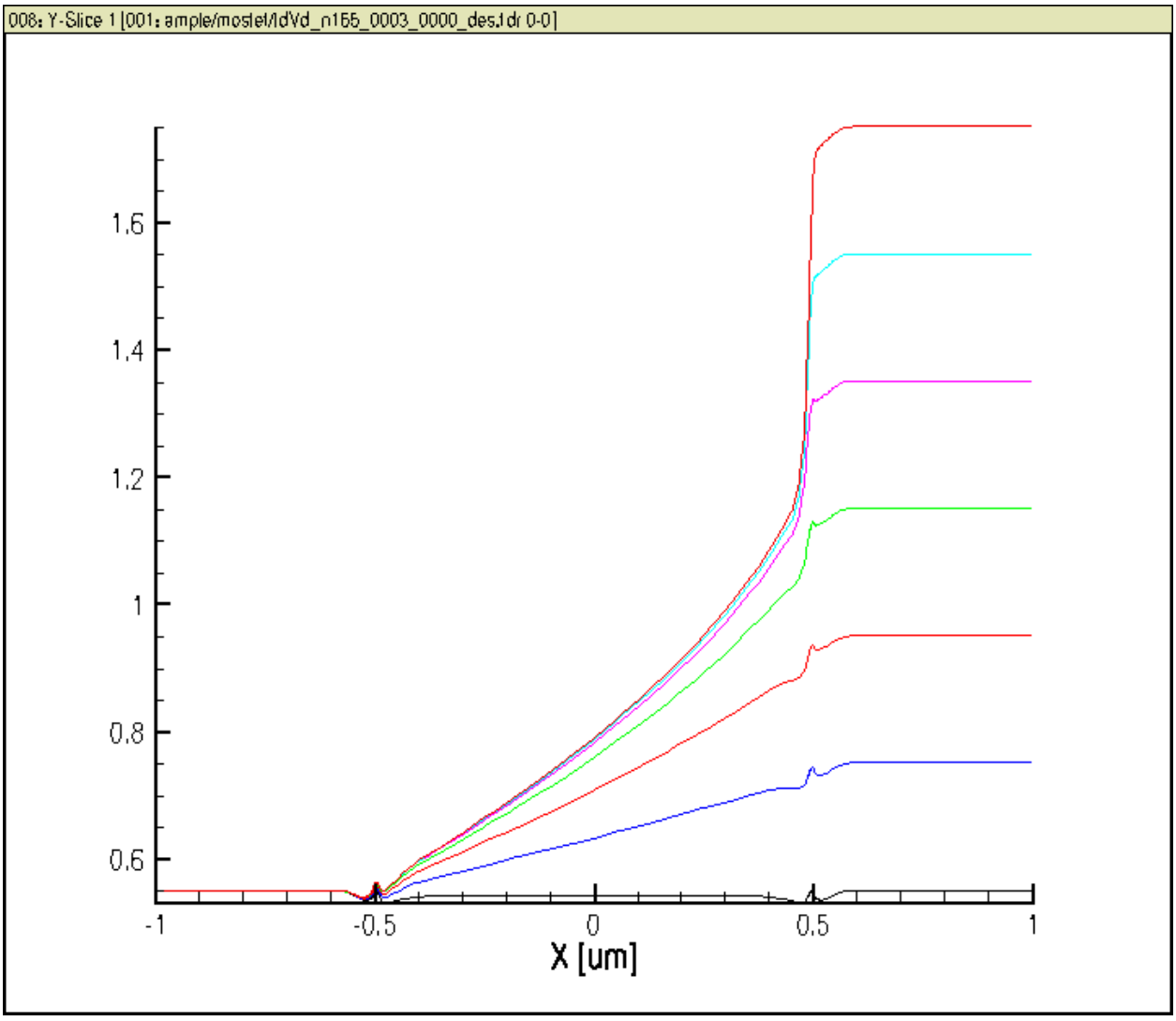


Figure 4.5: Surface potential along channel. Vg=1.2V. Vd from 0 to 1.2V.

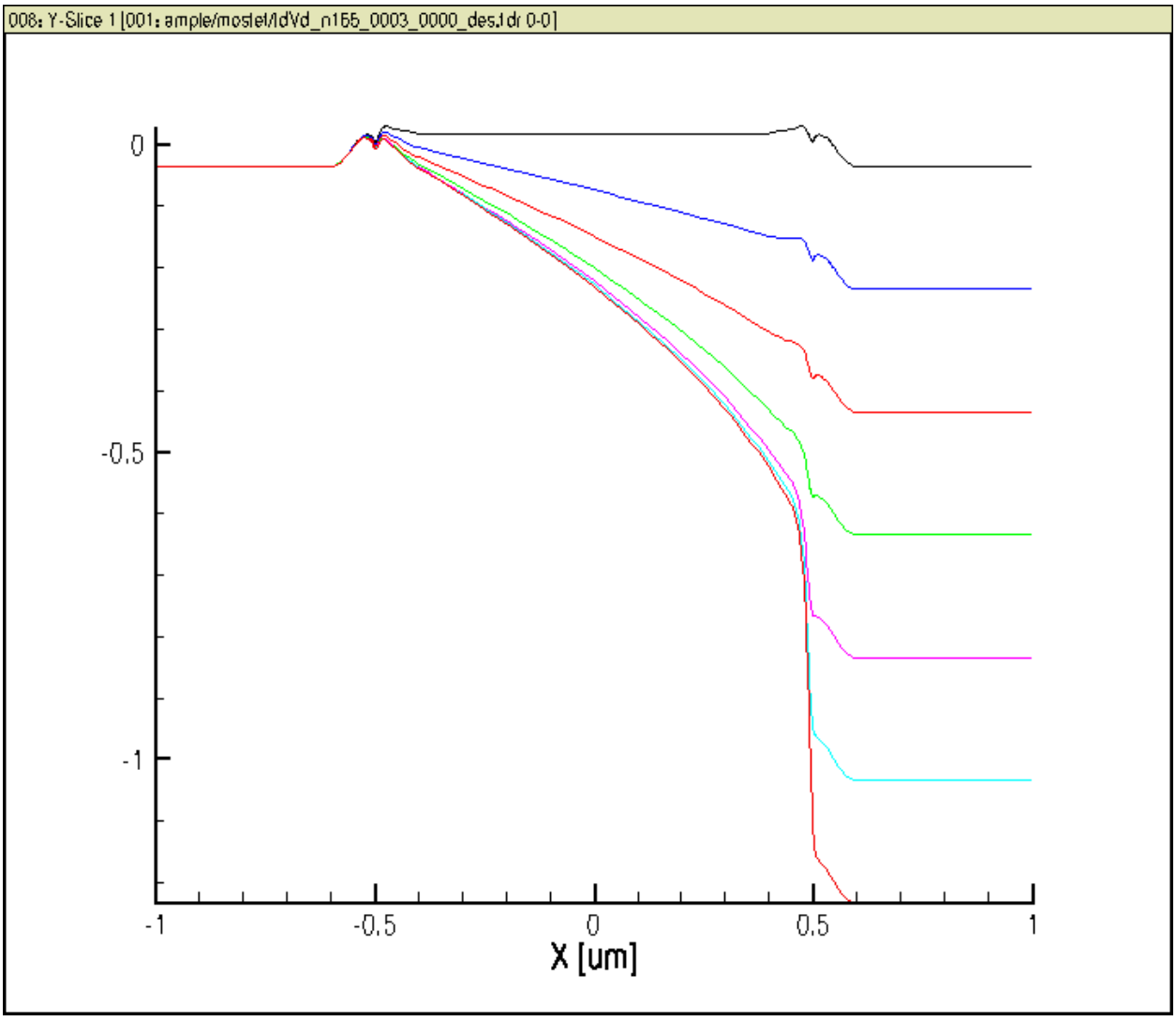


Figure 4.6: Ec along channel. Vg=1.2V. Vd from 0 to 1.2V.

Conduction Band Energy, Journey of Electron

This pretty much correspond to the surface potential plots. Electrons in the N+ source move from the source to the channel, and *drift down hill* towards thee drain.

Note the sharp band bending near the drain after Vd is above 0.8V.

The Ec in the active channel region no longer changes much after saturation.

Channel Length Modulation

The point at which we see rapid change of Ec moves towards the source a little bit wiht increasing Vd. This is known as *channel length modulation*. Electrical channel length decreases a bit with further increase of Vd after saturation, causing the drain current to increase slightly. In circuits, this will cause some output conductance, or a finite amount of output resistance, limiting the so-called open-loop voltage gain of a transistor amplifier.

eQuasiFermiPotential

Recall that at strong inversion, which is the case for the whole channel except for near the drain, electron quasi Fermi potential follows the electronstatic potential (surface potential roughtly is Vcb + a constant).

We see this behavior in this plot.

eGradQuasiFermiPotential

Gradient of eQuasiFermiPotential is driving force of current. We see this clearly in the plot.

eQuasiFermiEnergy

This is very much the opposite of eQuasiFermiPotential. Recall that Vcb is equal to Vds at the drain.

eDensity

This is a plot you really would like to not use the y=0.001 cut. I recommend using the `cut-along-boundary` function. Reason? Even a 0.001um separation can make large difference in eDensity. This, however, cannot be done automatically without using mouse to position the start and end points of the boundary. Also I have not found a way to repeat the cuts on all biases loaded. Below is a cut-along-boundary result for the Vg=1.2V and Vd=1.2V case:

Still we see a clear picture of how electron density varies along the channel, that is, it decreases from source to drain, as we well expect from our 3-terminal MOS theory.

The amount of decrease is dependent on Vd which sets Vcb. From source to drain, gate voltage is the same, but Vcb increases from 0 to Vds in this case. Electron density thus decreases from source to drain.

This is a linear scale eDensity plot.

Electron Velocity and Pinch-off

Note that we are using a highly simplified mobility model, with no velocity saturation. So you see very high velocity, particularly near the drain at high Vd.

Again, we should understand that electron density does not go to zero anywhere along the channel. If it does, or it does pinch off completely at any point, it would be an open circuit, and block current flow.

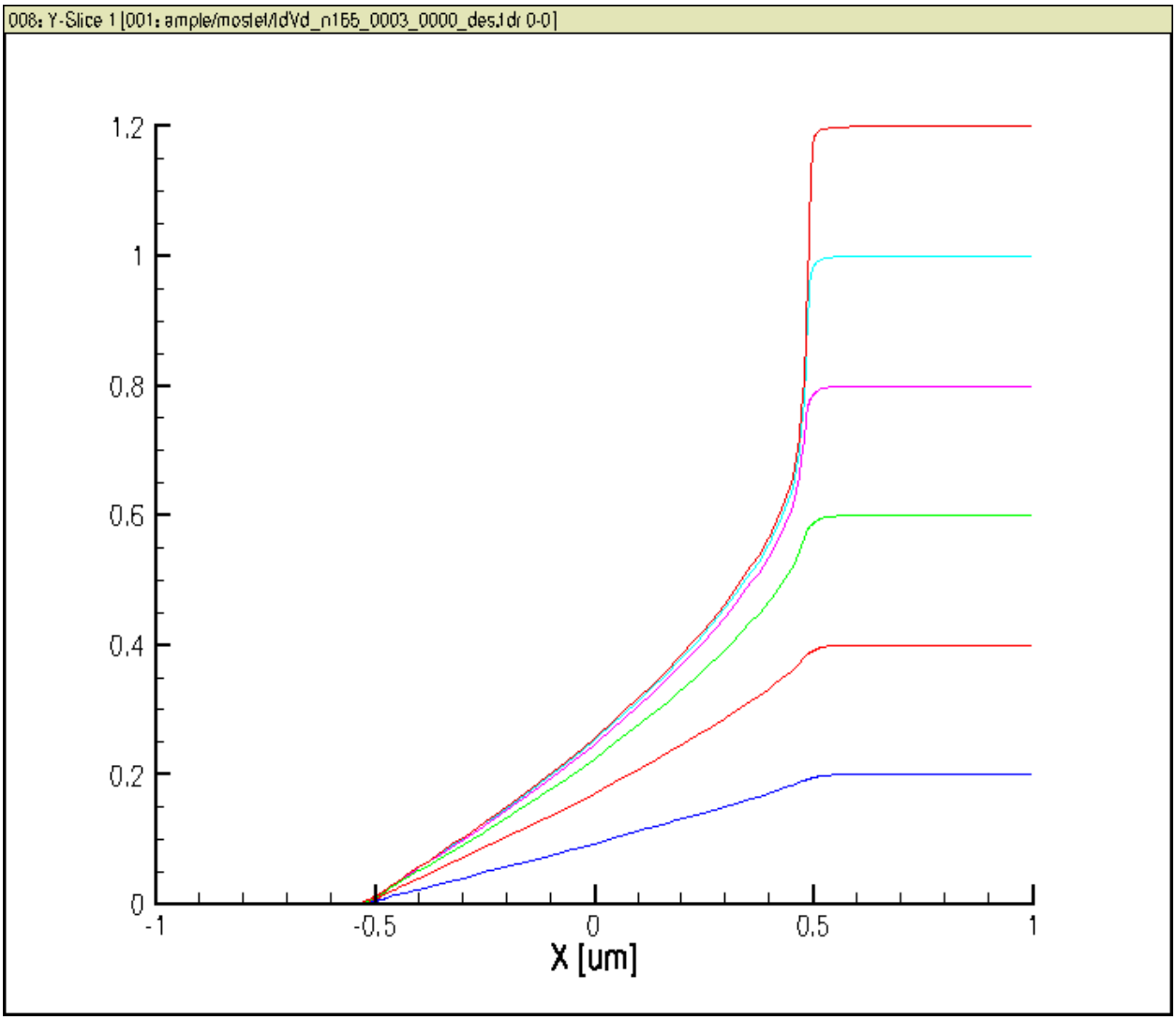


Figure 4.7: eQuasiFermiPotential along channel. Vg=1.2V. Vd from 0 to 1.2V.

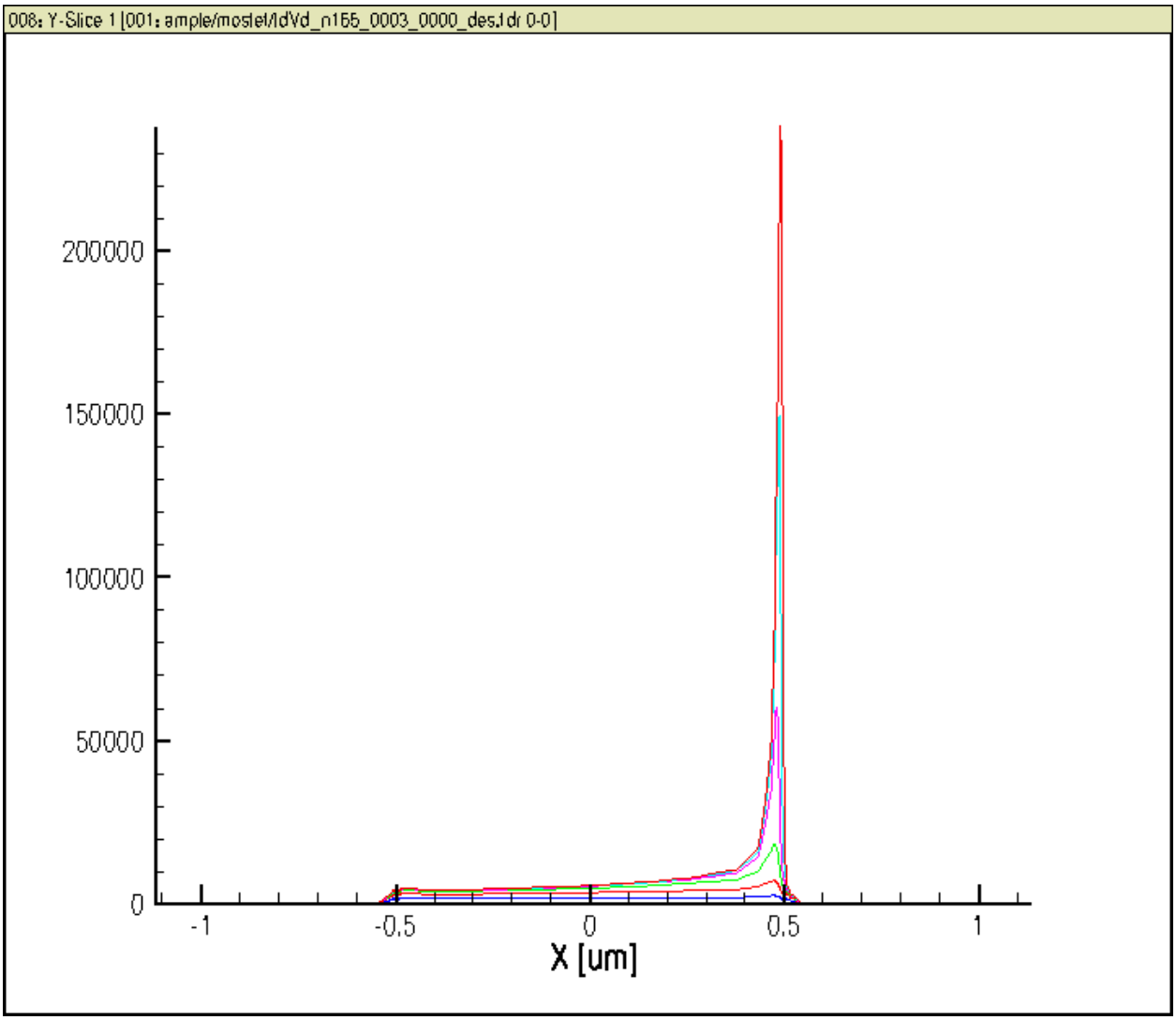


Figure 4.8: Gradient of eQuasiFermiPotential along channel. Vg=1.2V. Vd from 0 to 1.2V.

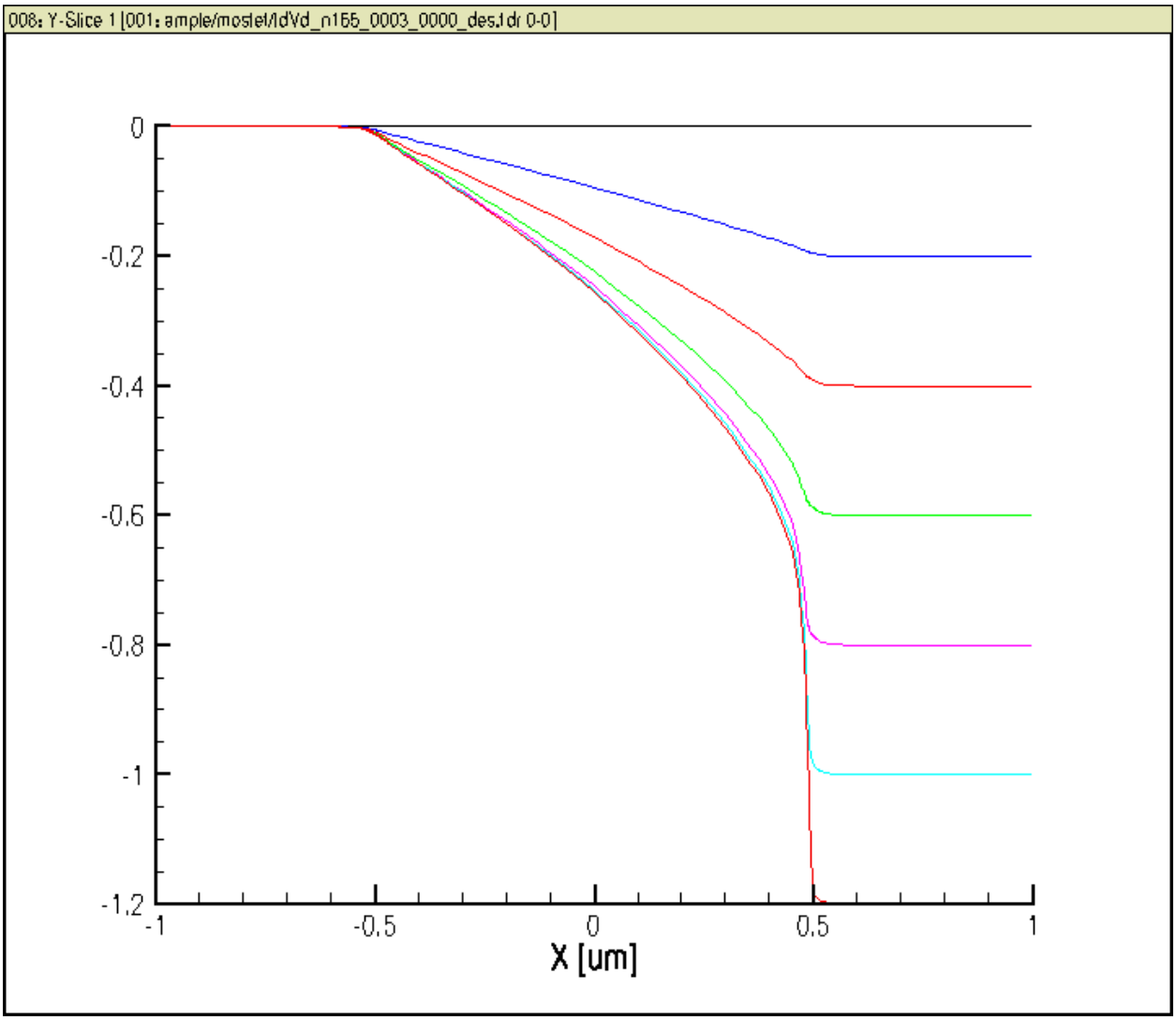


Figure 4.9: eQuasiFermiEnergy along channel. Vg=1.2V. Vd from 0 to 1.2V.

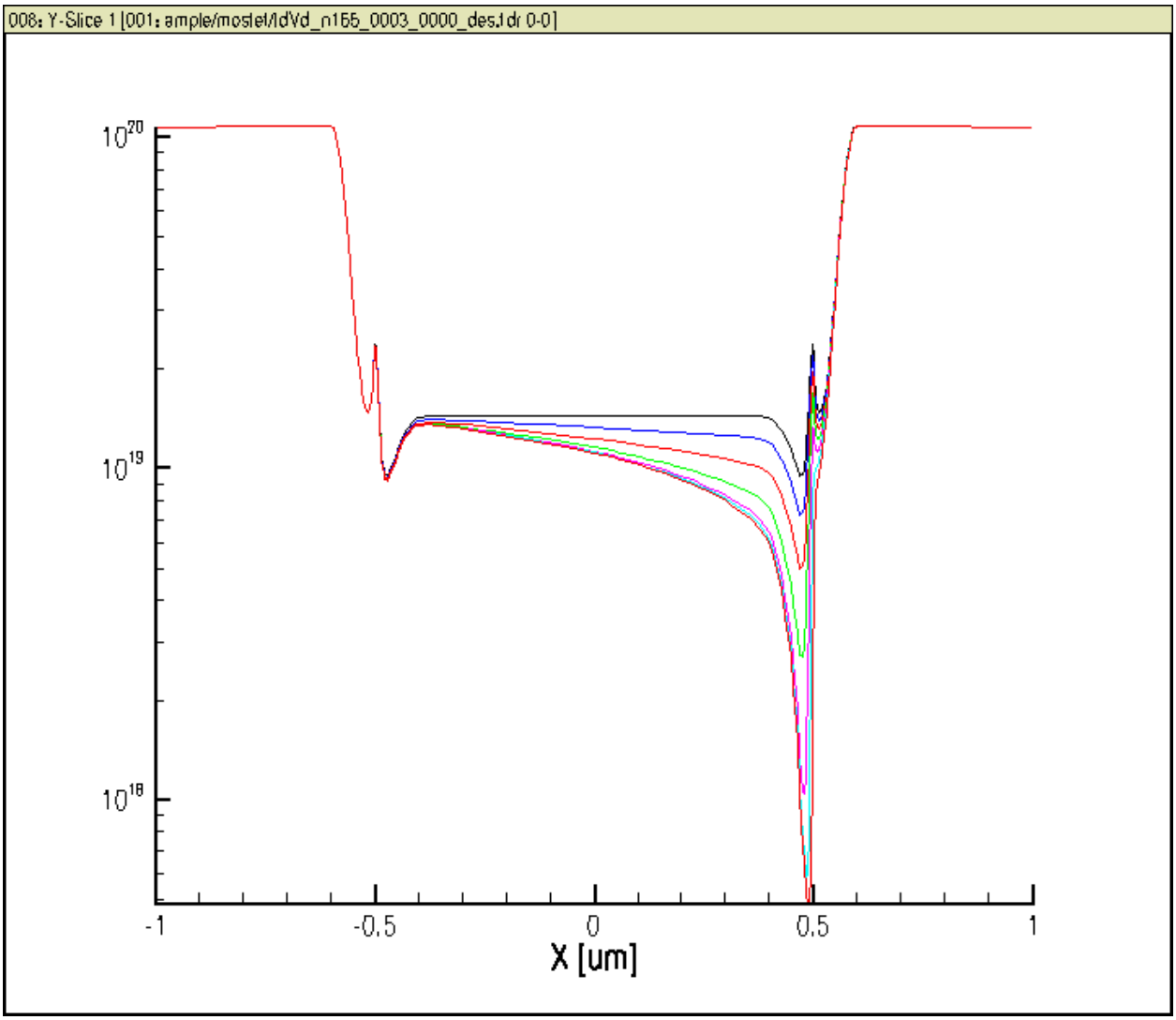


Figure 4.10: eDensity along channel. Log scale. Vg=1.2V. Vd from 0 to 1.2V.

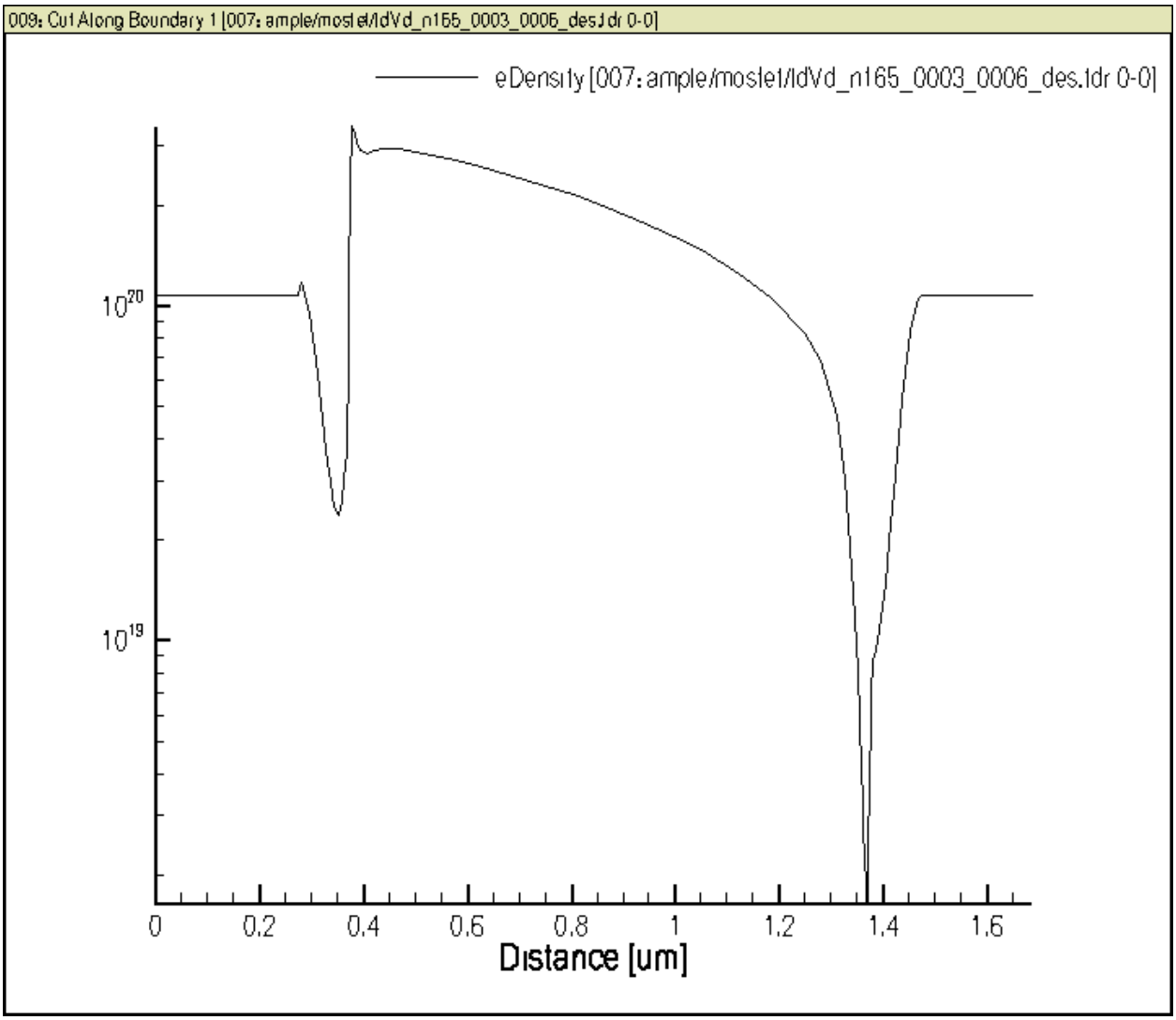


Figure 4.11: eDensity along channel obtained using cut-along-boundary. Log scale. Vg=1.2V. Vd from 0 to 1.2V.

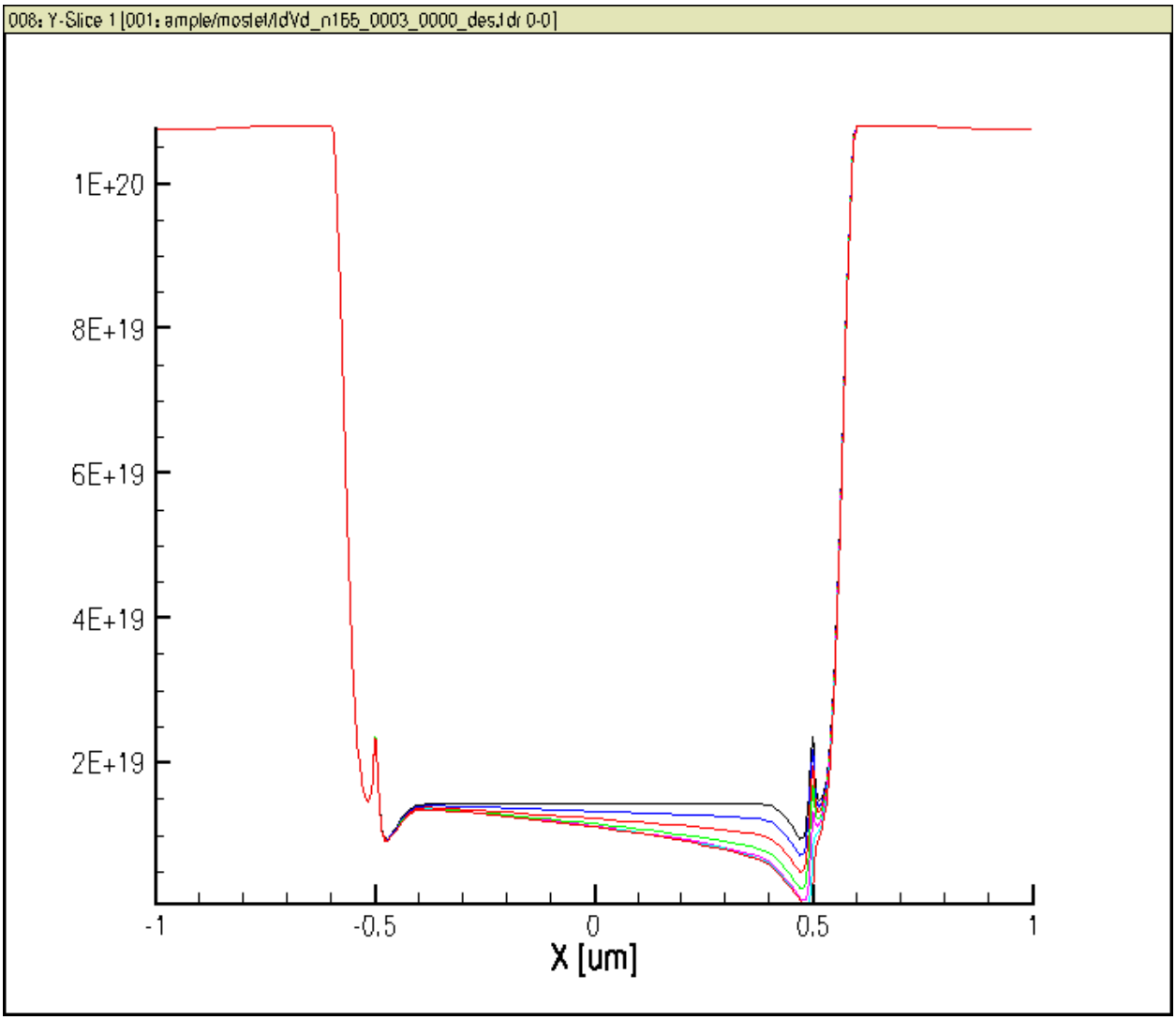


Figure 4.12: eDensity along channel. Linear scale. Vg=1.2V. Vd from 0 to 1.2V.

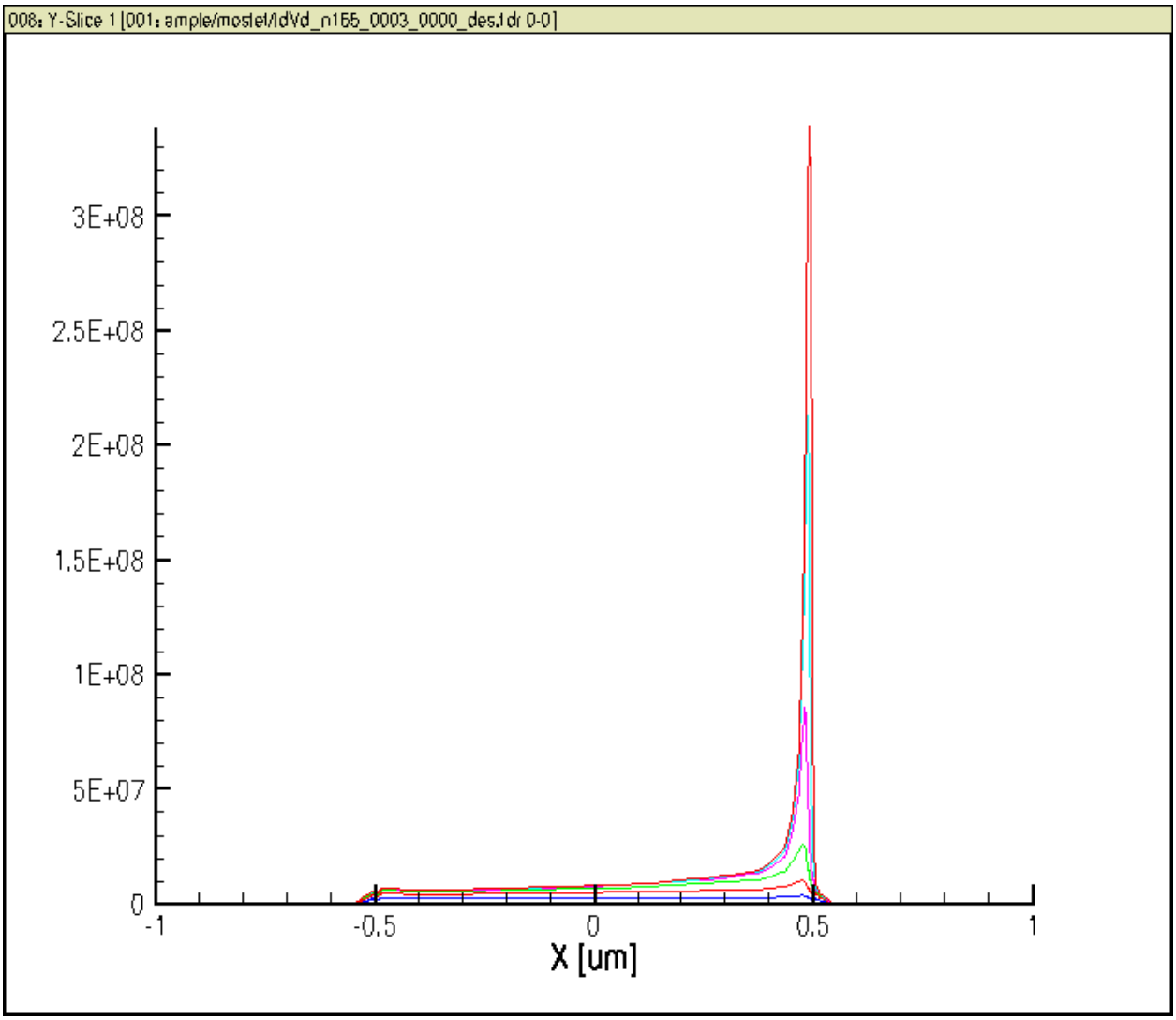


Figure 4.13: Electron velocity along channel. Linear scale. Vg=1.2V. Vd from 0 to 1.2V.

The so-called pinch-off is just a crude way of saying surface is no longer in strong inversion. Even in weak inversion there are finite electrons. A small amount of electrons can travel at very high velocity to still support a large current.

Again, note the high velocity and hence high kinetic energy of electrons near the drain, which can hit the oxide and create damages of the gate oxide/Si interface.

4.9 Homework

Pick one of the 3 problems below that is assigned to you, and complete it before Oct 14 class time. Bring your plots / files on a flash drive for sharing. You can work in groups if you wish, but each person needs to produce plots and analysis by himself/herself.

- (Olive, George, Wei-chuang, Jingshan) Follow the notes above, and repeat these plots and write down analysis as well as observations for Vg=0V and 0.6V. Associate the plots with your understanding and theory in any way you can. Save each plot as a .png file as well. Name each plot meaningfully.
- (Zhenzhen, Kun, Fang, Ruocan) Repeat these plots for the .tdr files generated from running `idvglin_des.cmd` and write down analysis.
- (David, Suraj, Pingye, Zhen) Repeat these plots for the .tdr files generated from running `idvgsat_des.cmd` and write down analysis.

Of course, you can choose to work on all of them as well.

CHAPTER

FIVE

PHYSICS SELECTION, EFFECT ON ELECTRICAL CHARACTERISTICS, AND TRACKING POSITION DATA

5.1 Objectives

- continue to develop physics understanding
- learn how to change physics models
- track internal physics quantities as a function of bias
- better understand how mobility changes affect I-V
- understand low-field mobility choices
- explore velocity saturation (high field mobility) impact
- go and find the manuals, download a copy. The directory is given below.

5.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect

5.3 Overview

See lecture notes on 3-terminal and 4-terminal MOS devices.

5.4 Files You Need

You can use the same files used previously for Id-Vd and Id-Vg. If you do not have them, download them. We will modify them a little bit as we demoed in class.

Download the files from our class sites. Use the large file which has more up to date contents. If you home directory is small, download it to a PC first and pick files you need. You can take out the *.cmd files and *.par files, and run them to generate all other files you need.

Once you download the file, go to file browser, then right click `Extract here`.

Typically you should put it under your home folder for ease of access, then do extraction.

5.5 Meshing

To generate the mesh, first edit structural parameters in the halomosfet_sde.cmd file. In a terminal, cd to the source file folder, just run

```
gedit halomosfet_sde.cmd&
```

For now, you do not have to change anything. Default gate length is 1.0um. You can change it to another value, e.g. 90nm to look at impact of gate length. You should find the script is written such that you can modify structural parameters easily. After editing, save the file. You do not have to close the editor as it is running in background.

Next, run

```
sde -l halomosfet_sde.cmd
```

You will see sde working to generate the MOSFET mesh. You can then close the sde program. At this point, you can also use `tecplot_sv` to inspect the structure produced.

5.6 Sdevice Commands

Recall that we have three *des.cmd files. Let us take the idvglin_des.cmd as an example.

```
File {
   * input files:
   Grid=   "n9_msh.tdr"
   Parameter="pp57_des.par"
   * output files:
   Plot=   "n57_des.tdr"
   Current="n57_des.plt"
   Output= "n57_des.log"
}

Electrode {
   { Name="source"    Voltage=0.0
   * Resistor=40
   }
   { Name="drain"     Voltage=0.0
   * Resistor=40
   }
   { Name="gate"      Voltage=0.0 }
   { Name="substrate" Voltage=0.0 }
}

Insert = "PhysicsSection_des.cmd"
Insert = "PlotSection_des.cmd"
Insert = "MathSection_des.cmd"

      CurrentPlot{
```

```
                eMobility((0 0.0000001))
                ElectricField/Vector((0 0.0000001))
                evelocity((0 0.0000001))
                ElectrostaticPotential((0 0.0000001))
                eDensity((0 0.0000001))
                }


Solve {
*- Creating initial guess:
    coupled(Iterations=100)  {Poisson
 }
   Coupled { Poisson Electron
 }
   Coupled { Poisson Electron Hole
 }

*- Ramp to drain to Vd
   Quasistationary(
      InitialStep=1e-2 Increment=1.35
      MinStep=1e-5 MaxStep=0.2
      Goal { Name="drain" Voltage=0.05
 }
   ){ Coupled { Poisson Electron Hole
 } }

*- Vg sweep
   NewCurrentFile="IdVg_"
   Quasistationary(
      DoZero
      InitialStep=1e-3 Increment=1.5
      MinStep=1e-5 MaxStep=0.04
      Goal { Name="gate" Voltage=1.2
 }
   ){ Coupled { Poisson Electron Hole
 }
      CurrentPlot(
                Time=(Range=(0 1) Intervals=30)
                  )
*     plot(FilePrefix="n57_snap" Time=(Range=(0 1) intervals=6) NoOverwrite)
   }
}
```

5.6.1 Physics Selection

The statement

```
Insert = "PhysicsSection_des.cmd"`
```

inserts the file `PhysicsSection_des.cmd` whose content is:

```
Physics{
*   EffectiveIntrinsicDensity( OldSlotboom )
}

Physics(Material="Silicon"){
```

```
   Mobility(
*       PhuMob
*       eHighFieldSaturation( GradQuasiFermi )
*       hHighFieldSaturation( GradQuasiFermi )
*       Enormal
   )

*  Recombination(
*    SRH( DopingDep )
*    Band2Band
* )
}
```

Note that the * at the line beginning is for commenting out the line. So by default, in the mobility section, none of the options is active.

So default simulation will use a constant mobility. We can remove the * one by one to produce difference combinations. For now, let us still keep the high field saturation turned off. We will turn on *Phumob* only, and then turn on *Phumob* and *Enormal*.

```
Physics(Material="Silicon"){

   Mobility(
       PhuMob
*      eHighFieldSaturation( GradQuasiFermi )
*      hHighFieldSaturation( GradQuasiFermi )
*      Enormal
   )

*  Recombination(
*    SRH( DopingDep )
*    Band2Band
* )
}
```

and

```
Physics(Material="Silicon"){

   Mobility(
       PhuMob
*      eHighFieldSaturation( GradQuasiFermi )
*      hHighFieldSaturation( GradQuasiFermi )
       Enormal
   )

*  Recombination(
*    SRH( DopingDep )
*    Band2Band
* )
}
```

We talked about the meanings of these choices in class. You can find more information in the sdevice_ug.pdf manual under /linux_apps/synopsys/v2.5/sentaurus/tcad/E-2010.12/manuals/PDFManual/data.

You may want to create a new folder or sub-folder for different options, as files with the same names will be overwritten. Instead of changing all file names in your command files, you can simply create a new folder with names like `muvertical` or `phumob` to indicate what the folder is about.

Of course, do not forget about changing the content after you create or duplicate the folder content via copy and paste.

5.6.2 Tracking Data at a Position Varying with Bias

A very advanced option in running TCAD simulation is to track how the internal physics data, e.g. mobility, or velocity, at a given position changes with bias.

This can be very useful for insight. For instance, one can look at how eVelocity changes with high field saturation on and off at different points along the channel. One will see different degree of velocity saturation.

In our case, let us plot out a few quantities at the center of the channel between source and drain, at the surface. As y=0 is not so good a choice, because it can be treated as inside the oxide by the simulator, let us use y=1e-7 instead. It ensures we are going to be looking at a point really close to surface, but inside Si.

This is achieved above by these lines:

```
CurrentPlot{
            eMobility((0 0.0000001))
            ElectricField/Vector((0 0.0000001))
            evelocity((0 0.0000001))
            ElectrostaticPotential((0 0.0000001))
            eDensity((0 0.0000001))
            }
```

It is very *important* to use parentheses to specify the coordinate, which is given as one to three (depending on device dimensions) numbers in parentheses. In this case, we are running 2-D simulation. So we are asking the program to give us required quantities at x=0, y=0.0000001.

The parentheses distinguish coordinates from node numbers, which you can find out using tecplot_sv. Do not worry about it for now. Just use coordinates. The advantage is that no interpolation is required. When specifying coordinates, the program will have to do interpolation using values at the nearby nodes in your grid / mesh.

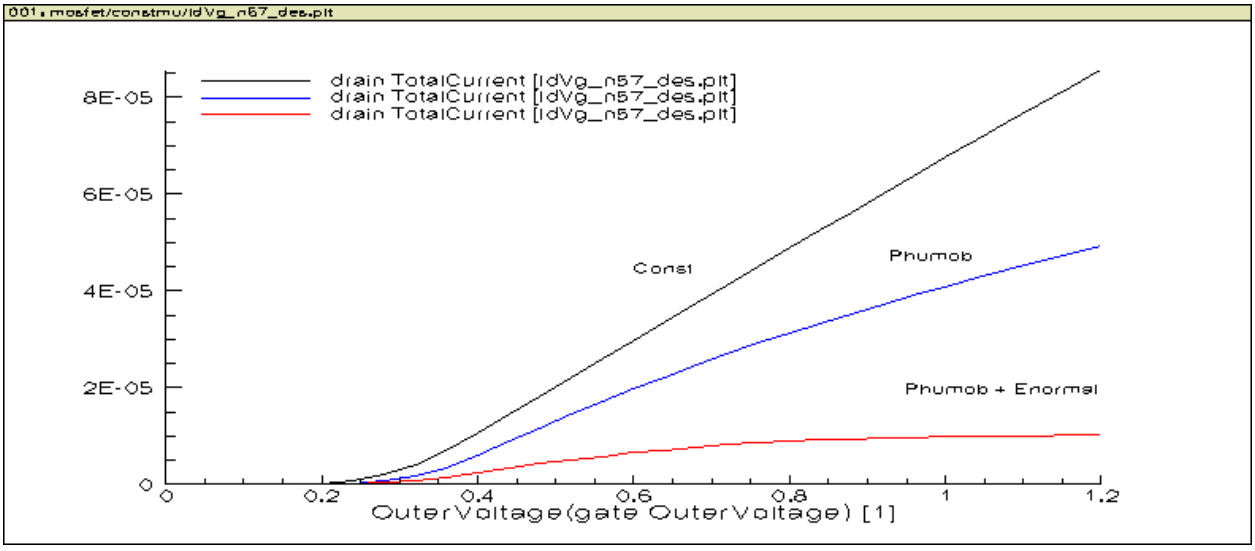


Figure 5.1: Mobility effect on Id-Vg linear characteristics. L=1um.

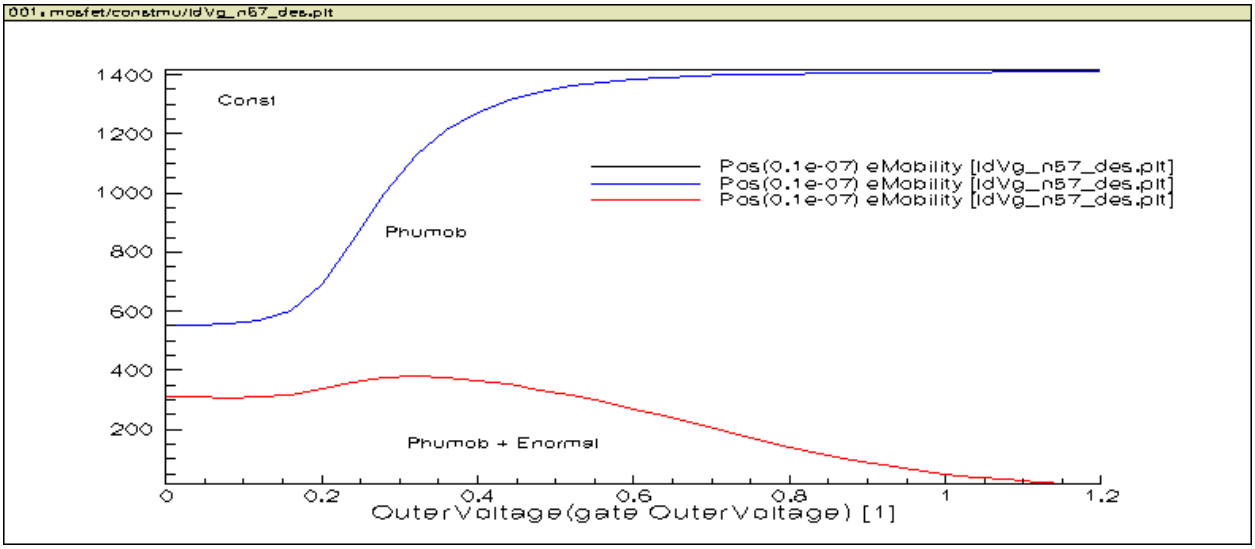


Figure 5.2: Mobility-Vg. L=1um. Think about how this will affect Id-Vg. Does it match the Id-Vg difference between different mobility choices? Why? Hint: also look at the lateral variation of mobility as we discussed in class.

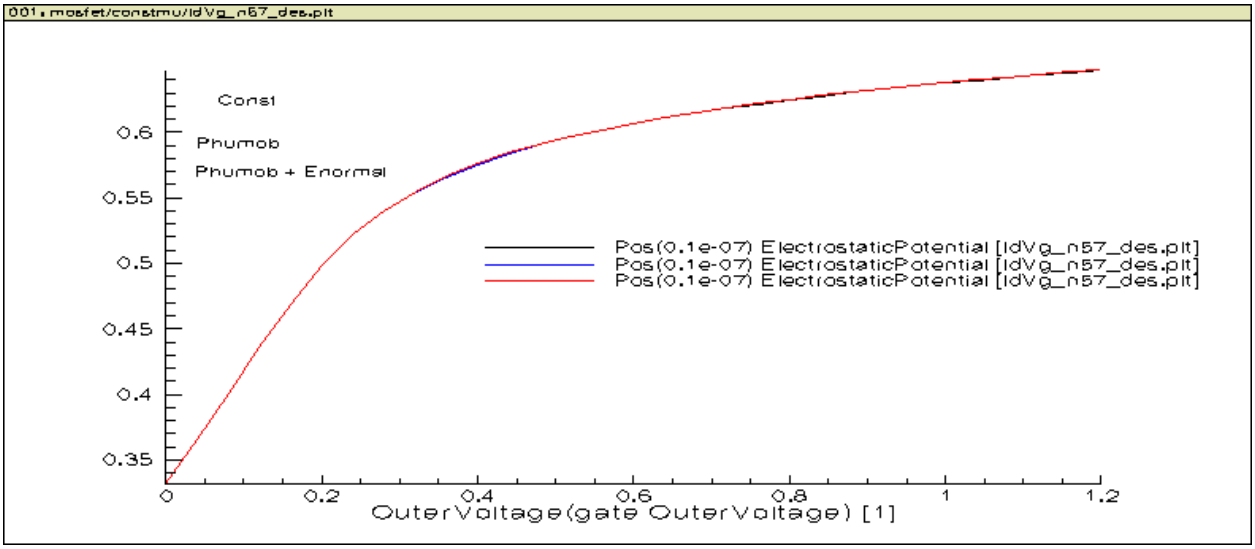


Figure 5.3: Phi_s-Vg linear characteristics. L=1um.

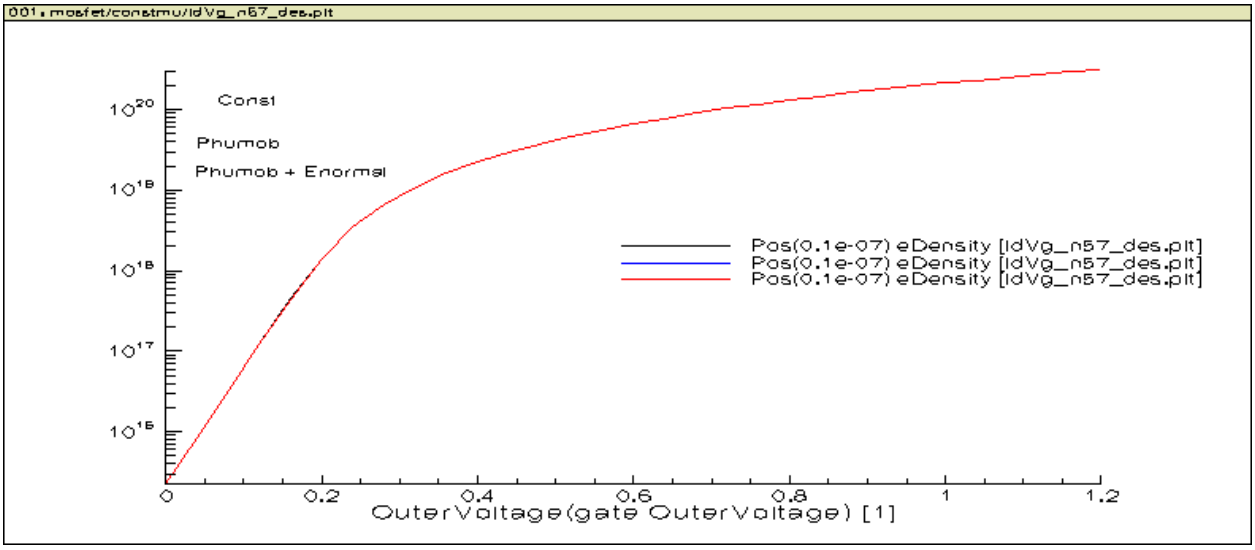


Figure 5.4: Mobility effect on n_s-Vg linear characteristics. L=1um.

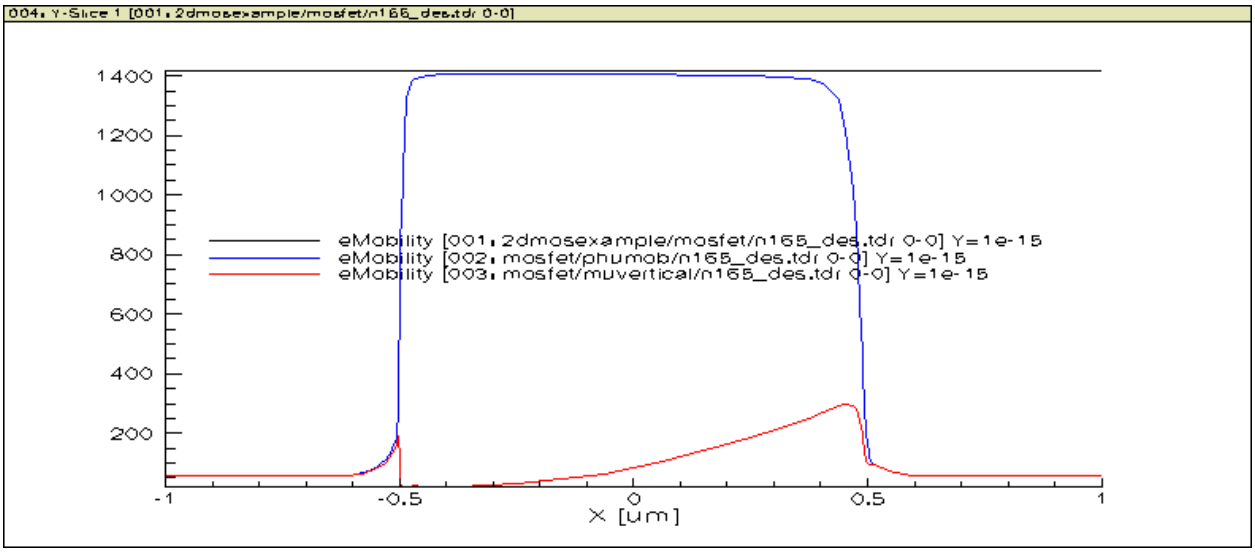


Figure 5.5: Mobility vs x comparison. L=1um. Think about how this will affect Id.

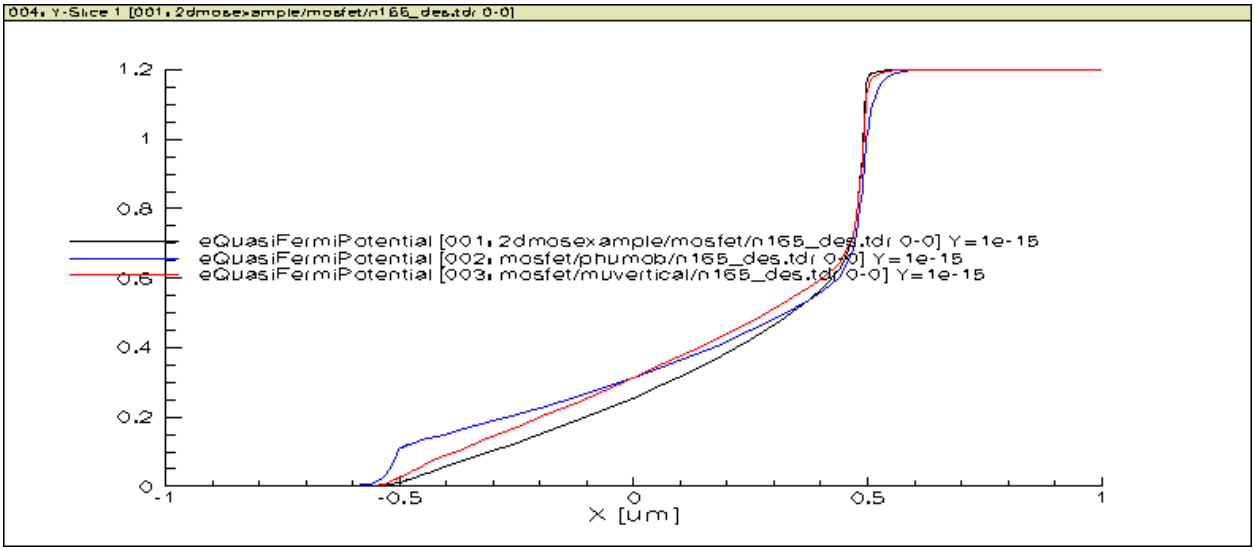


Figure 5.6: eQuasiFermiPotential vs x comparison. L=1um. Think about why.

5.7 Homework

Start with this. We will turn this into a midterm project next week.

- Try a 45nm, 65nm, or 90nm device with the mobility choices given above. Keep high field saturation turned off for now. Compare Id-Vg linear, saturation and Id-Vd curves.
- Turn on and off the two high field saturation switch with all other options in Mobility section turned on. Do this for 90nm and 1um. Compare Id-Vd curves by overlaying them. Do this one gate voltage at a time.
- Try to track lateral field (Ex) and velocity at the center of channel at surface as a function of Vd at the highest gate voltage.

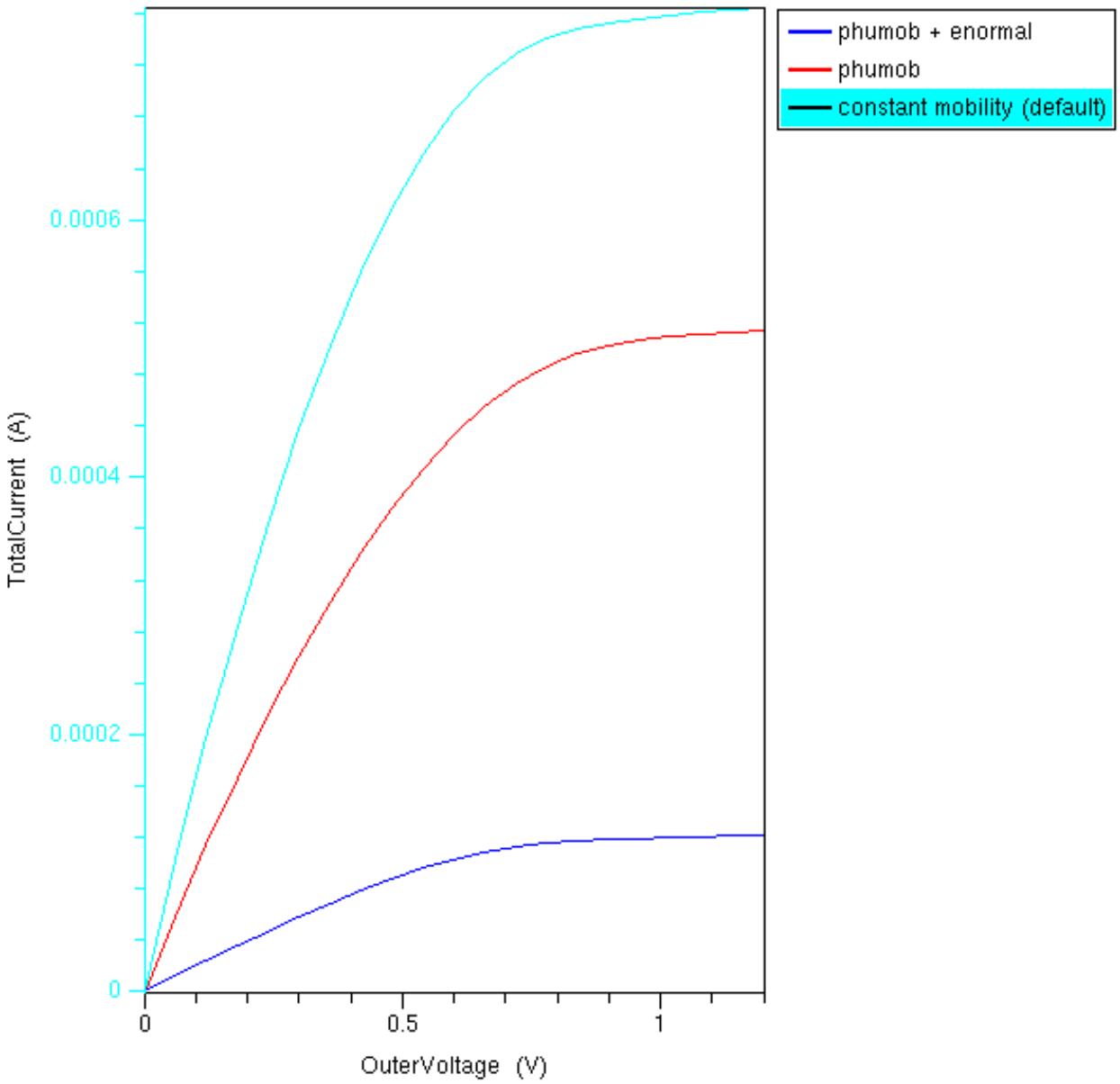


Figure 5.7: Mobility effect on Id-Vd characteristics. L=1um.

CHAPTER

SIX

VELOCITY SATURATION EFFECT EXPERIMENTS (45NM - 2UM)

6.1 Objectives

- set up your .bashrc to use Sentaurus tools 2011.9
- learn how to load in Sentaurus Work Bench (swb) Project
- learn how to use swb
- examine effect of velocity saturation on Ids-Vds, Vdsat, Idsat
- better understand what happens beyond “saturation”

6.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect
- swb

6.3 Overview

See lecture notes on 3-terminal and 4-terminal MOS devices.

6.4 Getting the files

First, log on to niu003.eng.auburn.edu.

Then in a command terminal, run

```
cd
```

This ensures you will change to home directory. Then, run

```
/scratch/getfiles.sh
```

This will copy a file physics.gzp to your home folder, and update your .bashrc so that you can now use the Sep 2011 version of the tcad tools.

As this is an existing window, we need to source the .bashrc file. Just run

```
source .bashrc
```

To launch swb, run

```
swb
```

6.5 Loading in a swb project

To load in the physics.gzp project, first look for the file physics.gzp as shown below in swb's file browser:

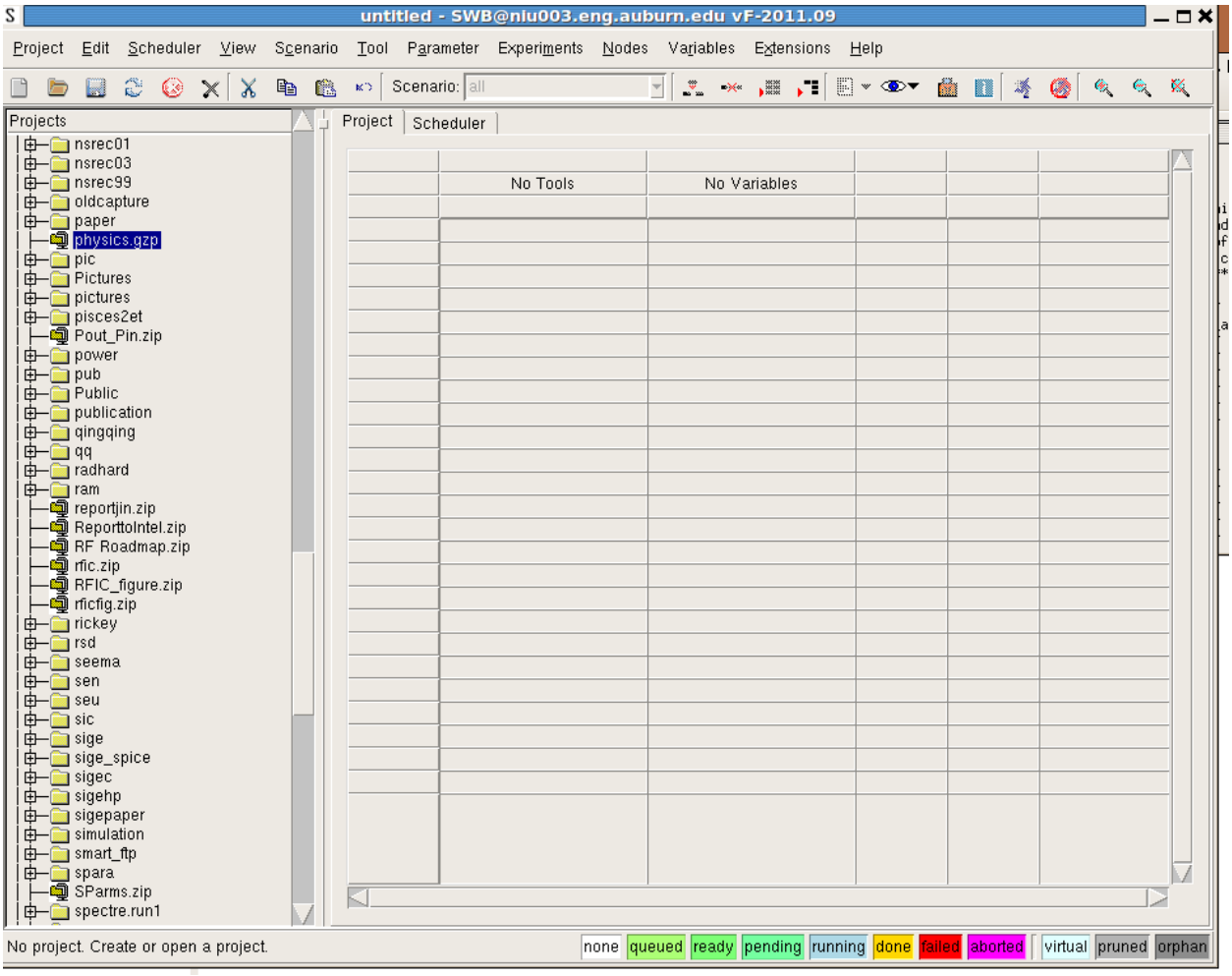


Figure 6.1: locate project archive

Right click the file, select `open`, you should see:

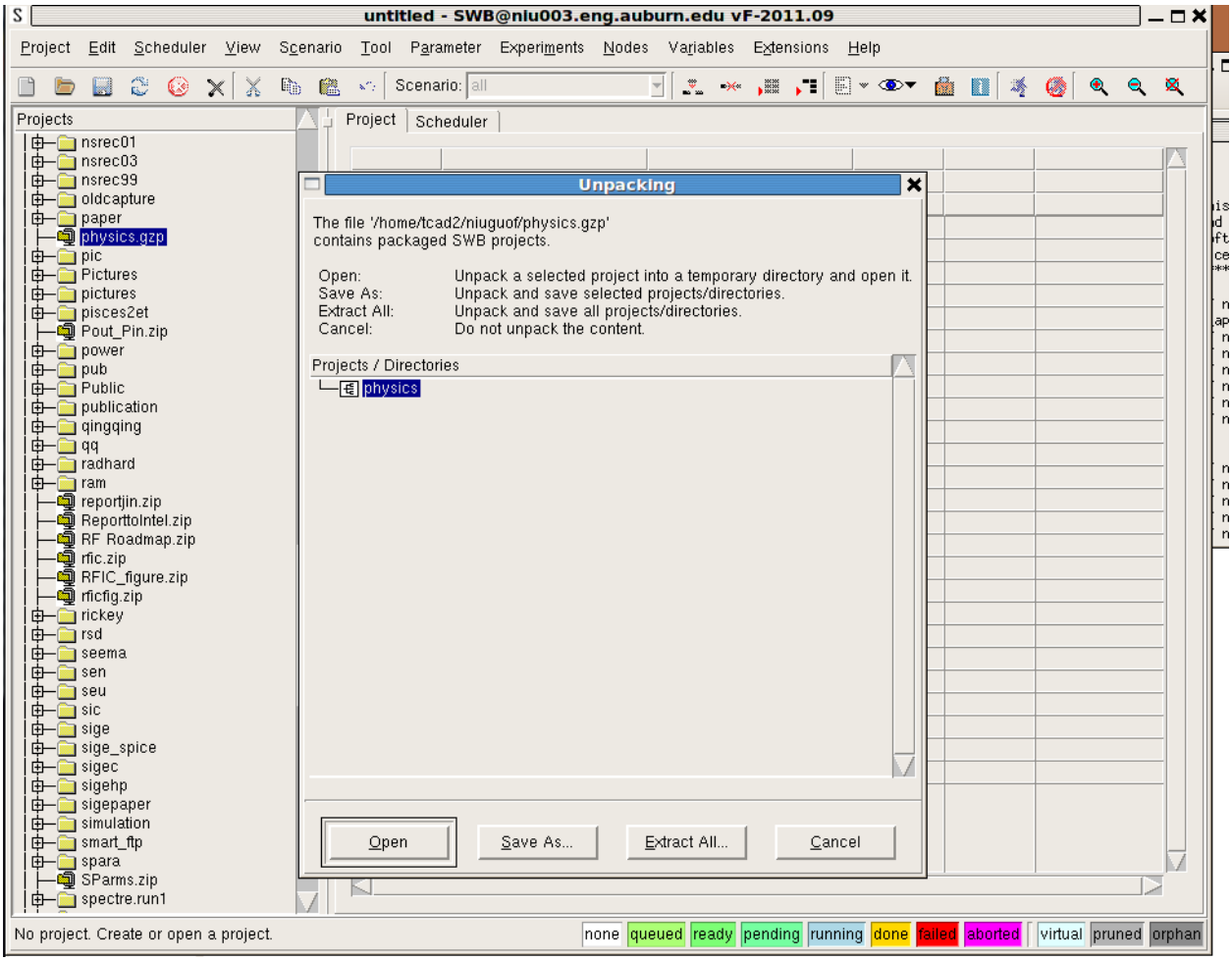


Figure 6.2: extract project archive

Select `extract all`, then create new folder or just use default to place the project.

Wait a little while. It is about 25 MB zipped, so big. You shall see:

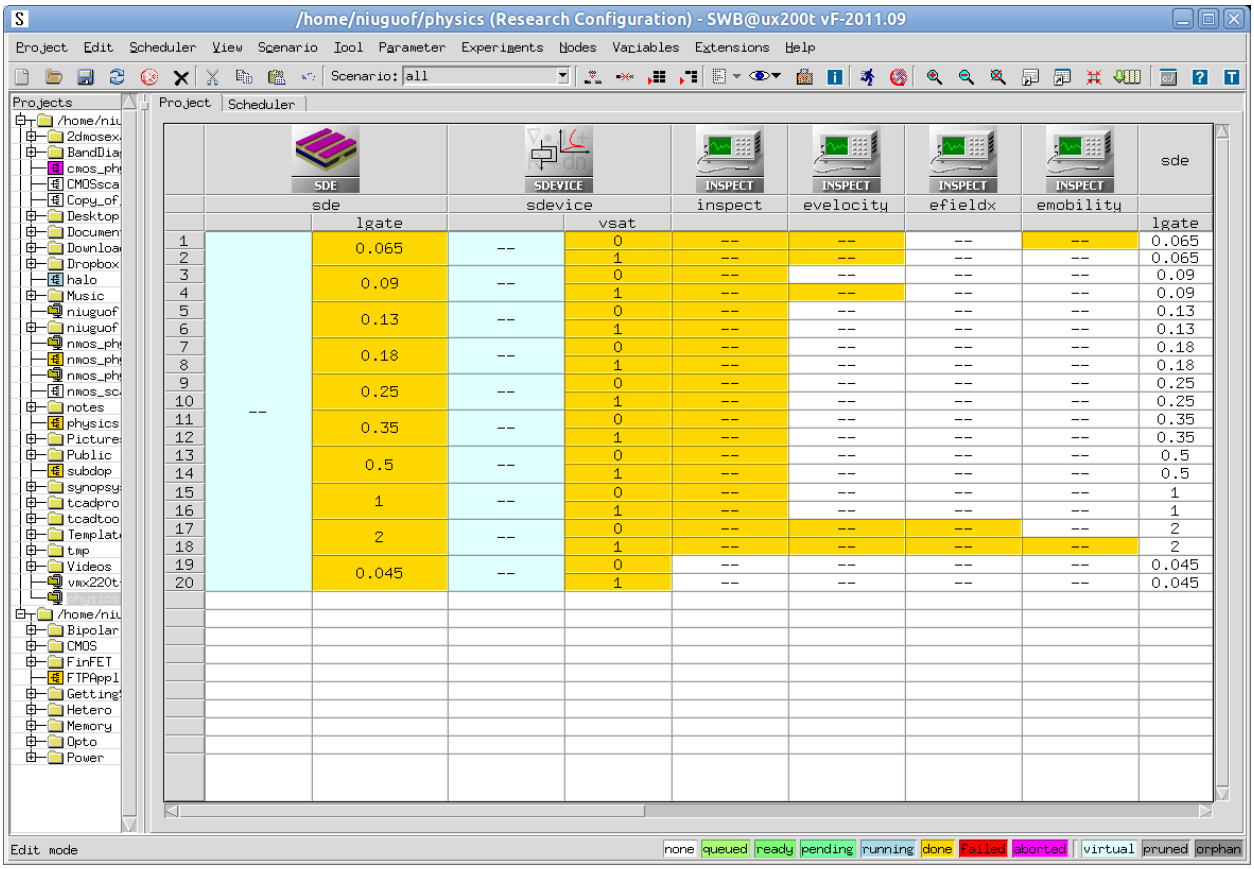


Figure 6.3: Velocity Saturation Physics swb project

You are now in a position to examine the impact of vsat. You can check out the Vds dependence, Lg dependence, etc. You can also on the command line look at the plt and tdr files produced. You can use tecplot_sv to look at those as well.

To save your disk space, I have commented out the commands for saving .tdr files during Vd sweep. There, however, is still at least one tdr file, which is saved in the very end, that is, for the highest Vgs and Vds. Check your folder, sort it by date, you will see which files are new.

6.6 Mesh

Always examine your structure and mesh before running a lot of simulations. A sample mesh for the 65nm device is shown below:

6.7 Id-Vd Curves

The Id-Vd curves for each gate length obtained with and without turning on velocity saturation effect are overlayed, so that we can see the impact of velocity saturation directly.

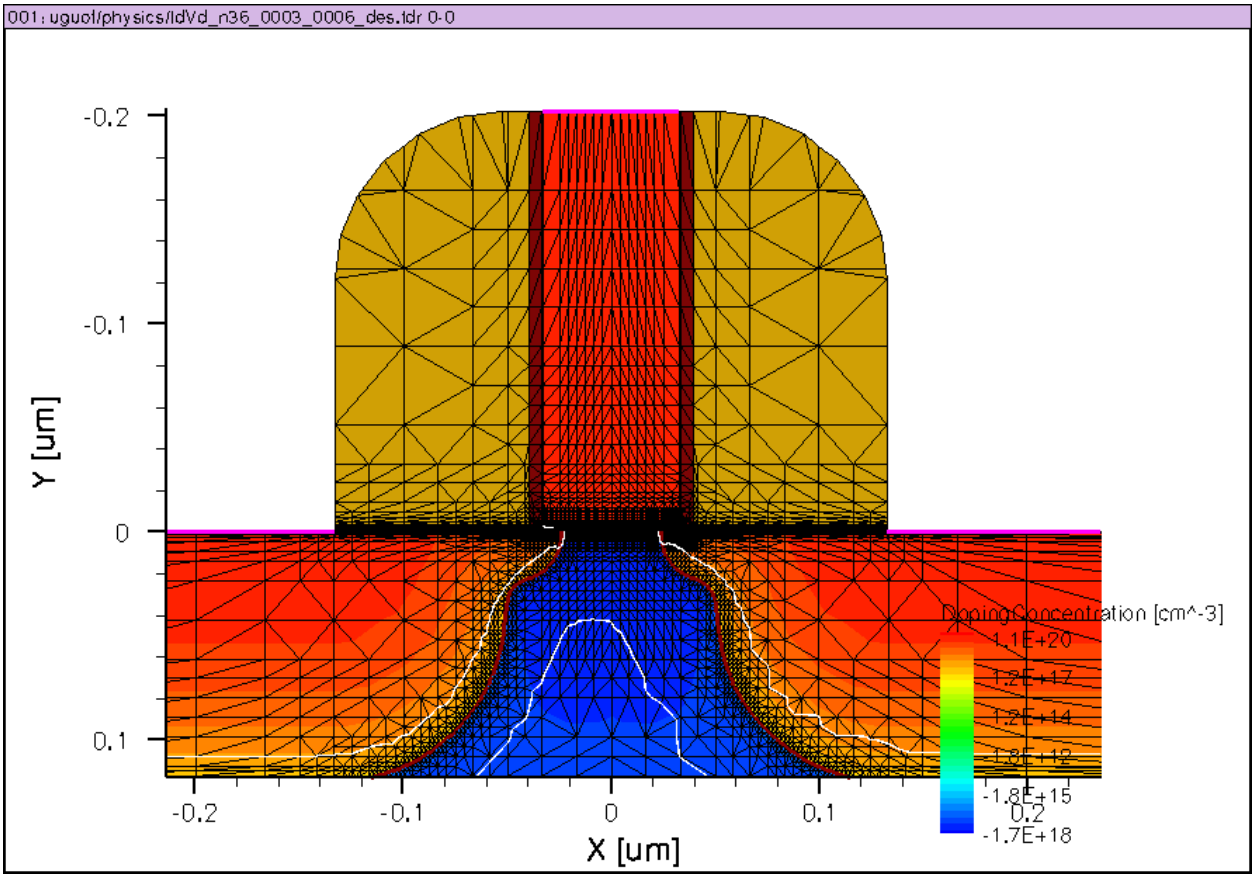


Figure 6.4: 65nm device mesh.

Simulations were run for all the gate length from 45nm to 2um, which produce the following Id-Vd plots:

6.7.1 2um

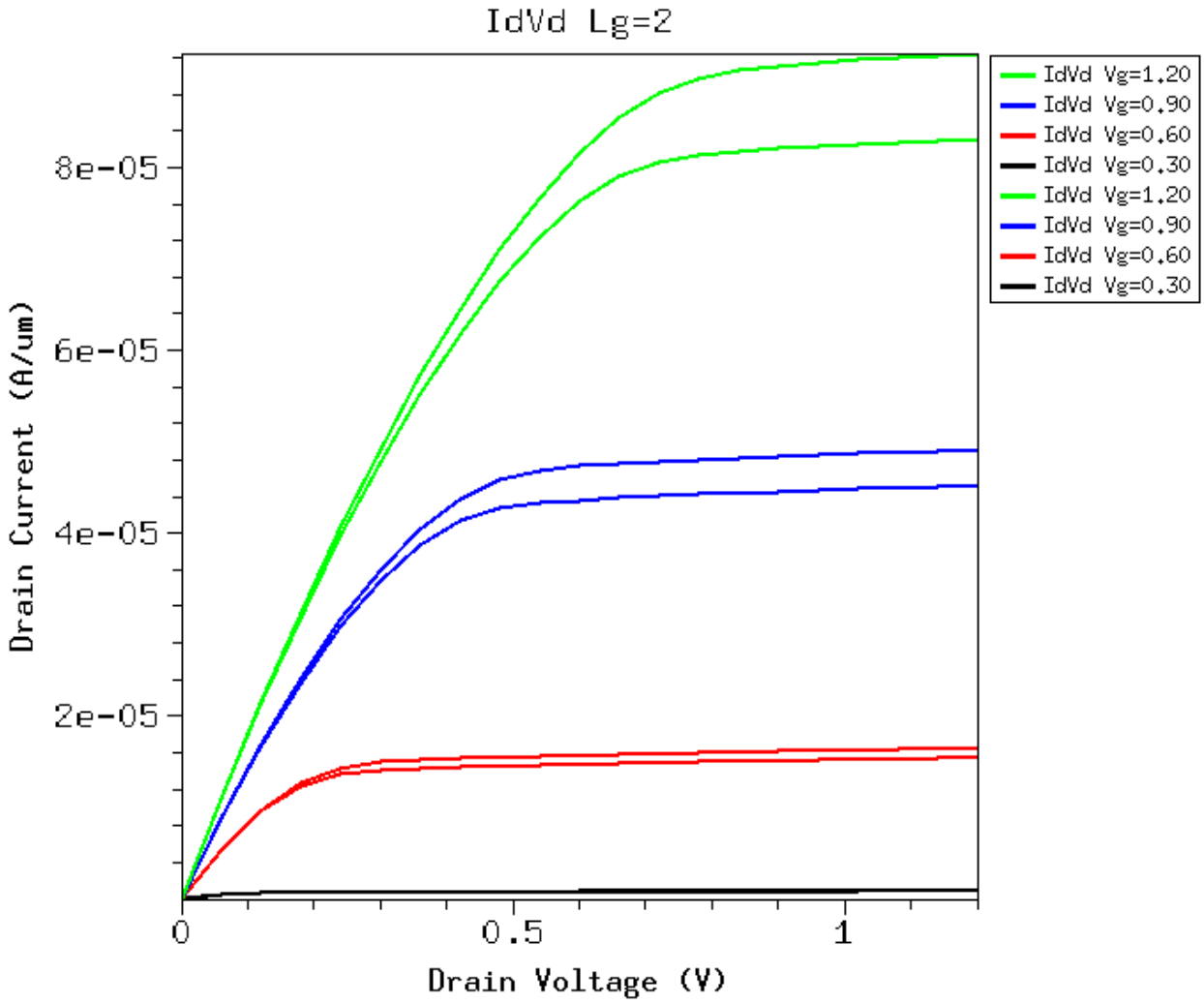


Figure 6.5: Id-Vd, lgate = 2um.

There is visible but small difference for 2um gate length.

6.7.2 1um

Note there is more difference compared to 2um lgate.

6.7.3 0.5um or 500nm

Vsat makes much more difference.

Id is smaller, and Vdsat is smaller too, as expected.

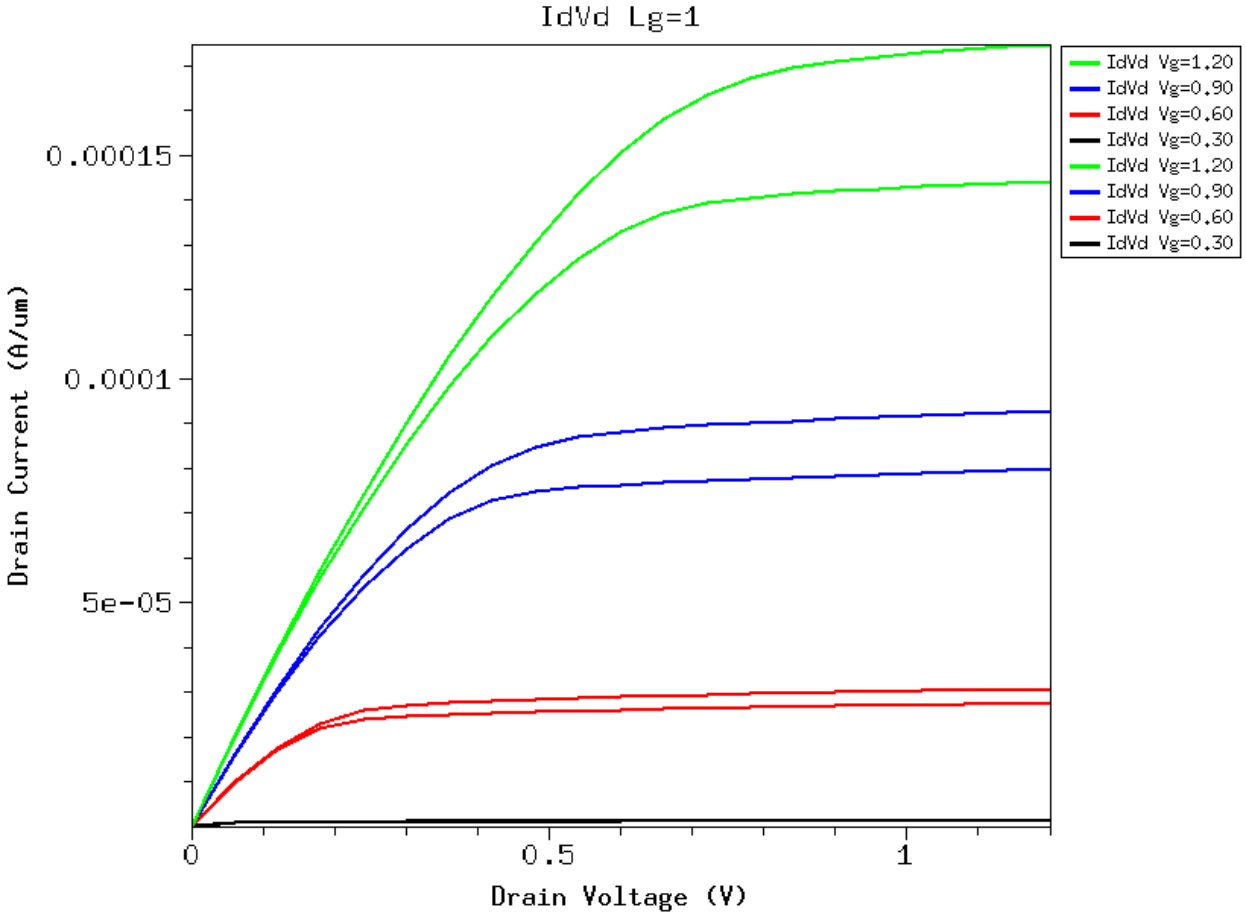


Figure 6.6: Id-Vd, lgate = 1um.

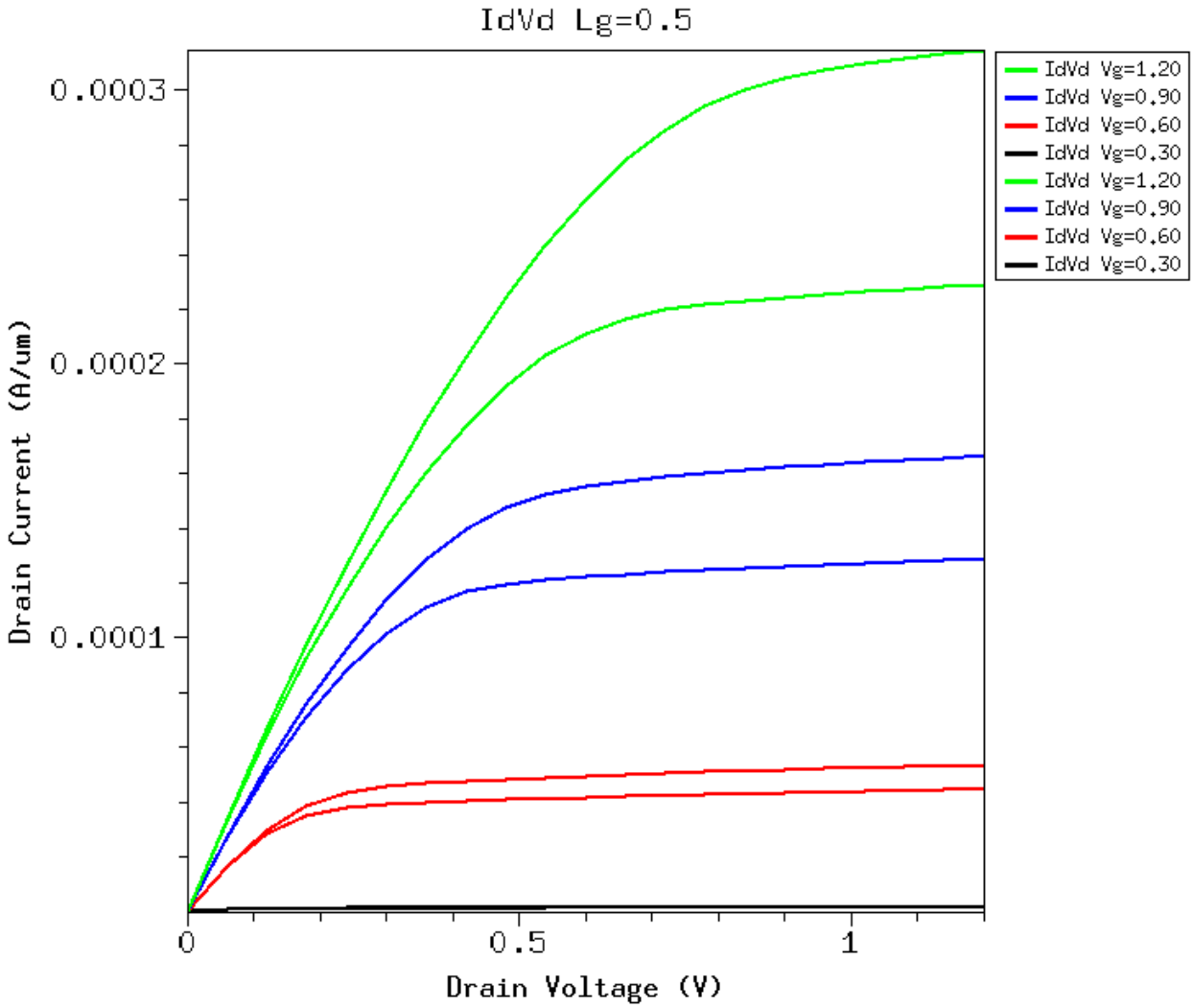


Figure 6.7: Id-Vd, lgate = 0.5um.

Velocity saturation occurs at a smaller Vd than drain end channel pinch-off, therefore, Id starts to *saturate* earlier with velocity saturation.

Now pay attention to the Vds at which saturation starts to occur, or Vdsat, for the same gate voltage as gate length varies further.

6.7.4 0.35um or 350nm

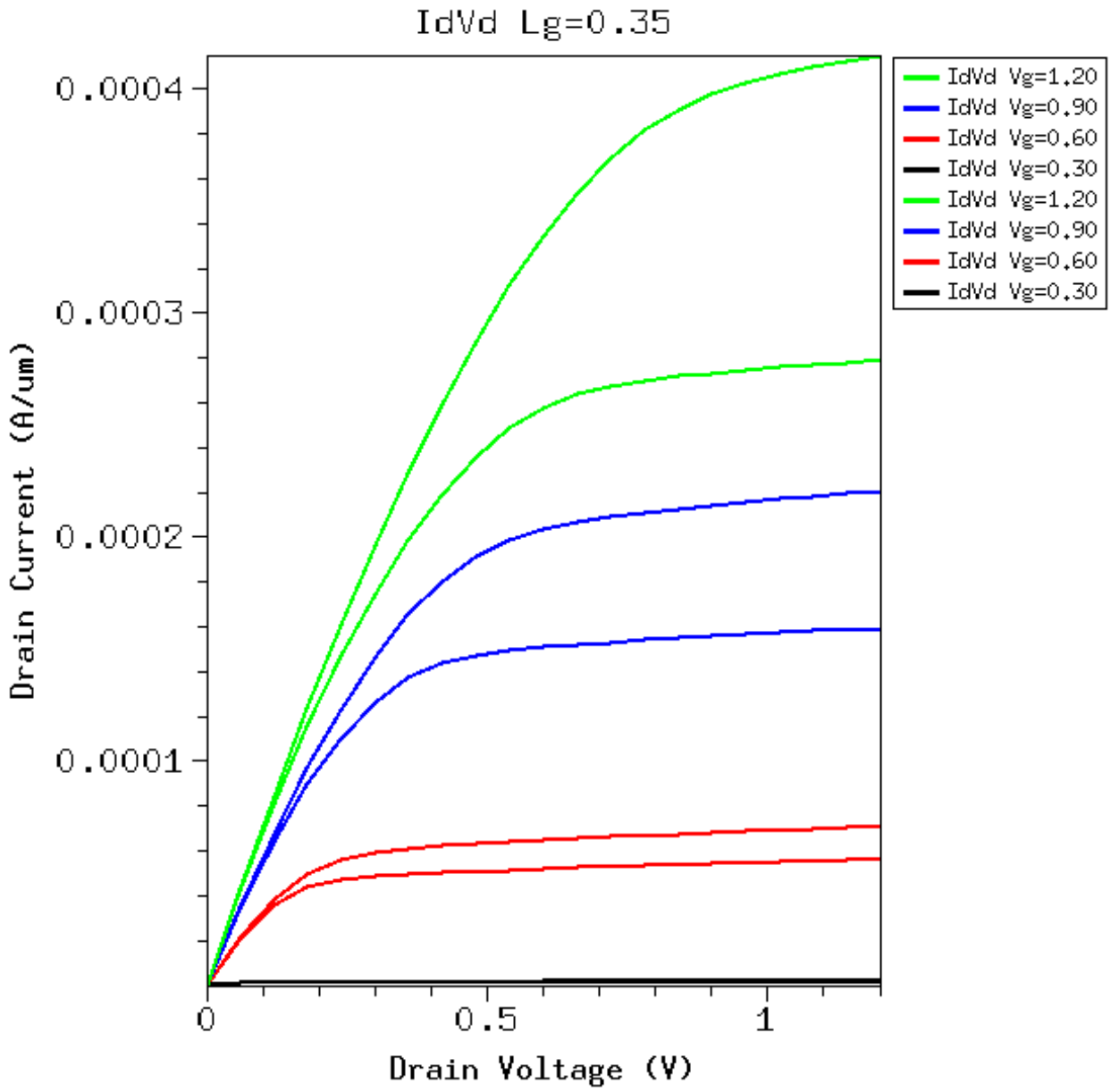


Figure 6.8: Id-Vd, lgate = 0.35um.

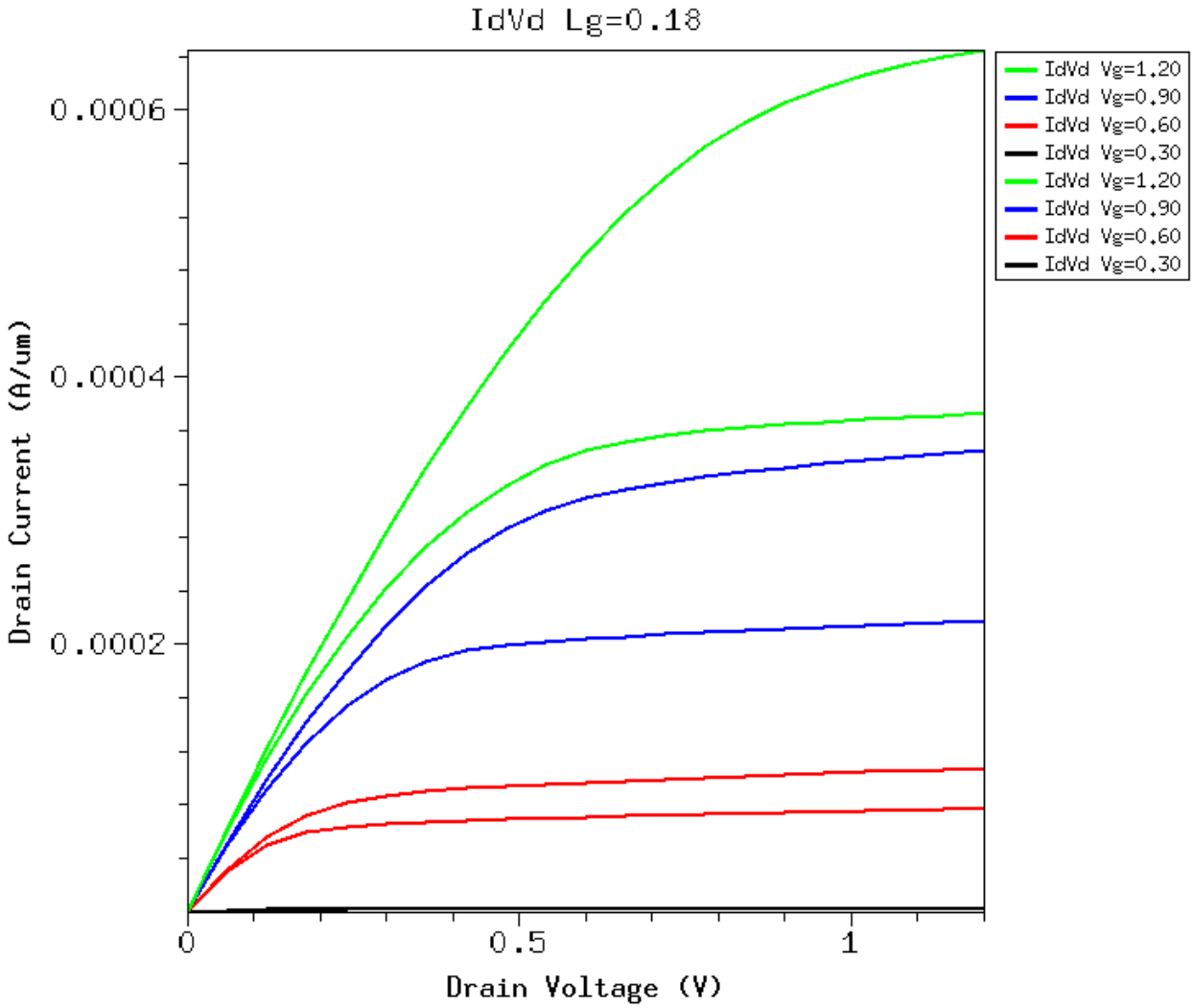


Figure 6.9: Id-Vd, lgate = 0.18um.

6.7.5 0.18um or 180nm

6.7.6 0.13um or 130nm

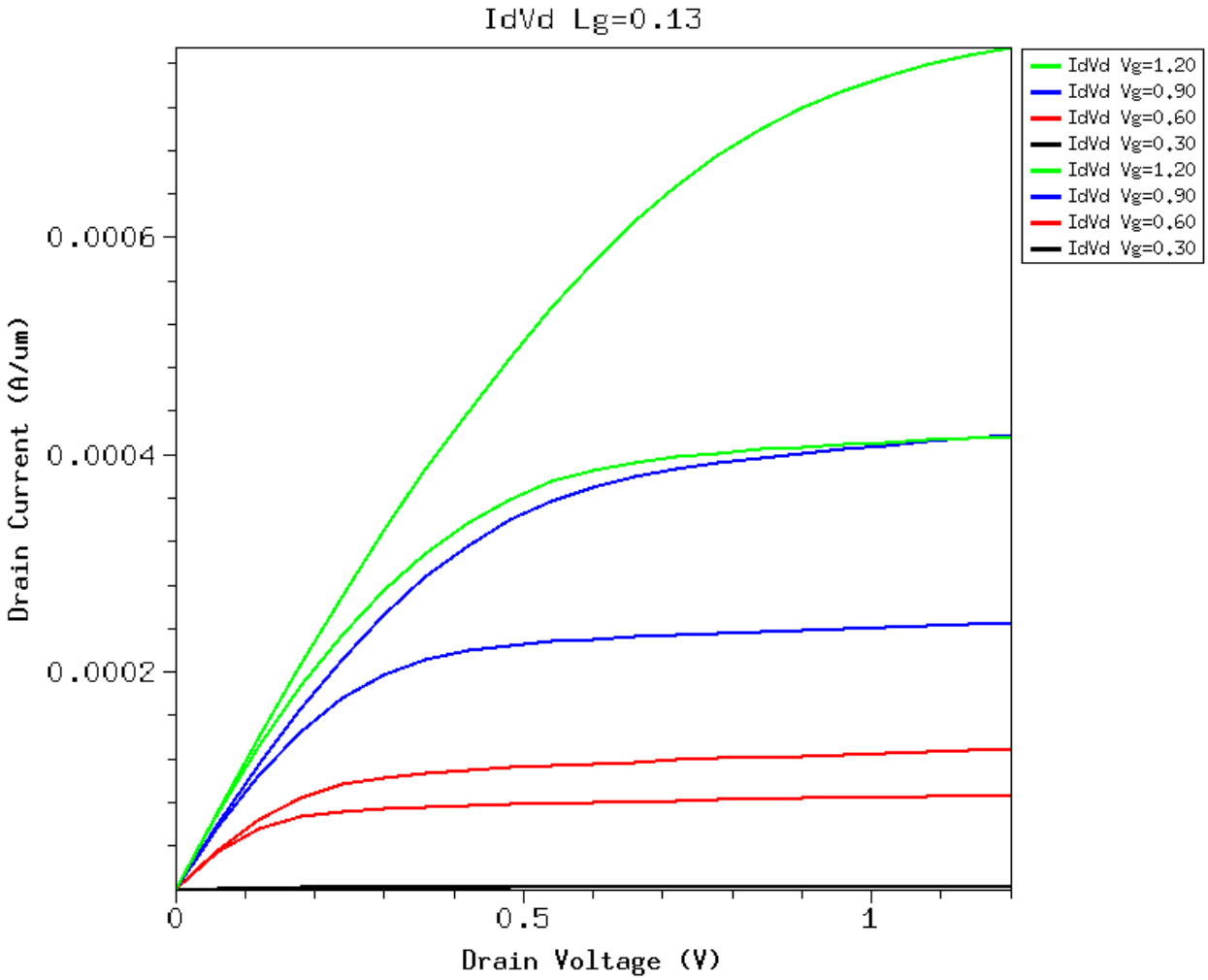


Figure 6.10: Id-Vd, lgate = 0.13um.

6.7.7 0.065um or 65nm

Now you can summarize how the impact of velocity saturation changes with gate length. Think about these questions:

1. Does the impact on Idsat and Vdsat become stronger with gate length scaling?
2. For the same gate length, why is the effect stronger at higher Vd?
3. For the same high Vd = Vdd, at the same lgate, is the impact on Id stronger at higher Vg or lower Vg? Why? Consider only the higher Vg values (above threshold), when the channel is inverted. This is a pretty challenging question without a simple answer. But I encourage you to think about it, and try to make some observations.

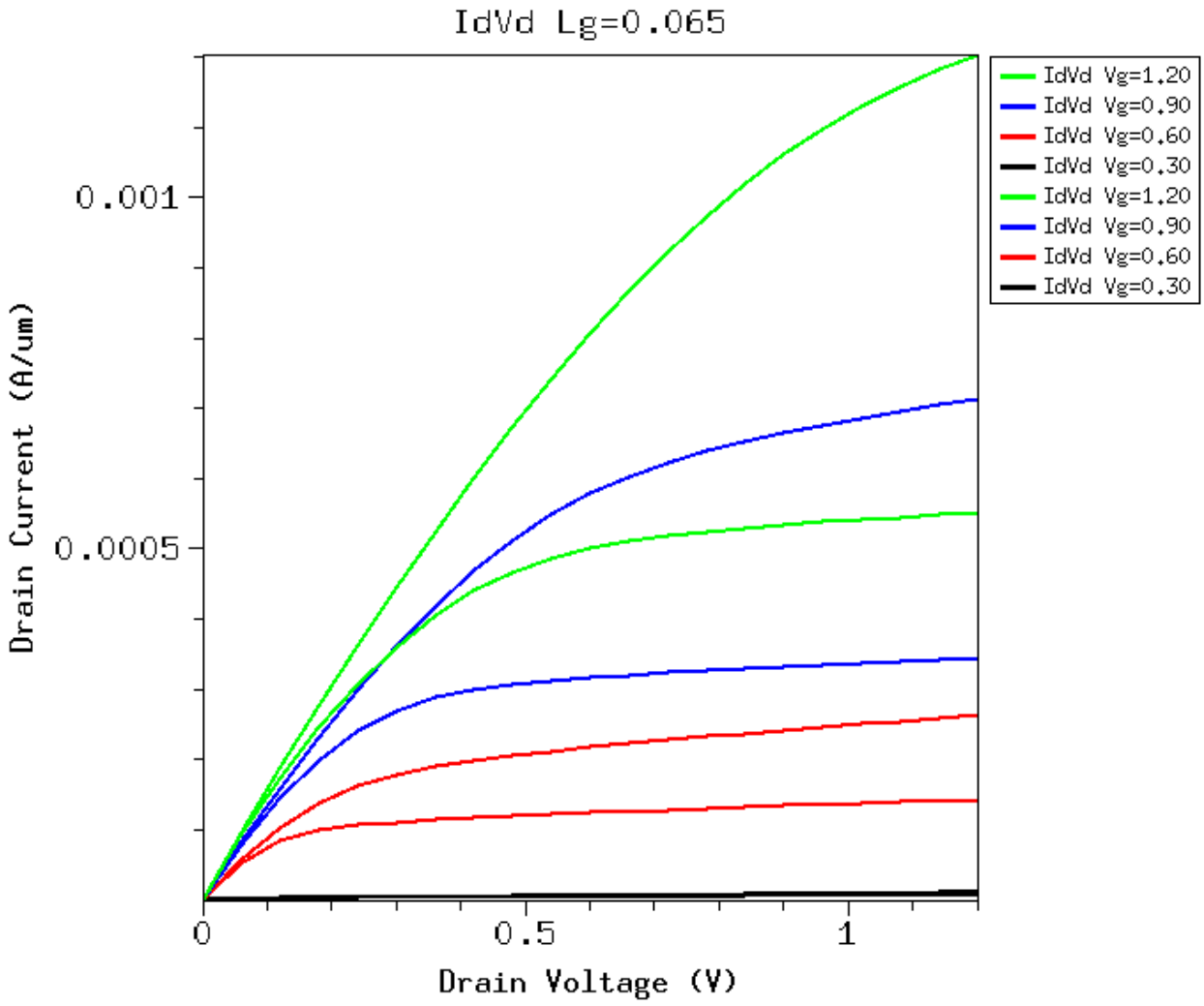


Figure 6.11: Id-Vd, lgate = 65nm.

6.8 Lateral field and Velocity at Mid Channel Point

Velocity saturation is a high lateral field effect. So let us look at how the lateral field in the channel varies with Vd. You should turn on the tdr file saving to see position dependence. Below are some mid channel point plots made using the CurrentPlot technique.

6.8.1 2um

Lateral field and eVelocity

The lateral field at mid channel for 2um lgate is shown below:

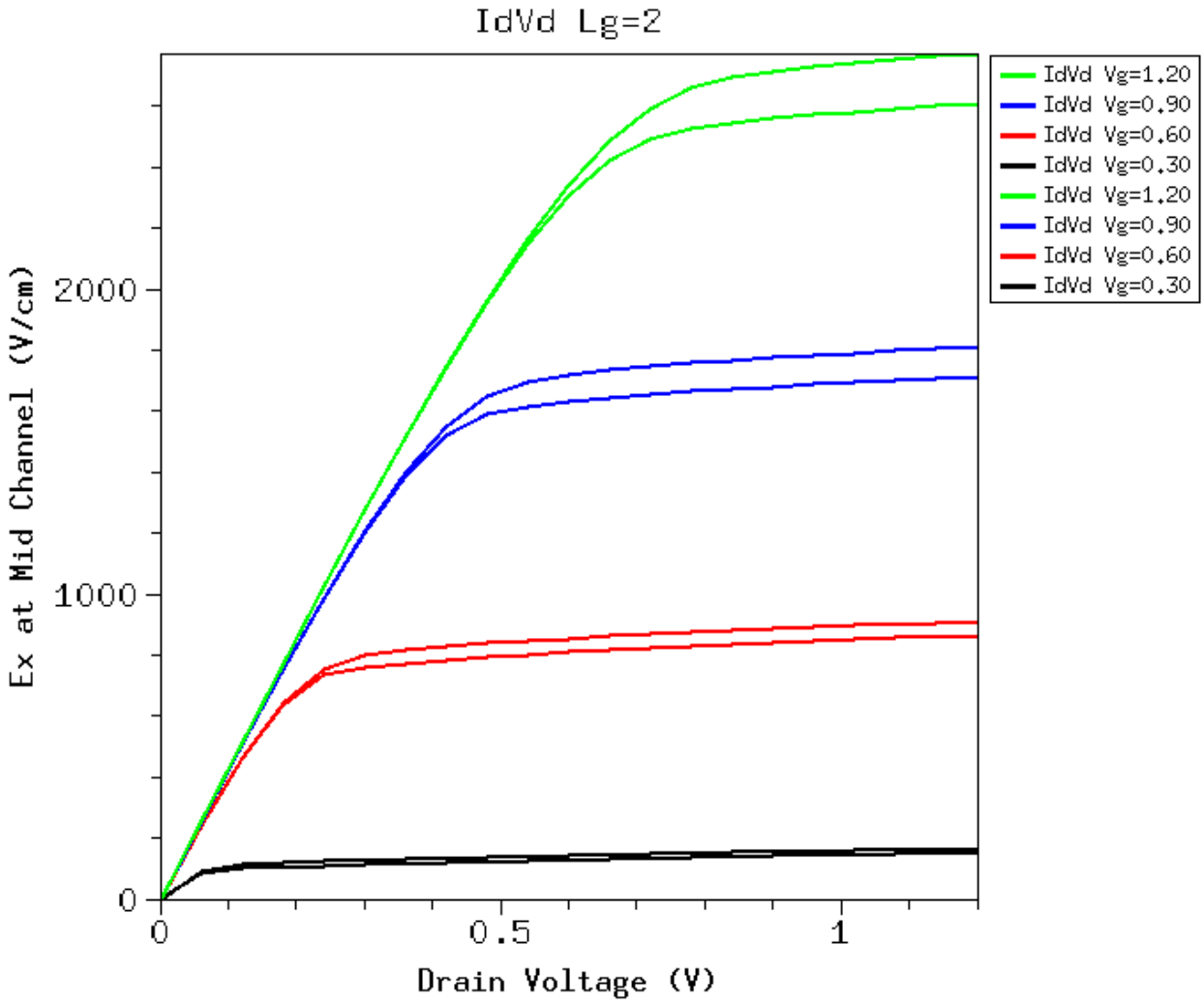


Figure 6.12: Ex at mid channel point at surface versus Vd, lgate = 2um.

The corresponding eVelocity is:

Pretty much at lower Vds, the channel is like a resistor, varying Vg simpy changes resistance, but the lateral field is simply Vds/L, and thus indpendent of Vg.

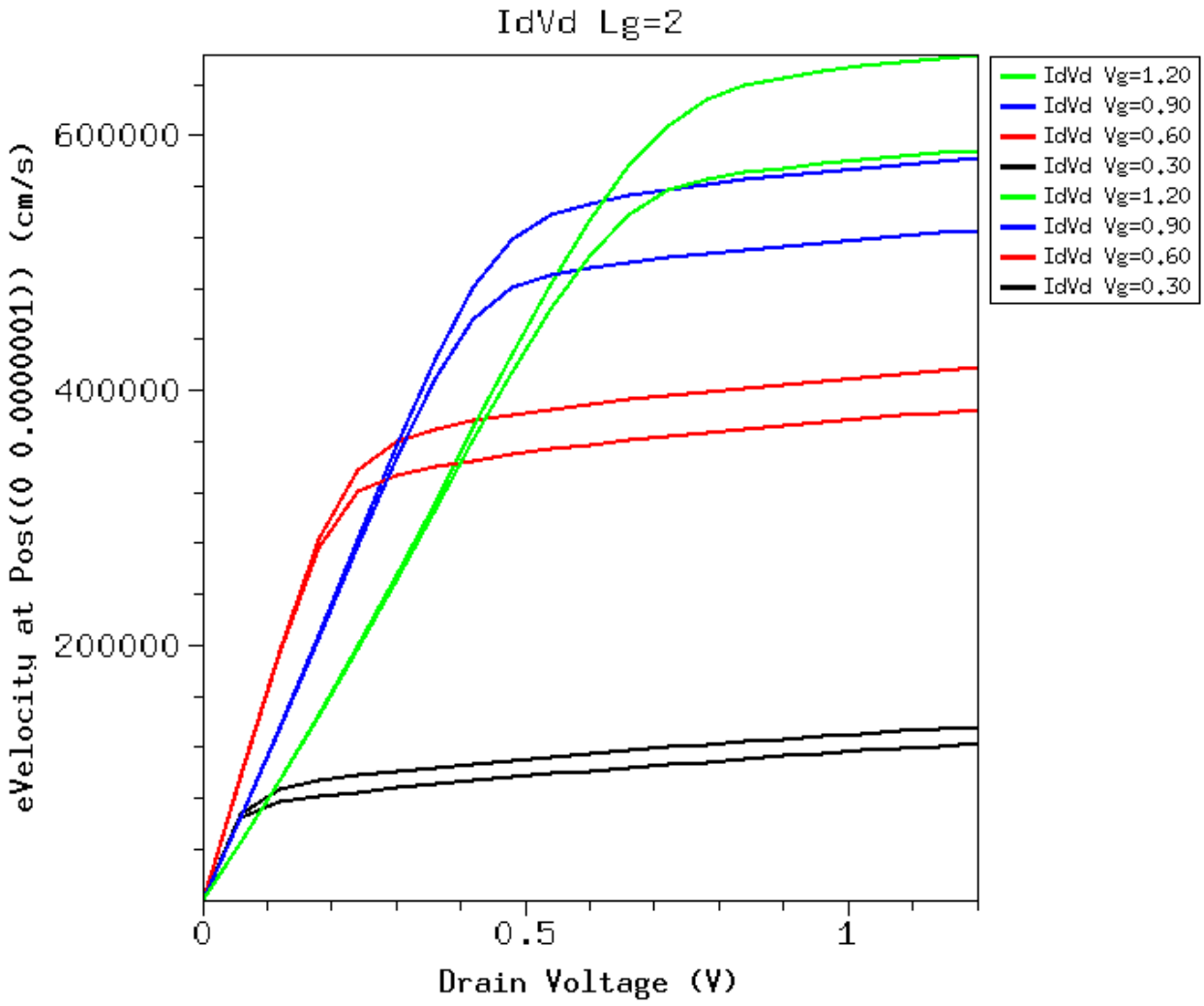


Figure 6.13: eVelocity at mid channel point at surface versus Vd, lgate = 2um.

The difference is at high Vd, when saturation occurs. With velocity saturation, drain current saturation occurs at a smaller Vd, which will then make the Ex inside the channel saturate too. Recall much of the extra Vd beyond saturation will be drop over a very narrow region near the drain.

After saturation occurs, the intrinsic channel's surface potential and hence lateral electric field will no longer change much. This means the mid channel lateral field Ex will also stay pretty much saturated.

Think about these plots and see if they make intuitive sense to you. For 2um lgate, velocity saturation plays little role. It is pretty much showing classic long channel behavior.

Velocity Saturation Region

This high field region near the drain used to be called "pinch-off" region. As velocity saturation is a fact of life for these short channel devices, and velocity saturation occurs before pinch-off, this high field region near the drain is more properly called *velocity saturation region* (VSR).

eVelocity vs Ex

Can you think about why the Ex plot does not have exactly the same shape as the eVelocity plot? Hint: think about mobility and how mobility changes with increasing Vg.

Note: Answer: The decrease of mobility with increasing Vg (recall the surface roughness effect, and the Enormal dependence we investigated earlier) caused this difference.

See also if you can correlate the velocity plot to the Id-Vd plot shown above for 2um. Does it make sense to you?

The eVelocity in the linear operation region is clearly higher for lower Vgs, but Id in the linear region is still higher for higher Vgs, can you think of a possible reason?

Note: Inversion charge density increases with increasing Vg. This increase dominates over the decrease of eVelocity, so the net result is still increase of drain current with increasing gate voltage.

6.8.2 65nm

At 2um, velocity saturation is not significant for most part of the channel, if you use tecplot_sv to look at the velocity vs x plot.

Let us look at the mid channel point.

eVelocity at Mid Channel

We now clearly see the strong impact of velocity saturation on the electron velocity at mid channel point, that is, *velocity at higher Vds is much reduced* with velocity saturation compared to without velocity saturation.

If we repeat the plots for other gate lengths, as we did for Id-Vd above, we find that mid channel velocity is affected more for shorter gate lengths. At 65nm, there is much more difference.

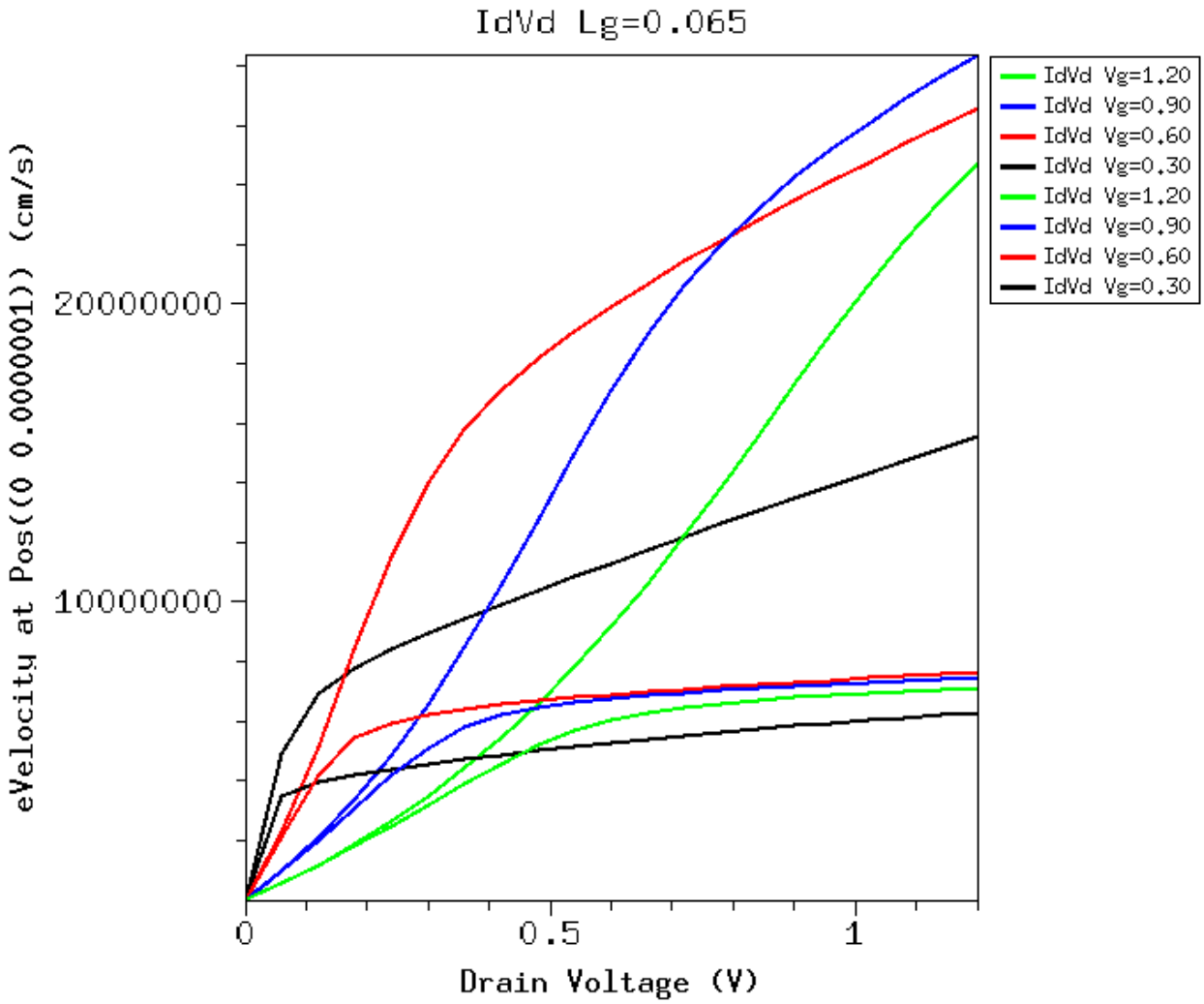


Figure 6.14: eVelocity at mid channel point at surface versus Vd, lgate = 65nm.

Ex at Mid Channel

Note that velocity saturation effect also has caused change of the e-field as well, so you cannot simply assume E will be the same when you turn on velocity saturation effect. It is a self consistent solution of all equations in the end.

The Ex at mid channel at surface versus Vd for 65nm is shown below:

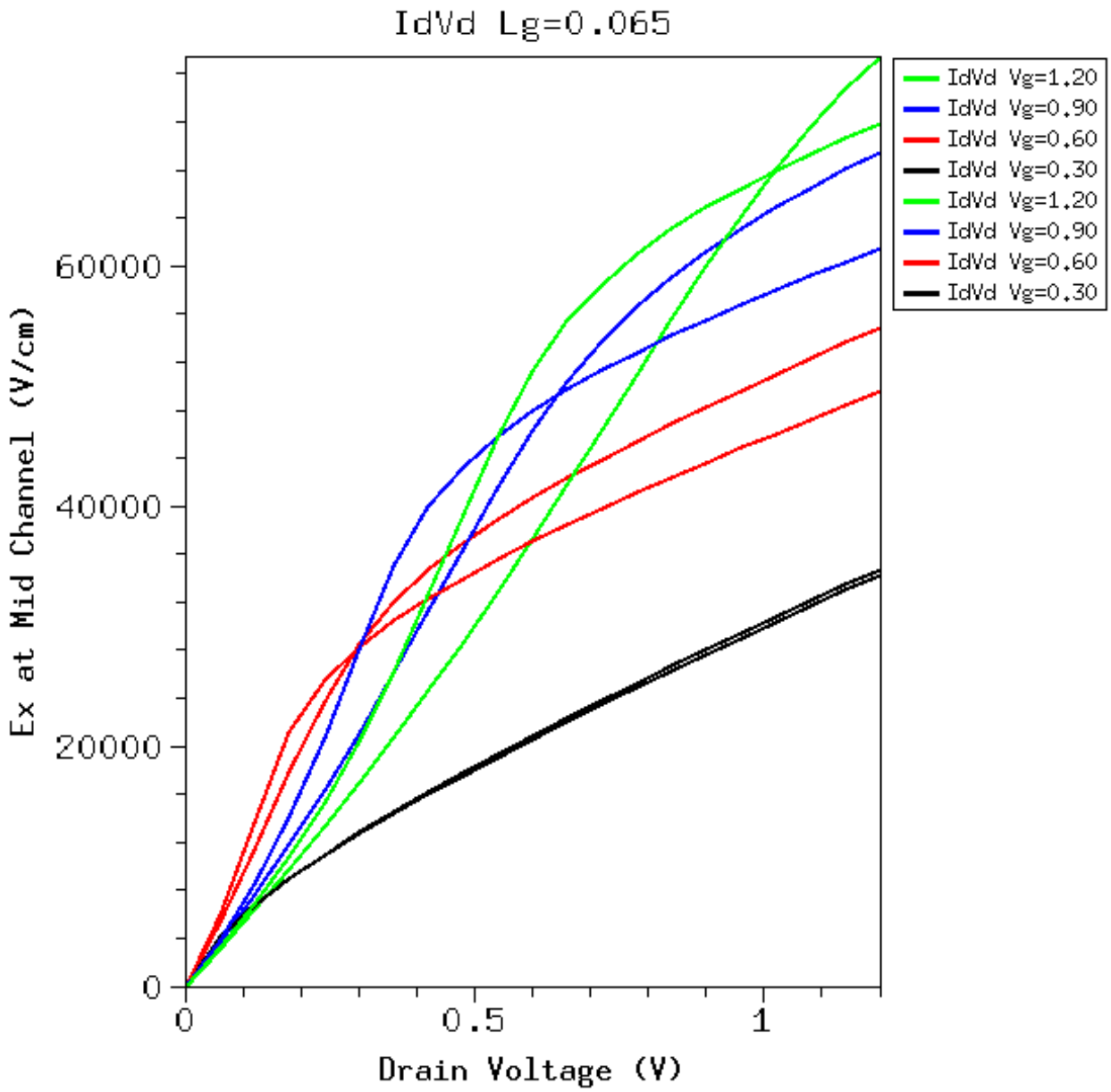


Figure 6.15: Ex at mid channel point at surface versus Vd, lgate = 65nm.

6.9 Inspecting Internal Details

Much of the above analysis is about the mid channel point. Ideally you always want to probe the 2D distribution of internal physical quanities, such as potential, eDensity, eVelocity etc.

Comparing 2D plots, e.g. eDensity distribution between with and without vsat, however, can be challenging. You can put them side by side, and observe their differences.

Ideally you want to "overlay" them, this is impossible if we make contour plots. We might make 3D surface plots instead of 2D countour plots. Even this can be hard.

For MOSFETs, a particular 1-D cut of the y-axis at the Si surface, can be used instead. Essentially this shows a cut along the channel from source to drain.

We can load in the .tdr files corresponding to the same Vg and Vd, simulated with and without vsat effect, make y-cut at the surface, then compare about everything we are interested in.

We have in fact made reference to insights gained from such 1D cuts. A number of plots are given below to assist understanding.

6.9.1 2um (long channel)

Surface potential

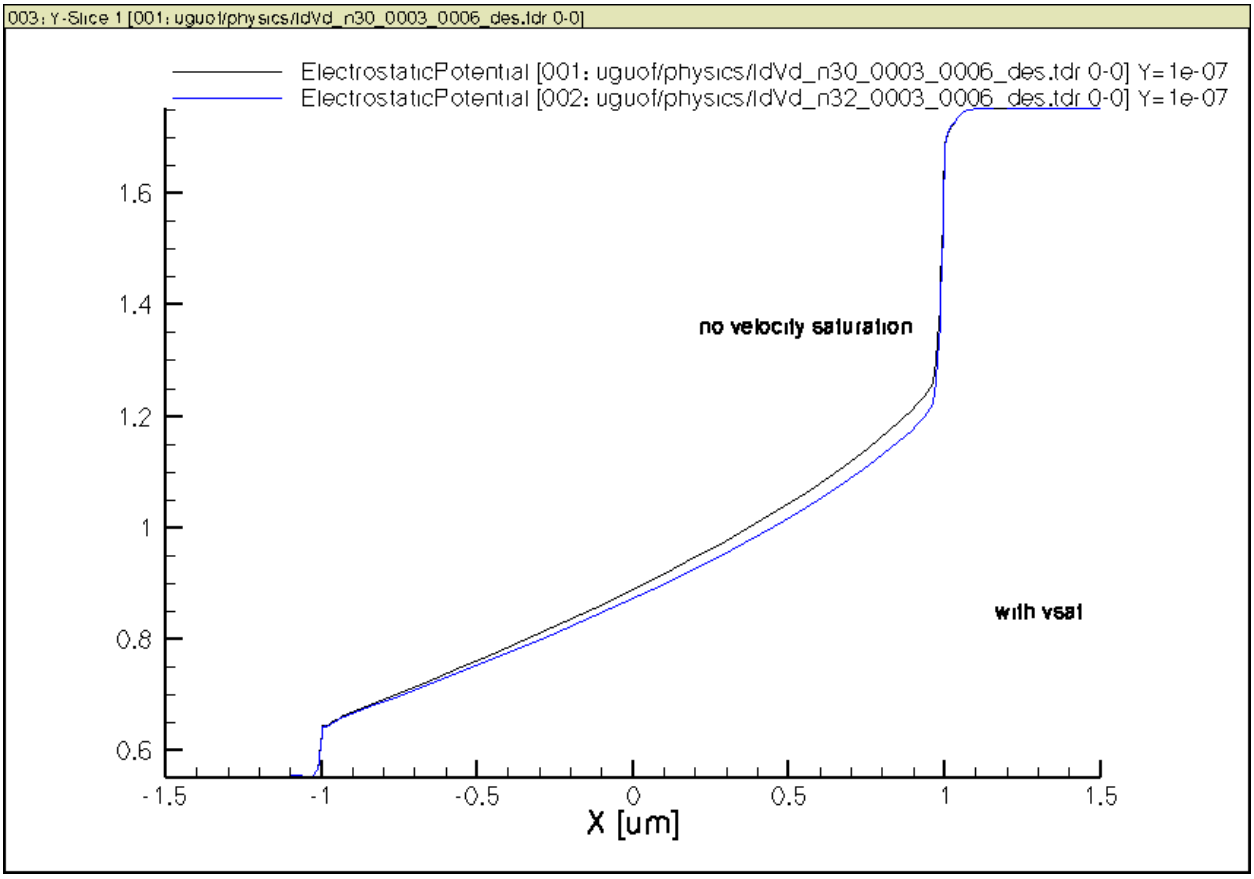


Figure 6.16: Surface potential along channel, lgate = 2um.

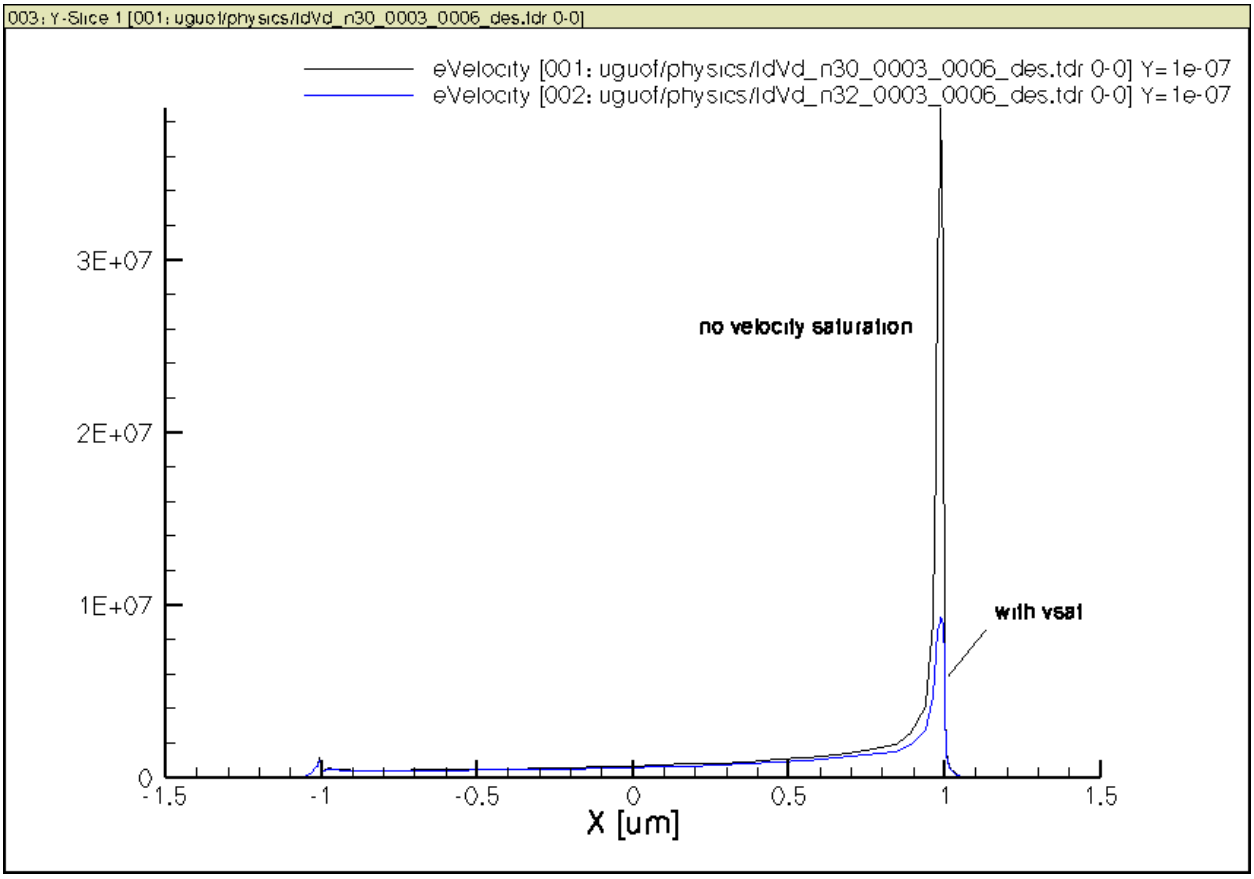


Figure 6.17: Surface evelocity along channel, lgate = 2um.

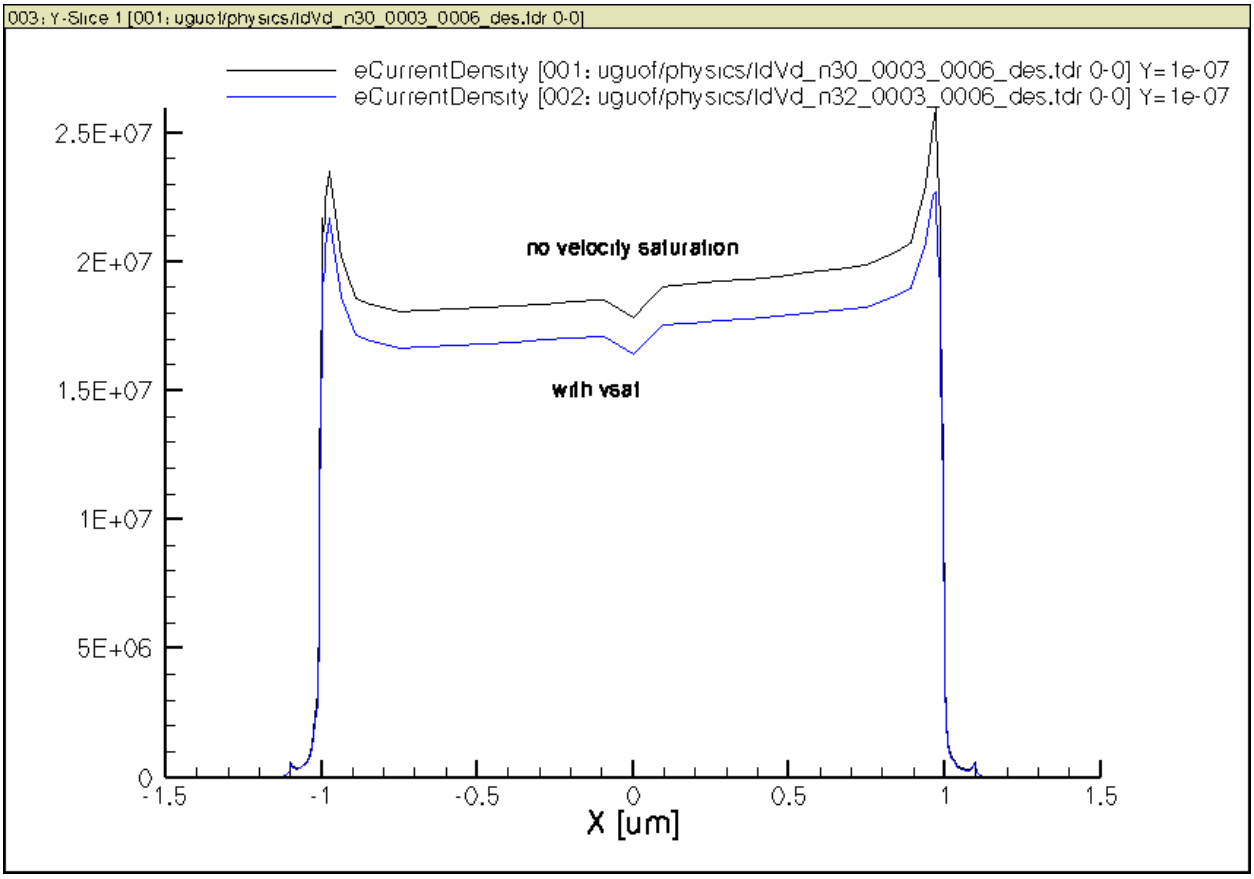


Figure 6.18: Jn along channel, lgate = 2um.

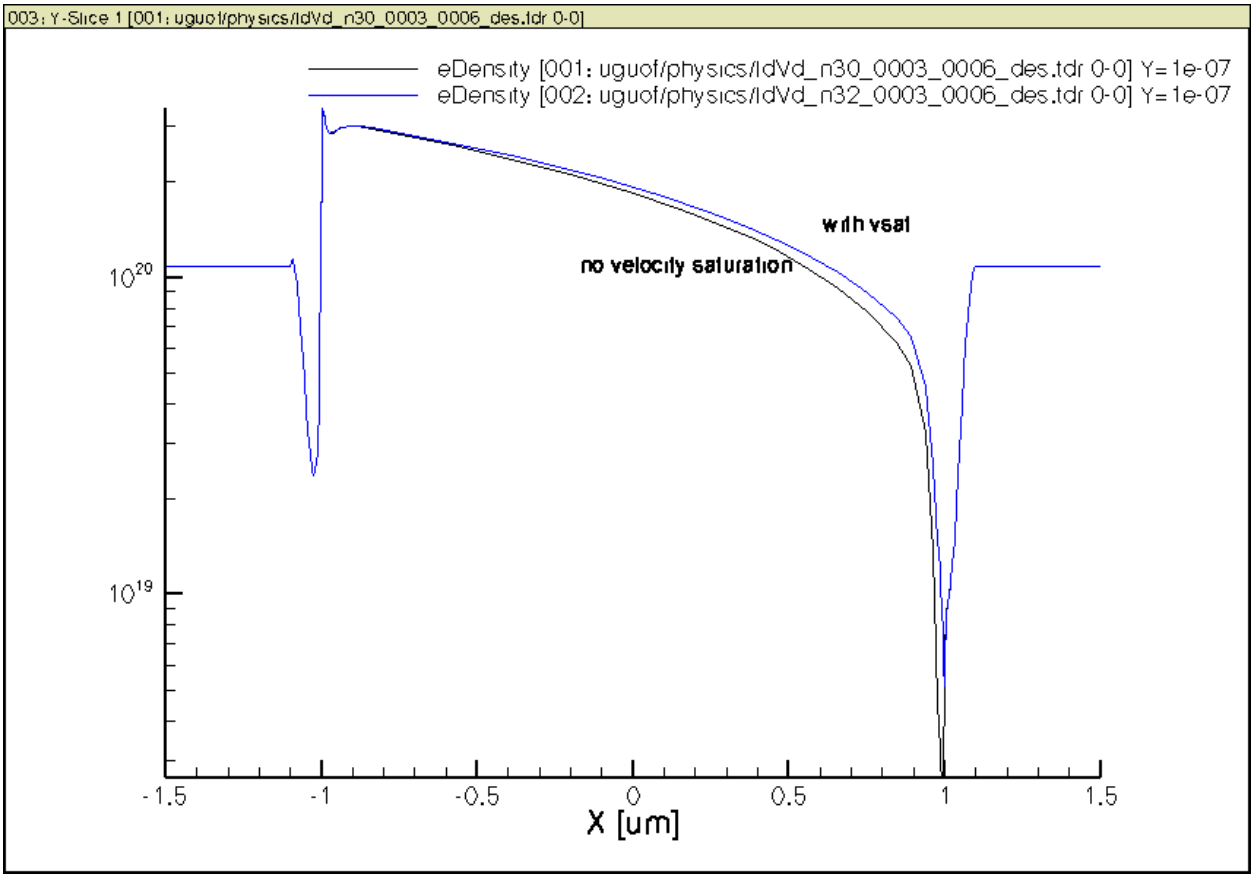


Figure 6.19: Surface n along channel, lgate = 2um.

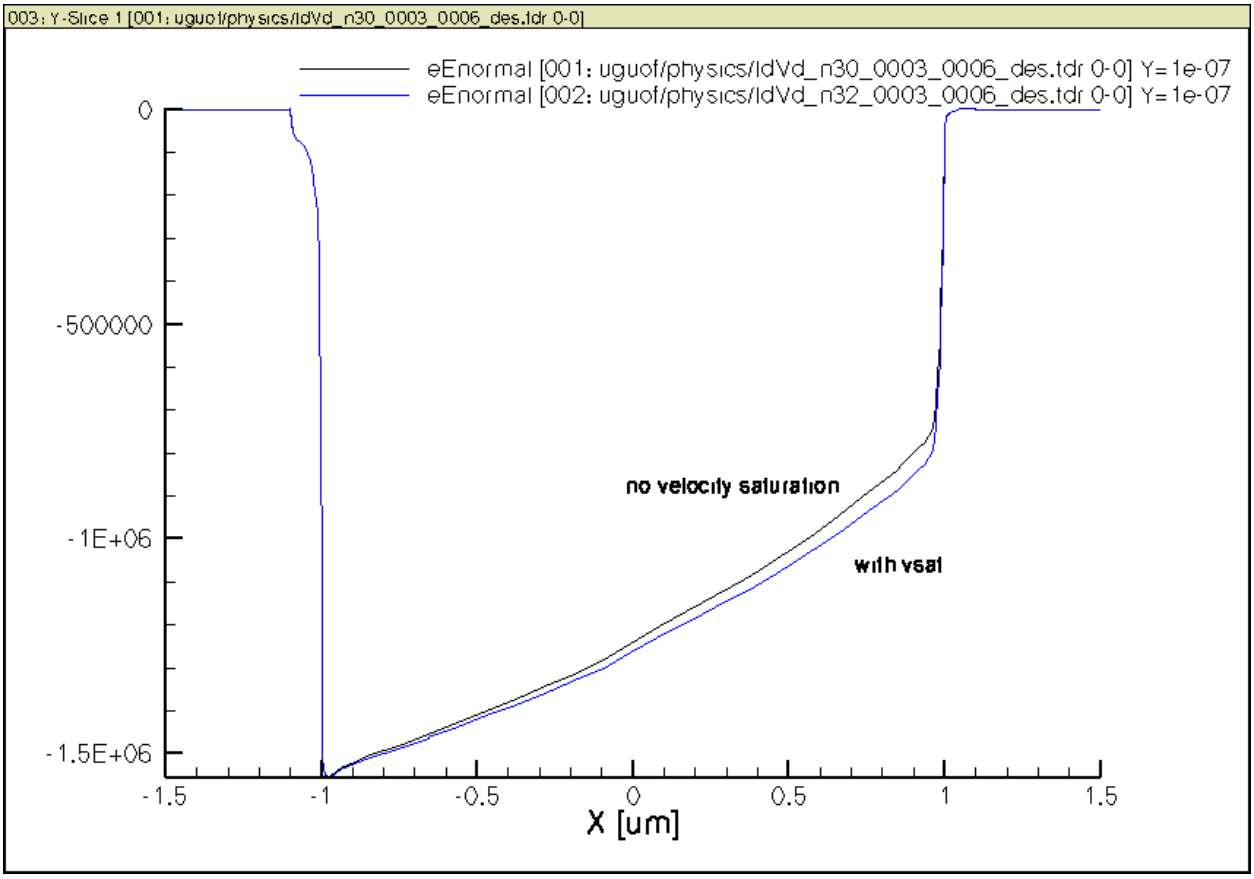


Figure 6.20: E-field normal to current flow along channel, lgate = 2um.

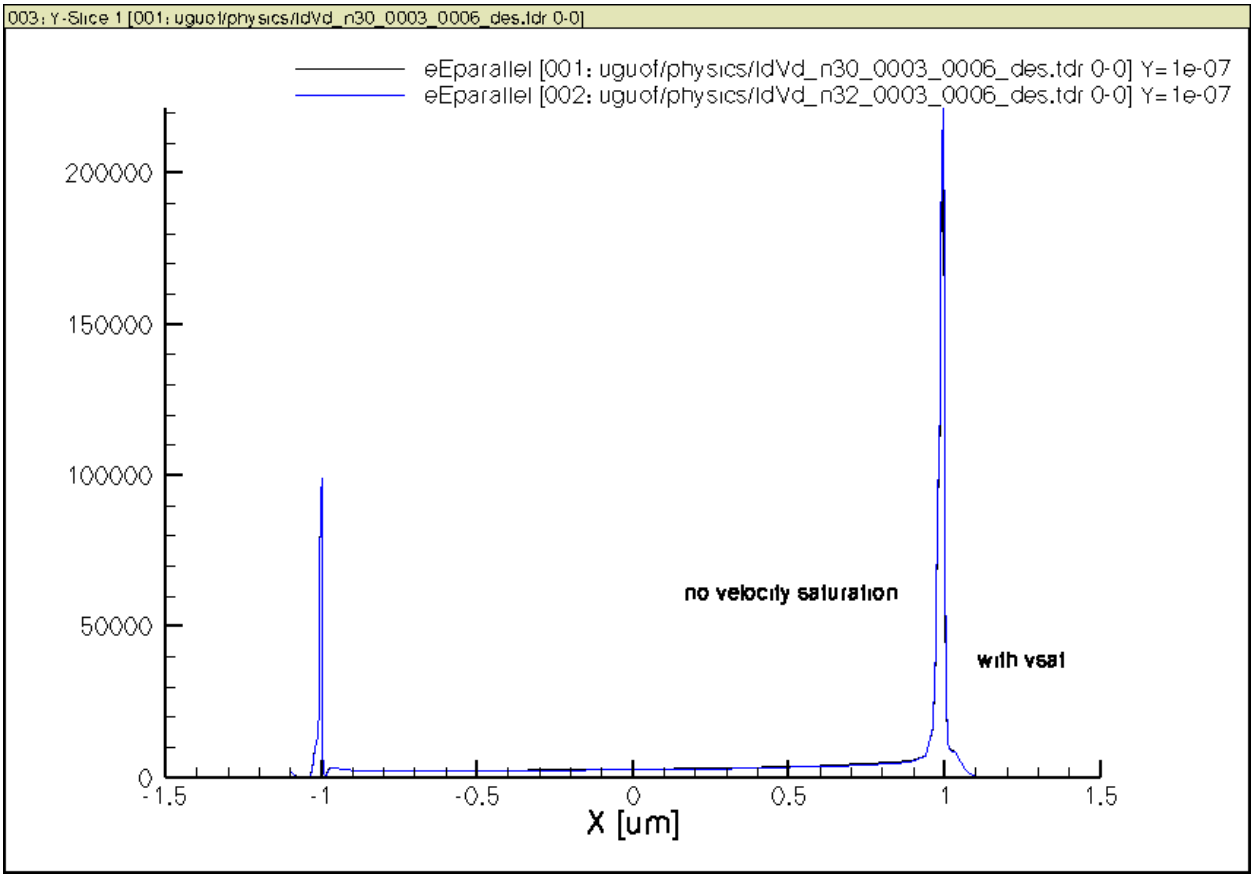


Figure 6.21: E-field parallel to current flow along channel, lgate = 2um.

Surface eVelocity

Other Plots

6.9.2 65nm (short channel)

Surface potential

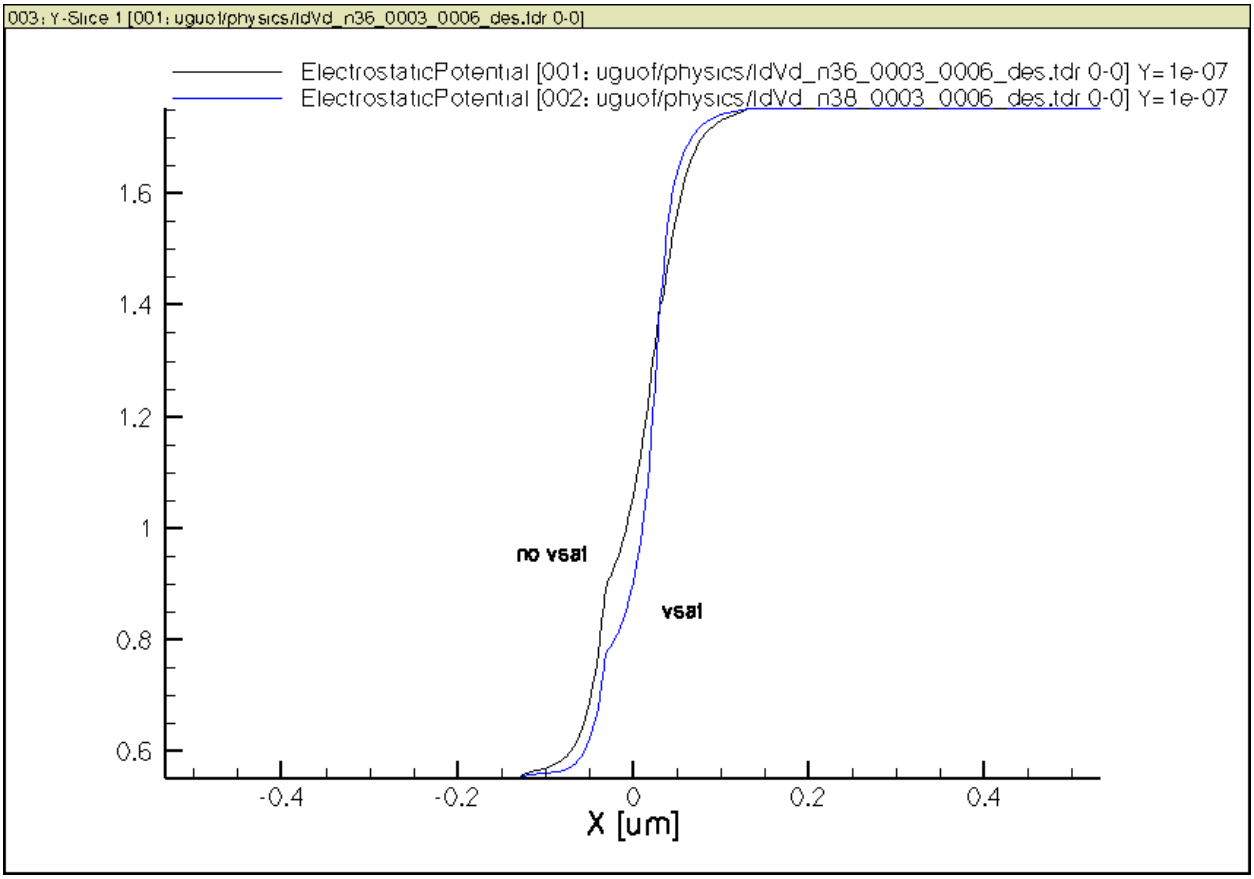


Figure 6.22: Surface potential along channel, lgate = 65nm.

Surface eVelocity

Surface eCurrentDensity

6.10 Homework

Play with the swb project, save graphics, and write up any analysis you come up on the impact of velocity saturation, e.g. dependence on Lg, Vd, Vg. You can look at internal velocity too, even the .tdr files.

This will get you ready for our mid term project.

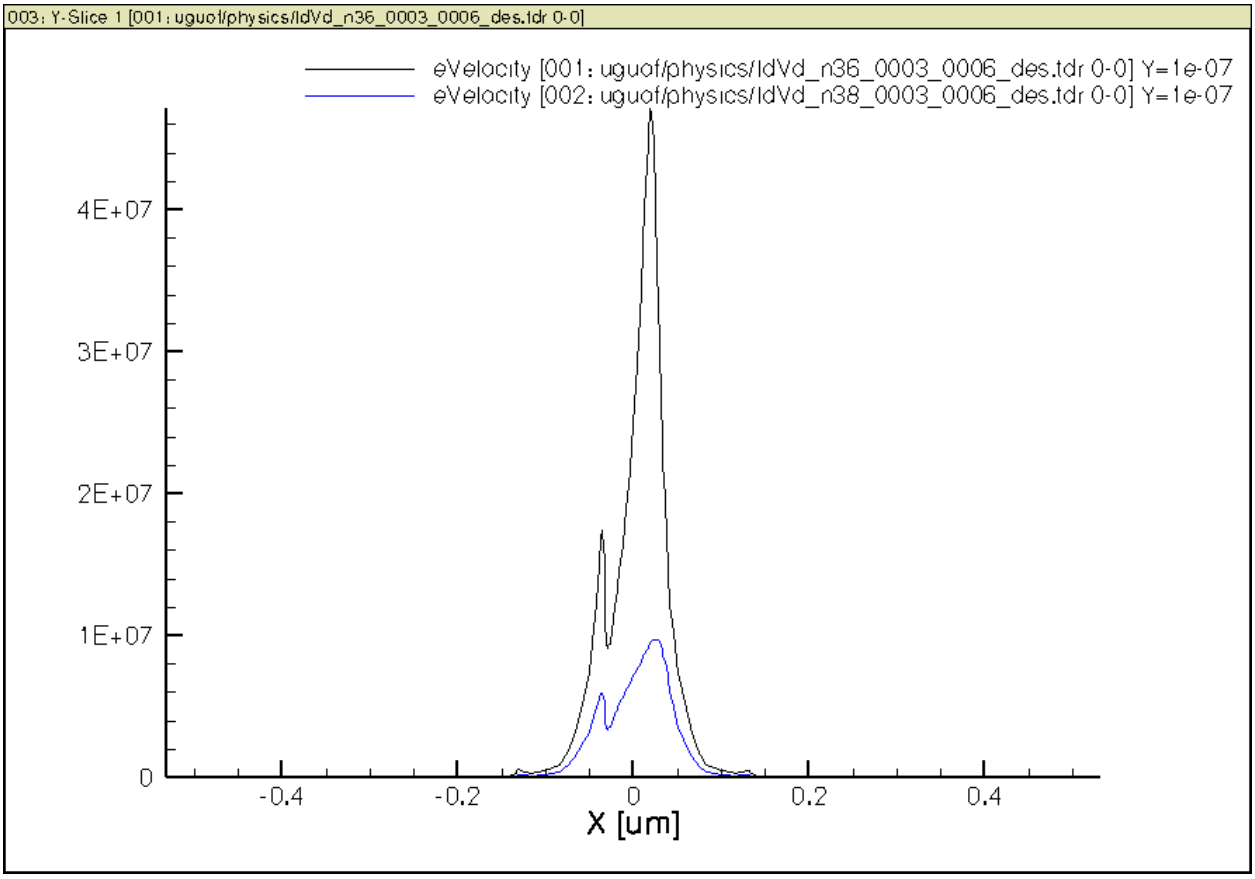


Figure 6.23: Surface evelocity along channel, lgate = 65nm.

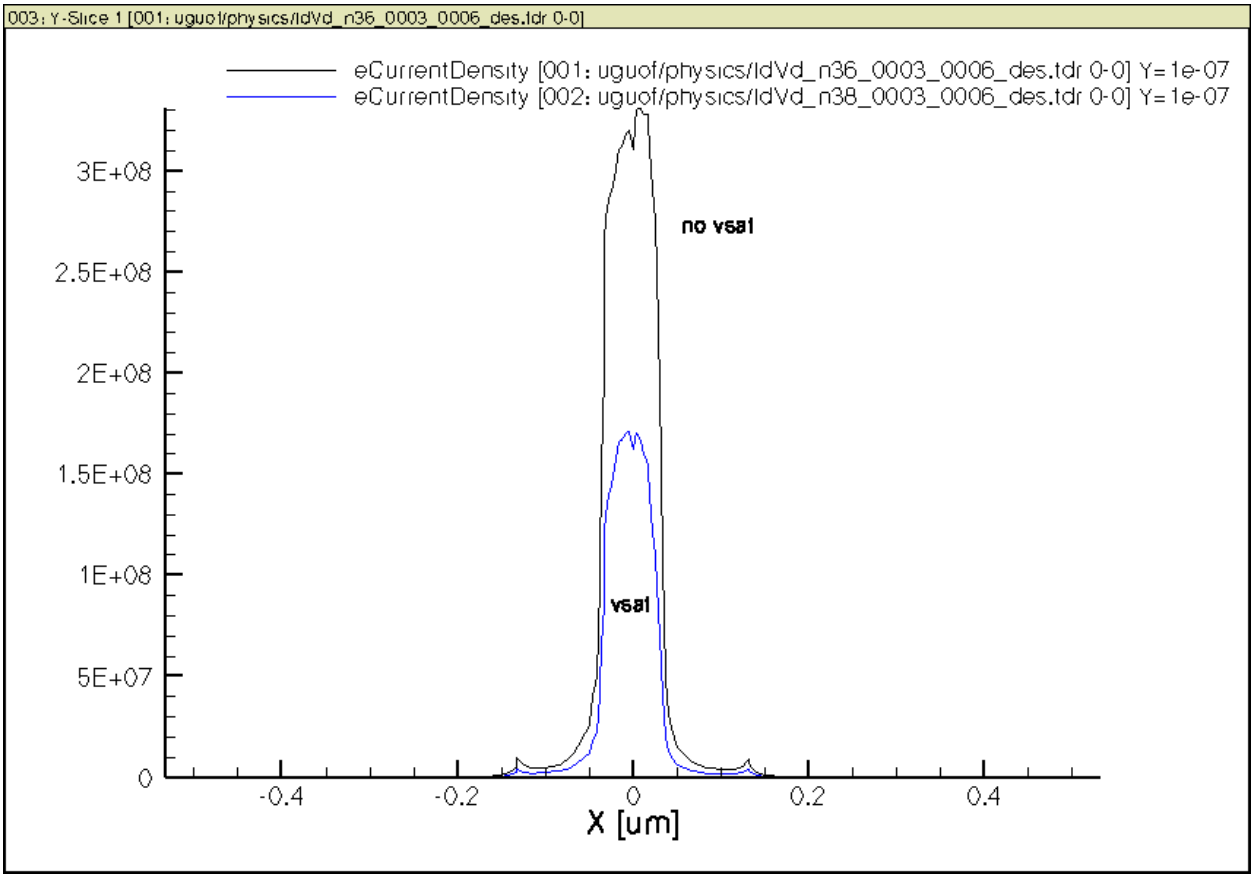


Figure 6.24: Surface Jn along channel, lgate = 65nm.

CHAPTER

SEVEN

MODERN CMOS TRANSISTOR PHYSICS USING PERFORMANCE PARAMETRIC ANALYSIS

7.1 Objectives

The objectives for this project are:

- understand CMOS transistor electrical characteristics, Id-Vg, Id-Vd
- understand body effect
- extract threshold voltage, subthreshold swing, Ion, Ioff
- quantify impact of mobility choices on Id-Vg and Id-Vd, velocity saturation in particular
- quantify impact of gate length
- quantify impact of poly gate depletion
- quantify impact of gate oxide dielectric constant
- quantify NMOS and PMOS differences
- quantify the impact of halo doping on Id-Vg and Id-Vd, Vth, SS, Ion, Ioff
- understand drain-induced-barrier lowering through examining Id-Vg for different Vds, and internal details, e.g. potential contours, surface potential, and electron current stream traces

7.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect
- swb

7.3 Basic Requirements (Reduced)

Except for the halo doping part, all other parts of this project have been done before. When running the inspect programs, in the terminal window you run inspect from, extraction results such as Vth, SS, Ion, Ioff are reported.

You can use swb to manage your project, but you do not have to. It is a powerful tool, but also requires extra effort to be able to write new codes. If it is a task not too different from the example swb I gave you, you can try doing so.

I suggest that you start with just modifying the non-swb codes you have been using for homework. That will be our basic requirement.

Consider using 65nm or 90nm as your nominal short gate length, 180nm or 250nm as mid gate length, and then 1um or 2um as your long gate length. Make 3 folders, one for each gate length. Use Phumob, Enormal, and turn on high field saturation of velocity.

For each gate length:

- (required) simulate Id-Vg at several Vds, e.g. 50mV, 250mV, 500mV, 1.2V. You can use more Vds values as needed.

  Plot out Id-Vg on both linear and log scales. Make observations and give explanations, analyze relevant .tdr files using tecplot_sv as we have done in the past to support your explanations. Surface plots, i.e. y cut made at say y=1e-7, are useful. For instance, you can compare surface potential plots (at all Vg's simulated, overlayed) at Vd=Vdd for all 3 gate lengths, to explain why SS at Vd=Vdd is larger in short gate length transistor.

  Use the `currentplot` commands to plot out the mid channel point's potential, lateral and vertical field, eVelocity, eDensity etc.

  The subthreshold swing is a good parameter to take notice when examining Vds dependence of Id-Vg.

  Look at how your Vth (use Vd=50mV), SS at Vd=50mV and Vd=Vdd, Ion and Ioff (note for Ion and Ioff, Vg=Vdd and 0, Vd=Vdd) change with gate length.

- (changed to optional) simulate Id-Vd. Include Vg = 0V.

  Make observations on the slope of Id-Vd, particularly in saturation region.

- (changed to optional) repeat all simulations, analyze a selected number of .tdr files with velocity saturation turned off. Compare results and explain the difference.

  Discuss how velocity saturation affects Ion, Ioff, Id-Vd (Idsat, and Vdsat in particular), SS, Vth.

7.4 Explorer Further

With some changes to the codes provided, you can investigate more modern CMOS transistor physics. For instance, you can look into a selected number of topics that may interest you at your choice:

7.4.1 Halo Doping

We have seen halo doping's importance from the simulation results I showed with and without halo doping. You can do this by varying the concentration of halo doping, and run your simulation for 3 gate lengths.

An extreme case will be to set halo doping to a level that is negligible compared to substrate doping level. The sde command file provided has an entry on halo doping level, just modify that, and you can do this yourself.

Examine the tdr files as necessary for deeper insights.

7.4.2 Threshold voltage roll-off

With halo turned off, you can run the Id-Vg at 50mV simulations at many gate lengths from 45nm to 1um, record your Vth and SS, plot them out as a function of gate length. You shall see Vth decrease with gate length, due to the increasing impact of the drain to channel junction on surface potential with decreasing channel length.

Examine the tdr files as necessary for deeper insights.

7.4.3 Reverse Short Channel Effect (RSCE)

With halo turned on, repeat the Id-Vg at 50mV simulations for many gate lengths, e.g. 32nm, 45nm, 65nm, 90nm, 130nm, 180nm, 250nm, 0.5um, 1um, 2um, and 5um. Plot out Vth vs lgate. Vth will first *increase* with decreasing lgate, and ultimately decrease again at very short gate length.

You may even vary the halo doping level and see how this vth - lgate curve changes.

7.4.4 Body Effect

So far we have set the body bias to zero. You can vary this, and see how your Id-Vg is changed. See how Vth is changed in particular.

7.4.5 Poly Depletion Effect

Poly depletion effect is real and degrades performance by decreasing the effective gate to channel capacitance. Vary the poly gate doping level entry in the sde command file. Look up our text book or IEEE explorer for poly gate doping found in real CMOS technologies. Try a few values from practical levels to very high, e.g. 1e20, 5e20, and see how your Ion and Gm are affected.

Ultimately, you can also use an ideal metal gate in your sde / sdevice command files.

7.4.6 PMOS

While much of our in class discussions / simulations have used NMOS, we should have equally good understanding of PMOS transistor, as CMOS requires both NMOS and PMOS. By swapping all the doping types in the sde command file, you will get a PMOS transistor. You can then repeat simulations.

You want to compare the hVelocity in PMOS with eVelocity in NMOS to understand the difference. Hole mobility is lower than electron mobility, which results in less drive current for the same set of voltages. That difference will also change with channel length due to velocity saturation.

Keep in mind all voltages in your codes need to be multipled by -1.

You can plot out -1 * Id too as the Id for PMOS will be negative. Current going into a terminal is positive in the simulator.

7.4.7 High K gate dielectric

You can also vary the gate `oxide` material, or simply use silicon diode, but modify the dielectric constant value in the parameter file. A higher epsilon means stronger vertical field effect, which will give you better SS, better Ioff, and better Ion.

7.5 Report

Use as many pages as you need, but think about the best figures to use to support your analysis as well. Every figure you put in the report should serve a purpose. Always make observerations, add analysis, use tdr plots as necessary. Save all your images as png files.

You can use word, latex or any of your favorite word processor.

We will devote this week to this project.

You should have a zipped folder for all your simulations files, ready to run.

Submit your report to me next Monday class time.

CHAPTER

EIGHT

GATE LENGTH SCALING ON TRANSISTOR ELECTRICAL PERFORMANCE

8.1 Objectives

The objectives for this project are:

- use swb to manage large complex experiments
- understand impact of gate length on Vth, SS, Ion, Ioff, DIBL etc.

8.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect
- swb
- nmos_physics.gzp (25 MB - so large) from /scratch folder

8.3 Get Files

To get the swb project file, log on to niu003, run

```
cd
/scratch/getfiles.sh
```

Then, in a terminal, run

```
swb
```

Load in the nmos_physics.gzp, and `extract all` into a folder of your choice.

8.4 Gate Length Impact

8.4.1 Threshold Voltage Roll-off and Reverse Short Channel Effect (RSCE)

As gate length scales down from 1um to 30nm, run all the Id-Vg simulations, then run the very last "Roll-off" node, only the last is meaningful, it is a trick to make plots of extracted variables versus lgate.

You can always open up a spreadsheet - move your mouse over one of the icon that says `spreadsheet` if you wish to do so.

Follow on-screen instruction. The 2nd screen plots out the current defined threshold voltage at low Vd and high Vd (50mV and 1.2V), vti, and DIBL, defined as the amount of vti change per voltage change in Vd, as a function of lgate:

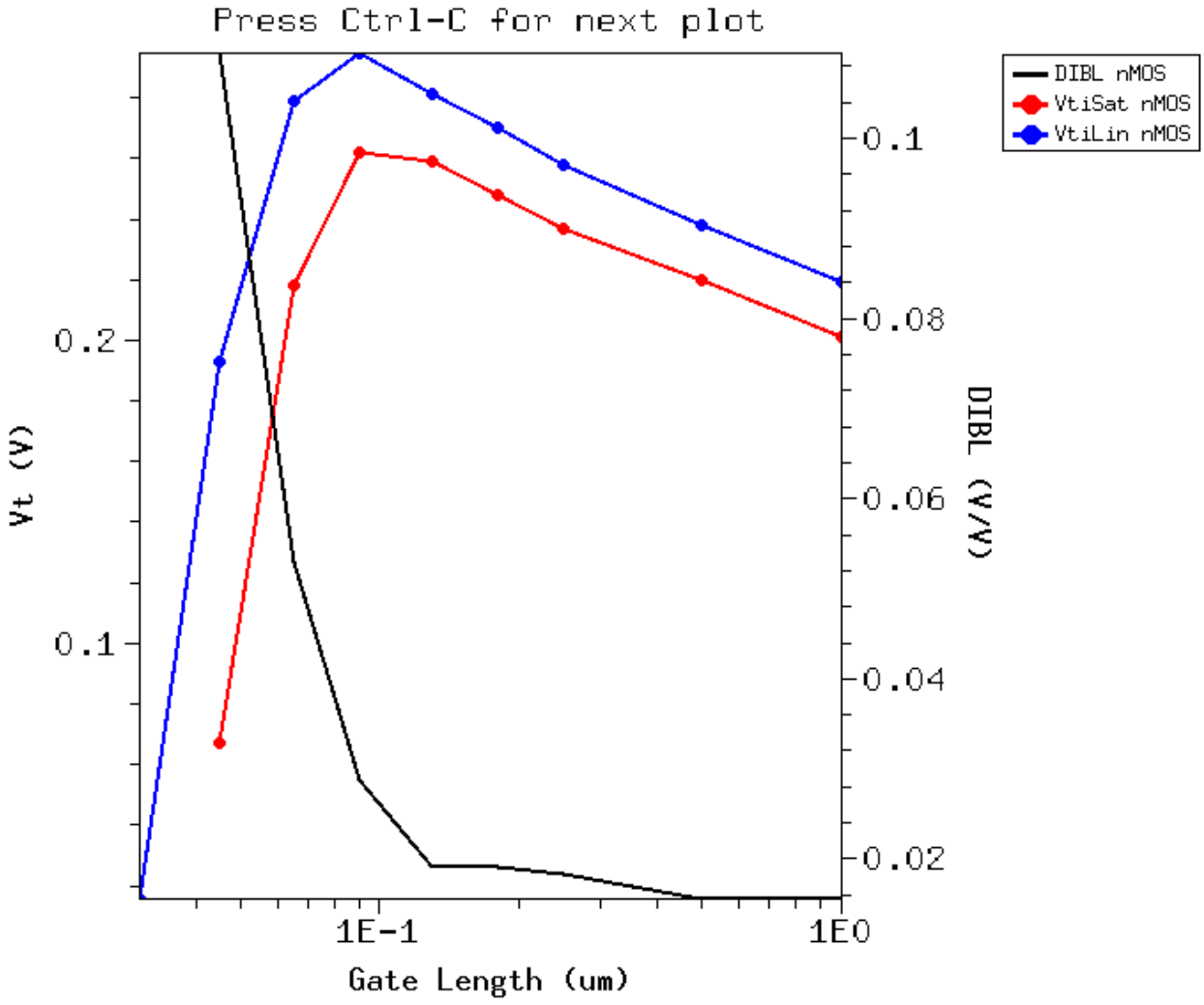


Figure 8.1: Vti linear, Vti saturation and DIBL vs lgate from 30nm to 1um.

We see clearly with decreasing lgate, the Vt's *increases* first, due to the rising *average* channel doping, because of the halo doping, but eventually *decreases* with further scaling down of lgate, due to the increasing impact of the drain to channel junction on surface potential with decreasing gate length.

You can examine the tdr files as necessary for deeper insights.

I suggest that you plot out the surface doping or near surface doping along the channel. It is a 2D effect, yes. But a 1D plot along the channel at the surface will allow you to better see the numbers, and how the doping along the channel changes as you scale down lgate.

Once you are in tecplot_sv, and have made the cuts, it is easy to quickly inspect other variables, like potential and n. Save some plots, you should then be able to link what happens microscopically inside your device with its electrical characteristics at the terminals.

Such insights will also be very useful later on in our next semester's Elec7710 class on more advanced topics, including device design and compact modeling.

You can also change halo doping level or make it very low compared to substrate doping, and then see how your results change.

Below are more plots from running the Roll-off node, I am going to leave analysis to you for now.

8.4.2 Off State Current Ioff

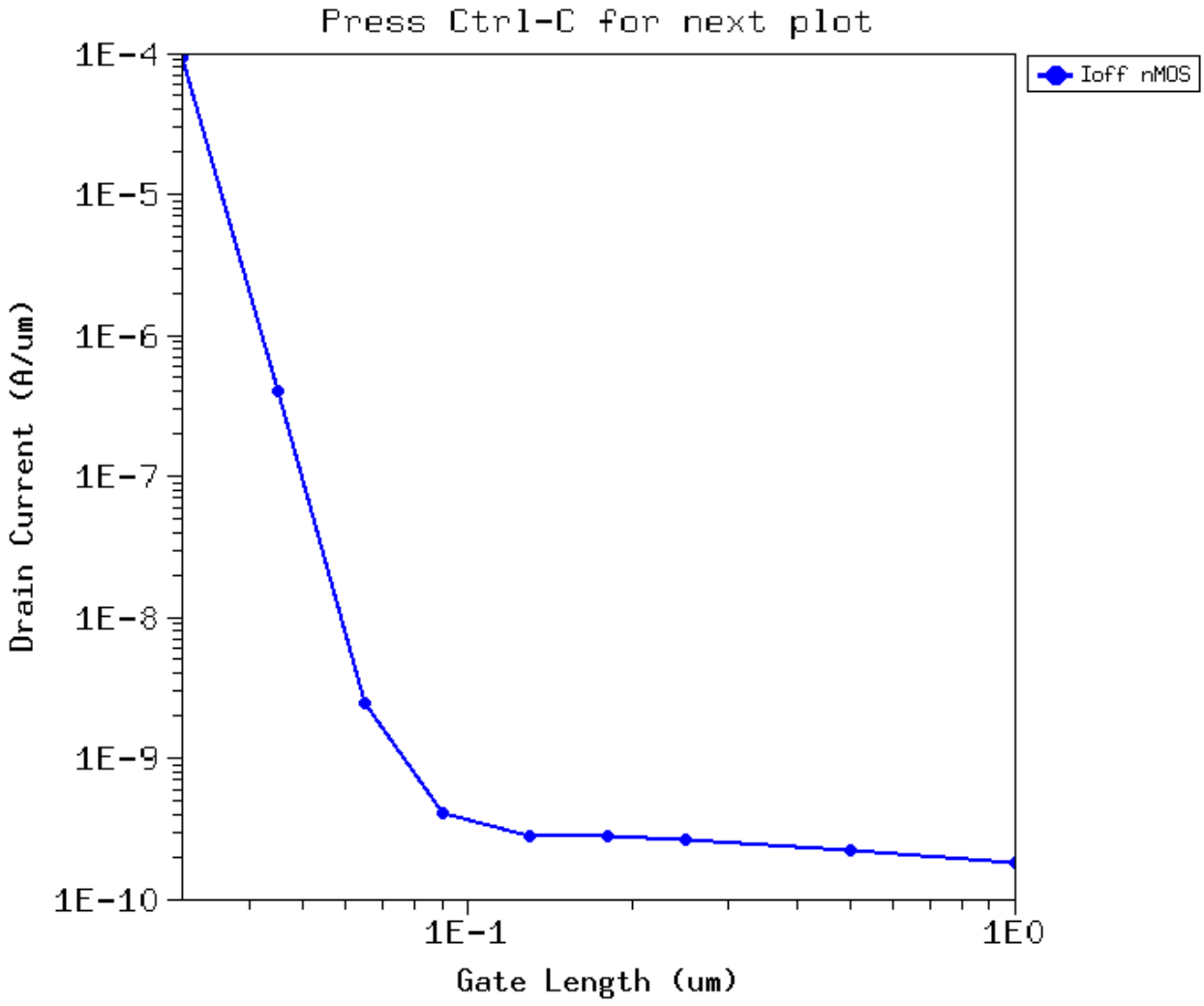


Figure 8.2: Ioff vs lgate from 30nm to 1um.

8.4.3 On State Currents and Ion

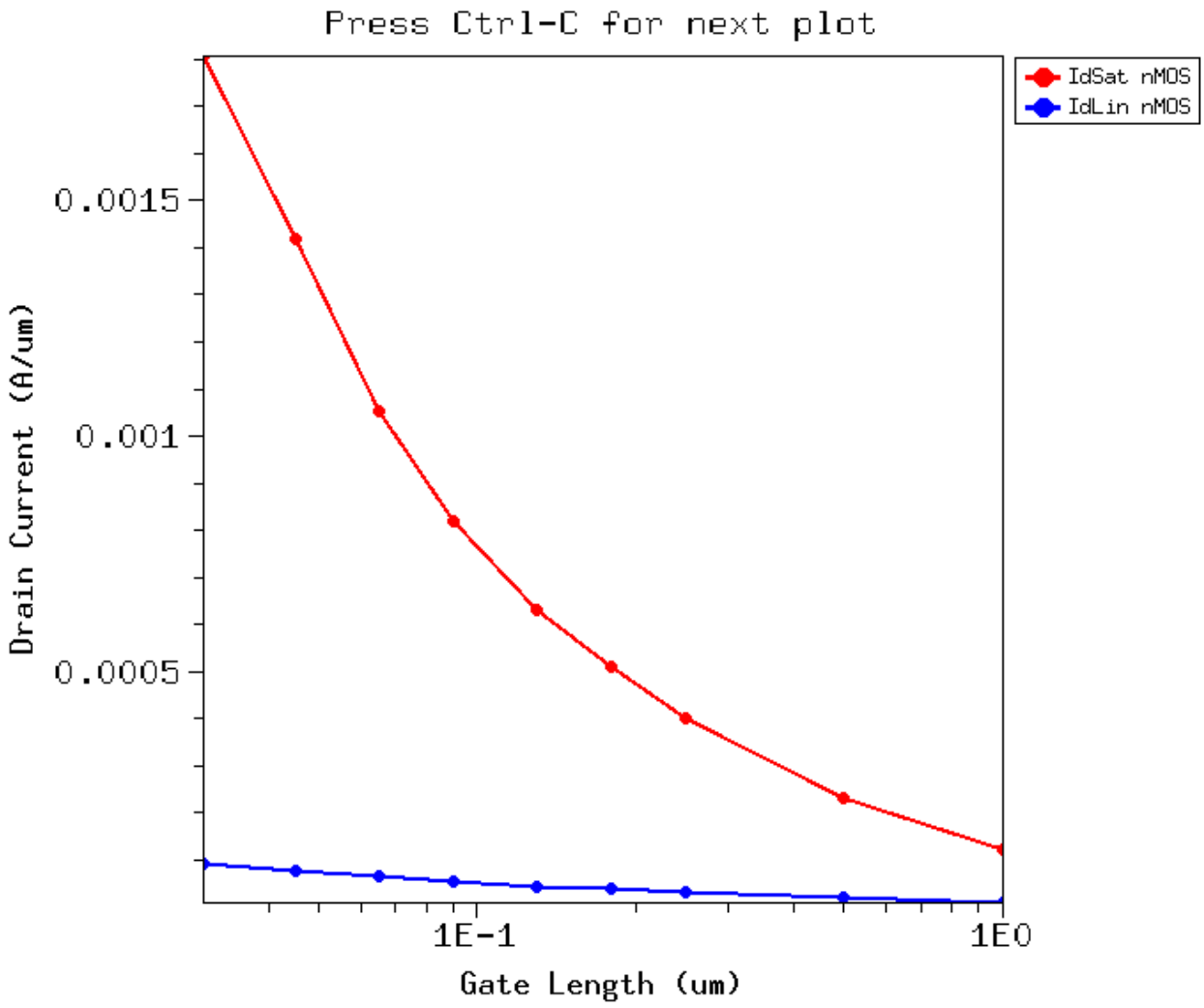


Figure 8.3: I(vg=vdd) vs lgate from 30nm to 1um. I(vg=vdd, vd=vdd) is called Ion.

8.4.4 Ioff vs Ion

8.4.5 Gm vs lgate

8.5 Operating swb

You can refer to the swb user guide located in that same folder as sdevice, swb_ug.pdf, for more information on using swb.

For your class project of IdVg_lin simulation, you can modify the Vd used in the Vg sweep. There are a few ways to do this. L The simplest way is to just modify the value of the parameter `Vdlin` from 0.05V to say 0.5V.

To modify the value of `Vdlin`, select it. The right click, `Edit values` as shown below:

Remember to preprocess and then run the IdVg_lin swb nodes involved. Hole control to select multiple nodes.

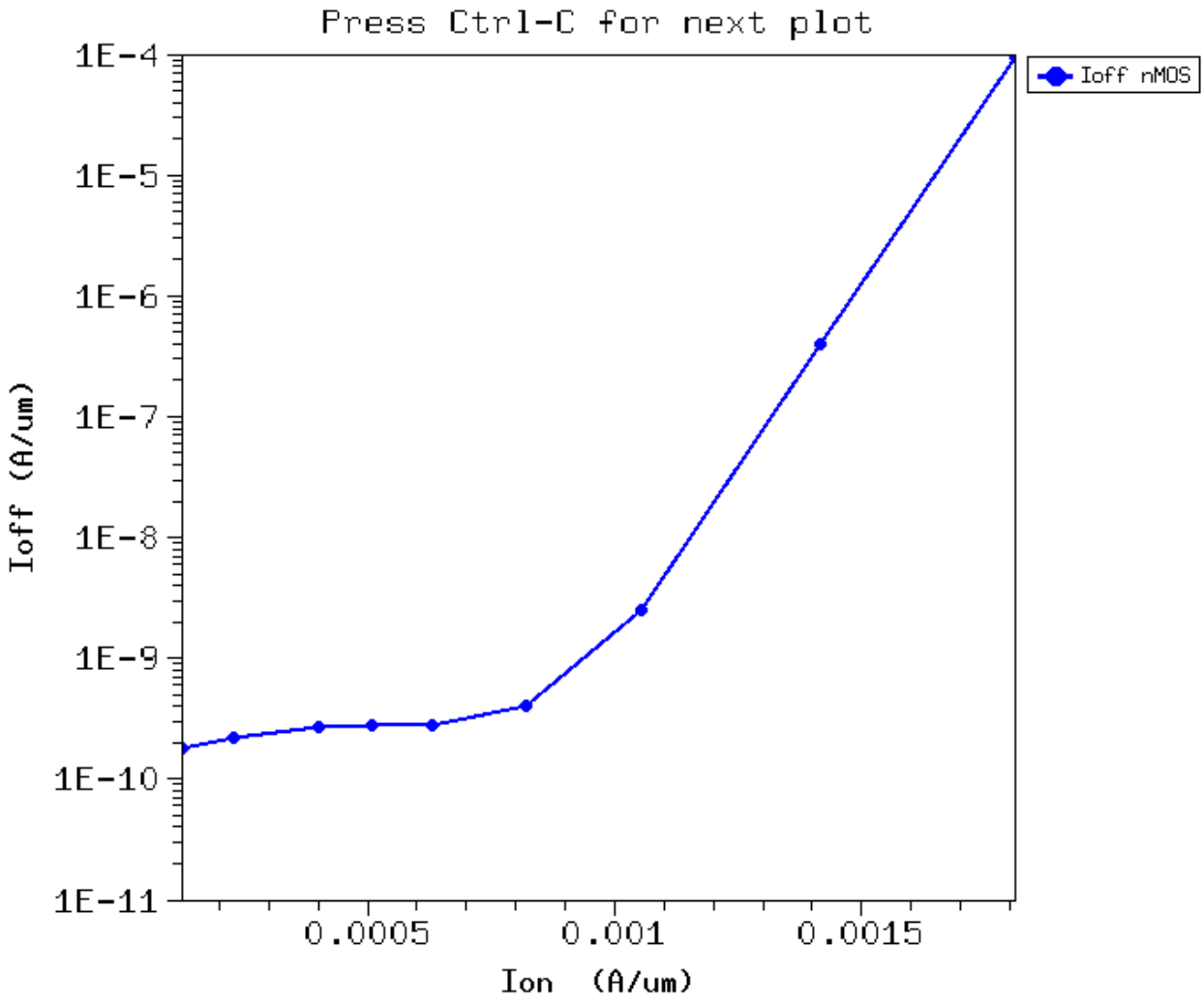


Figure 8.4: Ioff vs Ion from 30nm to 1um. I(vg=vdd, vd=vdd) is called Ion.

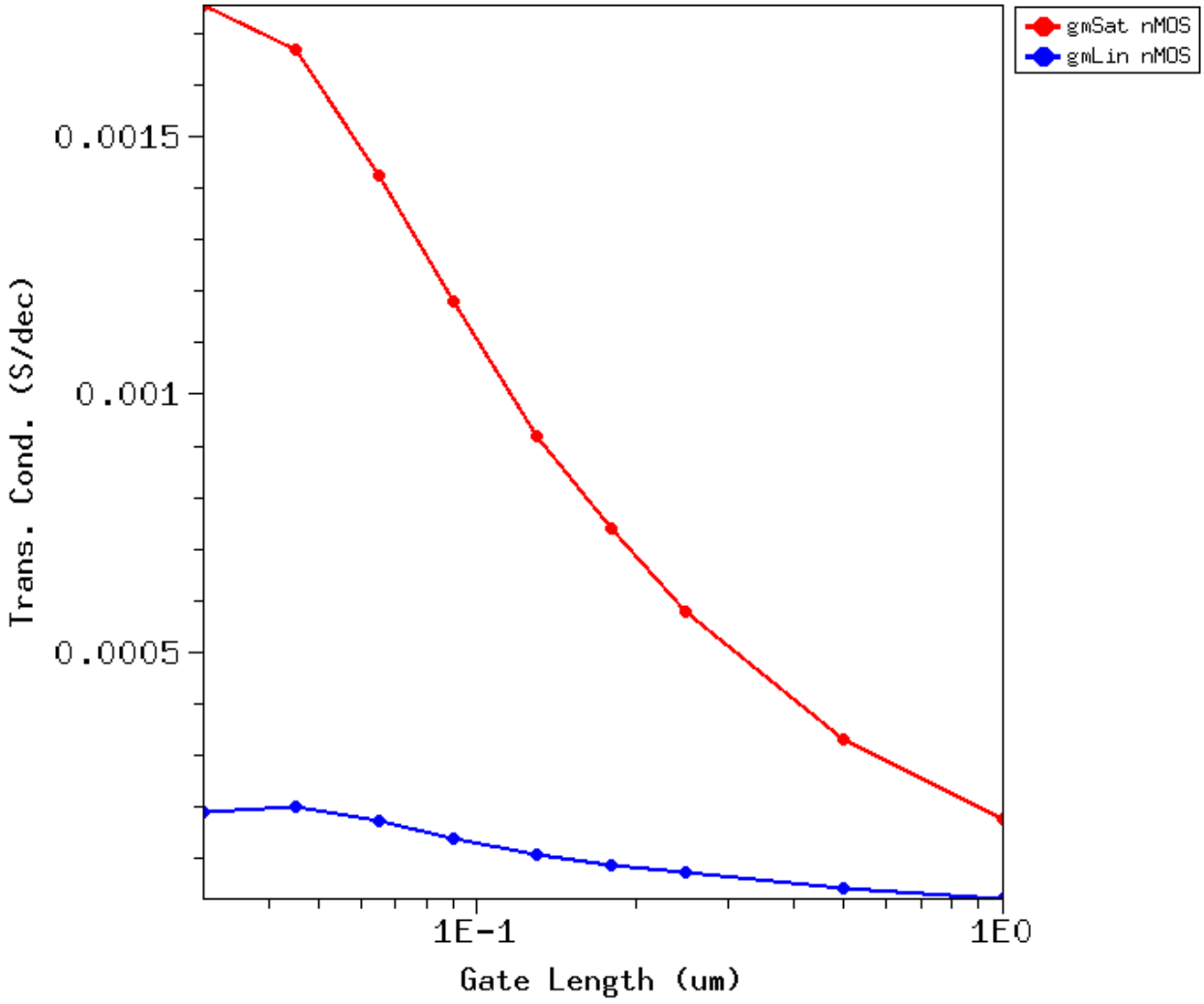


Figure 8.5: Gm in linear and saturation vs lgate from 30nm to 1um.

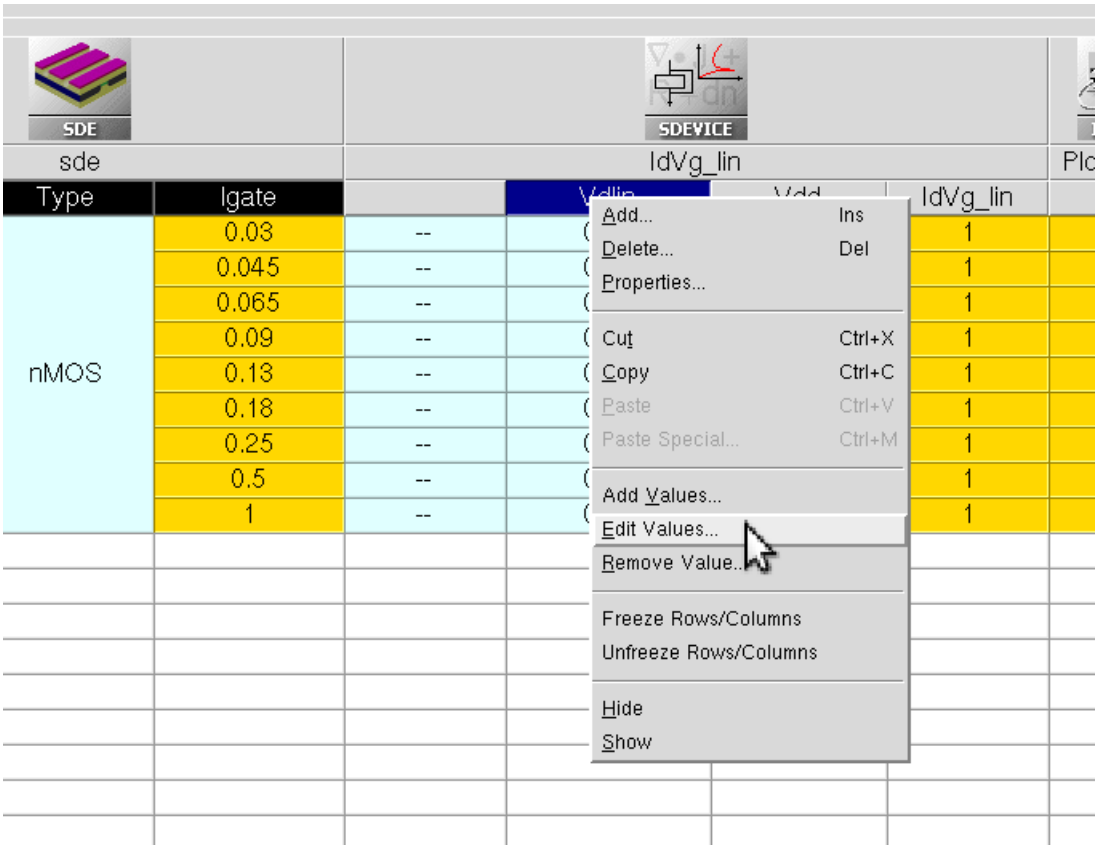


Figure 8.6: Editing parameter value.

Then you should preprocess and run the inspect nodes as well.

Finally, update the variables in swb manually if it has not been set to auto update.

CHAPTER

NINE

SHORT CHANNEL EFFECT AND REVERSE SHORT CHANNEL EFFECT

9.1 Objectives

- visualizing short channel effects
- understanding short channel effect
- understanding reverse short channel effect
- explore halo doping's impact
- understand halo doping's impact
- learn to formatting 2-D contours to make meaningful comparisons between structures

9.2 Required Programs

- sde
- sdevice
- tecplot_sv
- inspect

9.3 Reading Assignment

Read section 3.2, Short Channel MOSFET, on page 175. There is not much discussion on halo doping, other than a brief mentioning in passing in the later sections.

9.4 TCAD Simulations Required

You pretty much can take the existing .tdr files and existing command files to practice this yourself.

To see obvious short channel effects, try simulating Id-Vg and Id-Vd with the .tdr files saved for the 45nm and 1um MOSFETs. Try varying the halo doping level. A low halo doping of 1e16 is used as "no halo" in this chapter. The default halo doping is 1.5e18.

Below I will focus on the 45nm and 1um comparison, at Vg=0V, Vd=Vdd (1.2V), for both with and without halo doping.

9.5 2-D Potential Contours and Ec

Let us consider non-halo devices first.

Launch tecplot_sv, load in the tdr files for non-halo 45nm and 1um devices, at Vg=0V and Vd=Vdd (1.2V).

Choose your settings so that you can see the actions going on in the channel region. Make x and y axis independent in axis settings so that we can maximize the screen space. Most of the action is at the surface, so we want to use our vertical space to show only the actions at the surface.

Similarly, use different x-axis range for short channel and long channel devices, at least do this for the contour plots.

Then take a 1D cut along the surface. Play with your settings for best visual effects.

The 2-D potential contours and surface conduction band energy plots for the 45nm and 1um devices are shown in figure 9.1. Biasing conditions are Vg=0V and Vd=Vdd, which represent the worst case from a standby power standpoint in VLSI circuits.

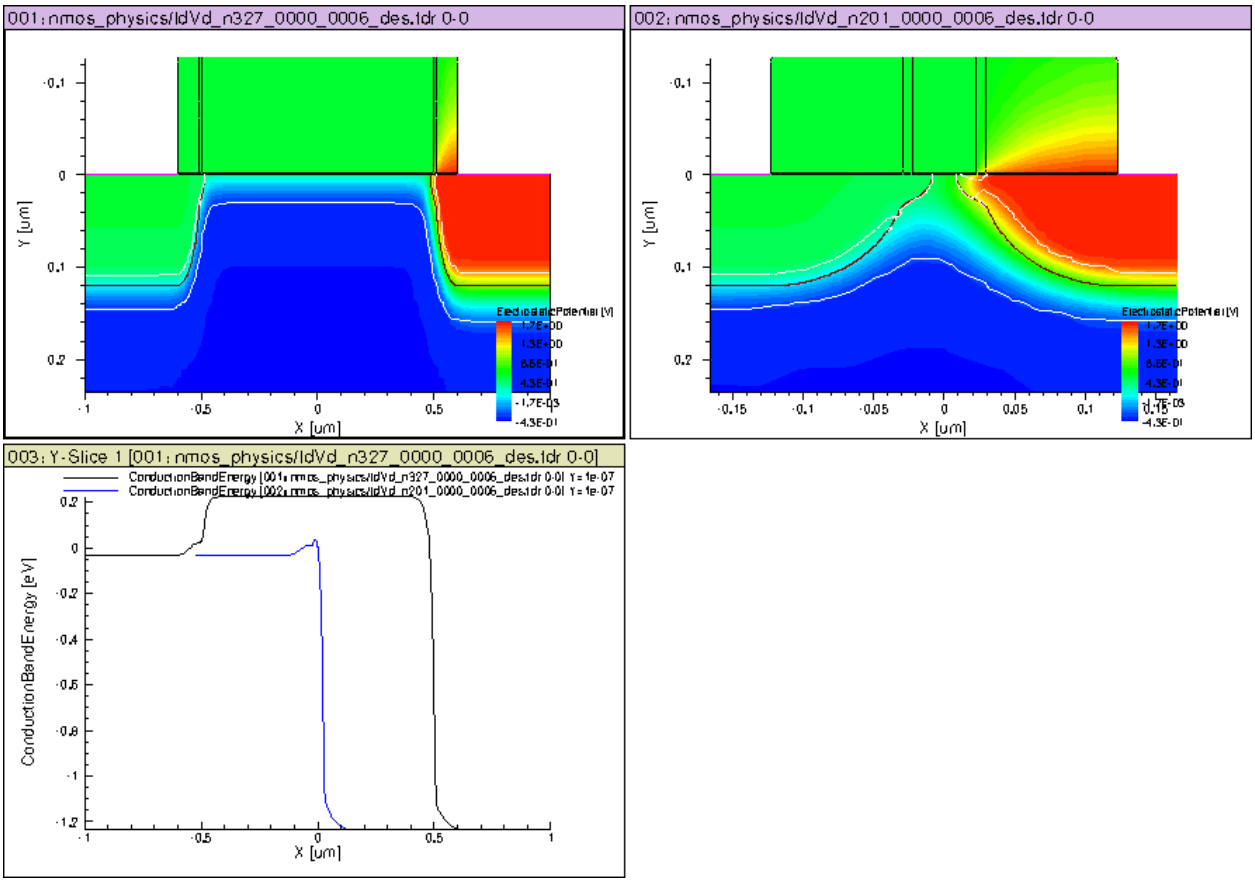


Figure 9.1: 2-D potential and 1-D surface conduction band energy along channel comparison between 45nm and 1um devices. Vg=0V, Vd=Vdd.

You could of course also plot out conduction band contour plots. I have indeed done so. Personally I find it more visually pleasant to look at potential contours and surface conduction band energy plots.

I have forced tecplot_sv to use the same contour levels in both channel lengths, to ease comparison. This is not default setting, so you want to keep this in mind in future comparisons.

A number of useful observations can be made:

1. Surface potential in the 1um device is well behaved, uniform along the channel. Potential contours are largely parallel to the surface.
2. Potential in the channel is much higher in the 45nm device.
3. Potential below the surface, also called sub-surface, is much higher in the 45nm device.
4. We cannot tell any laterally uniform region in the 45nm device.
5. The depletion boundary, indicated by the white line, is much deeper in the 45nm device for the same gate voltage. This is an indication of lower threshold voltage in the 45nm device.
6. At the surface, the potential barrier faced by source electrons is much lower in the 45nm device. As we have shown in previous classes, this barrier lowering at short channel length is strongly dependent on the drain voltage, and thus called drain induced barrier lowering (DIBL).
7. DIBL is non-existent in the long channel device.

Explore yourself to look at about everything you can think of, potential, conduction band energy Ec, eDensity n, eCurrentDensity jn, eVelocity vn.

Electron density is compared in figure 9.2. There are lot more electrons in the surface region of the channel in the 45nm device than in the 1um device.

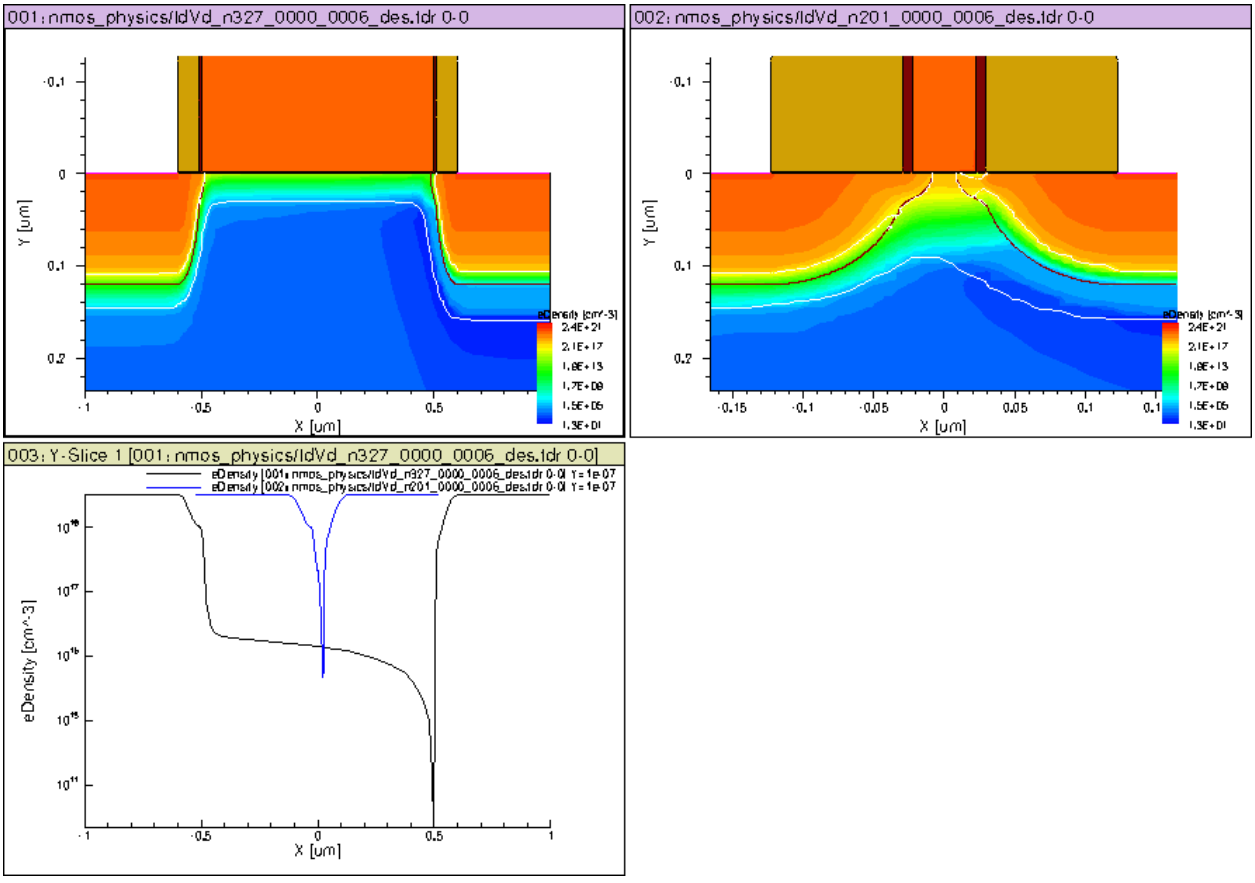


Figure 9.2: 2-D contours and 1-D surface plot of eDensity n along channel in 45nm and 1um devices. Vg=0V, Vd=Vdd.

As you go along with your exploration, save some screen shots for later use. I have generated tons of such plots in preparation for this year's lecture alone. You likely will need to redo your plots a few times before you are satisfied with its looks.

9.6 Id-Vg

figure 9.3 shows Id-Vg at Vd=0.05V and Vdd in 1um and 45nm devices with a low halo doping level that is much smaller than substrate doping level. We will refer to this as "without halo", for all practical purposes.

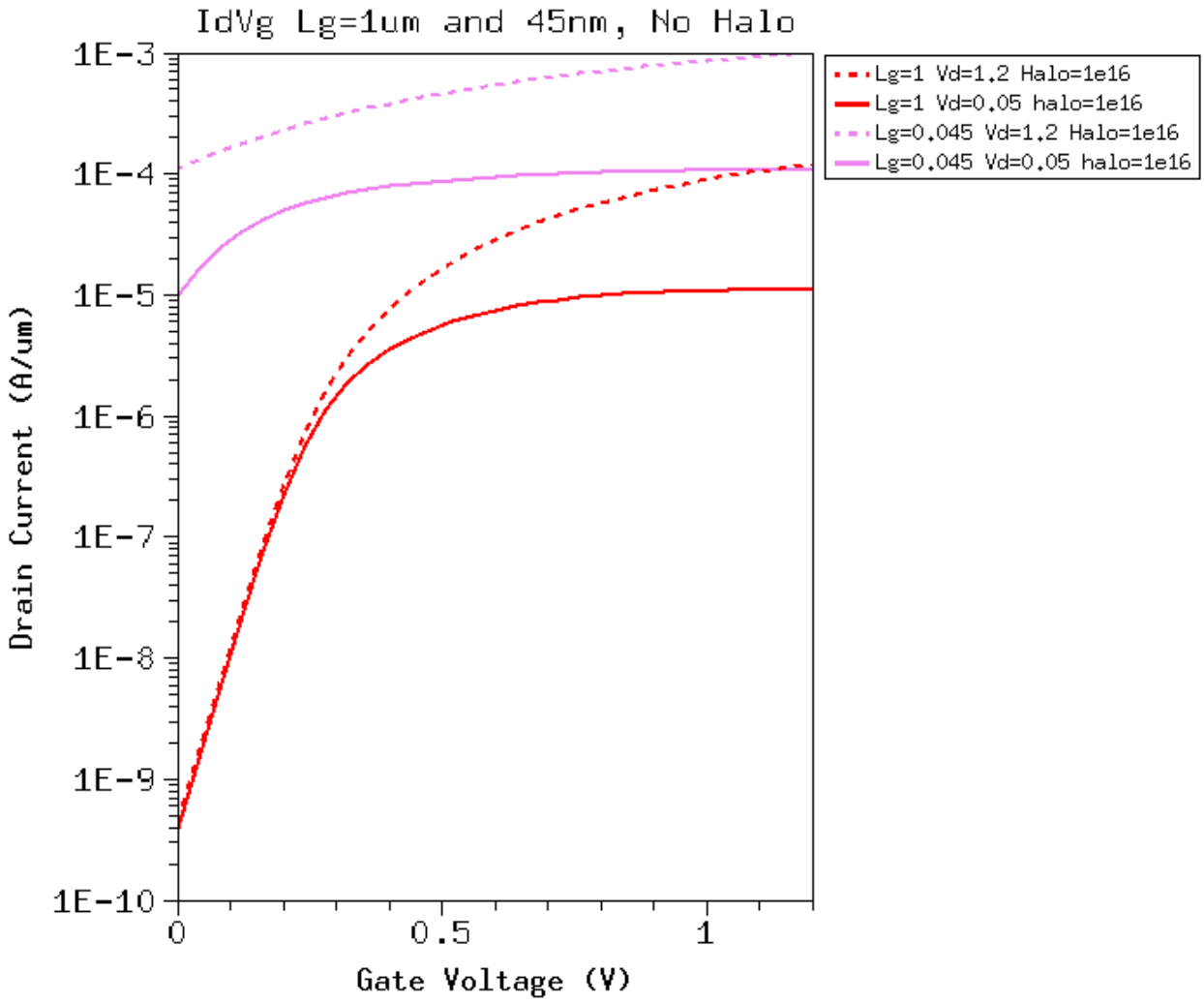


Figure 9.3: Id-Vg at Vd=0.05V and Vdd in 1um and 45nm devices without halo, realized with a low halo doping level.

Note that in subthreshold region, Id is about the same for Vd=0.05V and Vdd in the 1um device, as expected from ideal long channel transistor subthreshold theory. In the 45nm device, the Id at Vd=0.05V in subthreshold region shows very weak gate control. The situation is even worse at Vd=Vdd.

With a high halo doping level, the gate control over in the 45nm device Id is much improved, as can be seen from figure 9.4. Id turns off with decreasing Vg much more effectively. The drain voltage still has much more impact on subthreshold current than in the 1um device, but gate control is still quite obvious.

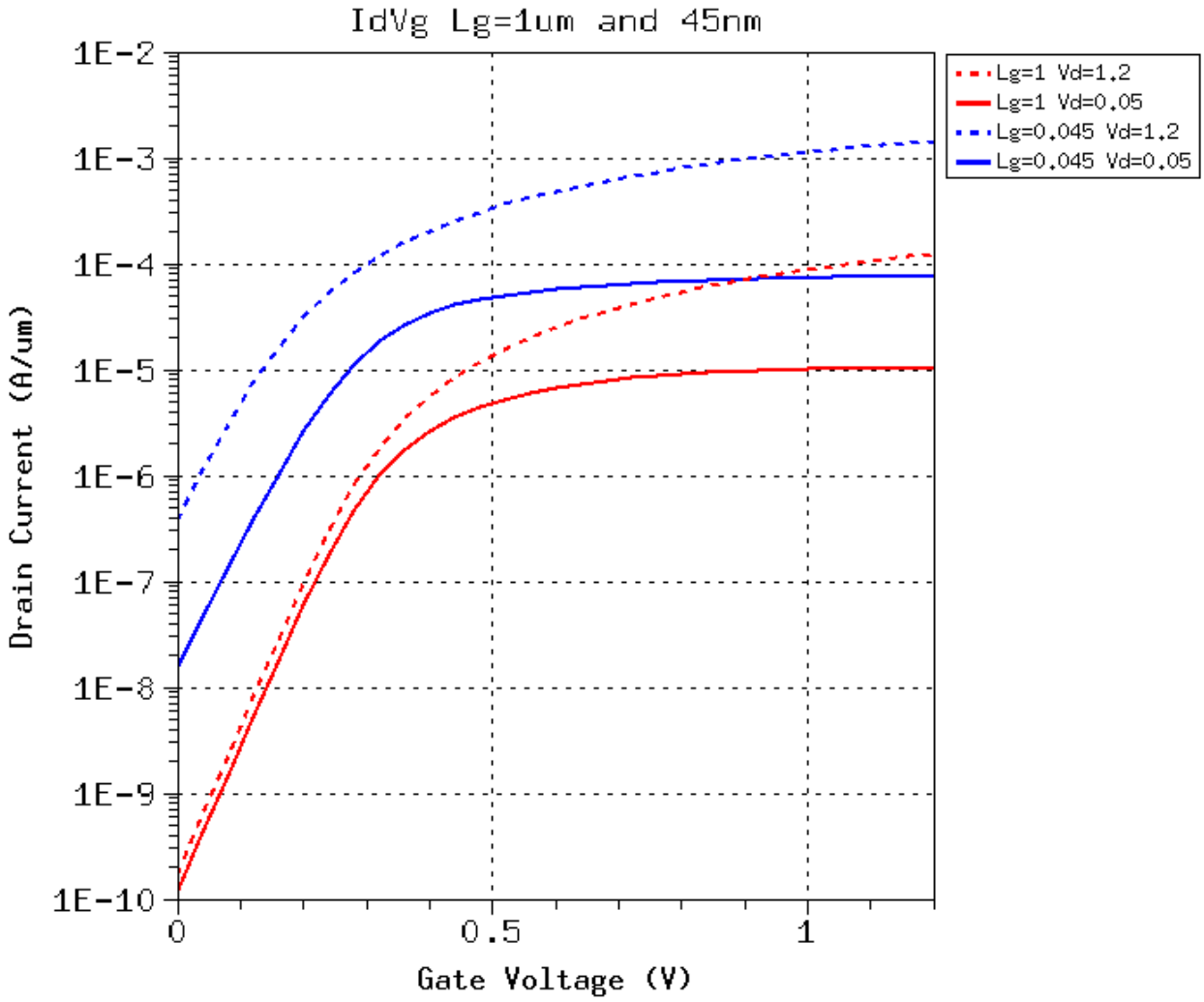


Figure 9.4: Id-Vg at Vd=0.05V and Vdd in 1um and 45nm devices with a high halo doping level.

Let me point out in passing that in the 1um halo device, the subthreshold Id at Vd=Vdd is higher than at Vd=0.05V by a *larger* amount than in the 1um device without halo. This cannot be explained by long channel transistor theory, and is unique to halo doping. Intuitive explanations were proposed in recent years, interested students are encouraged to read relevant research papers.

9.7 Threshold Voltage vs Gate Length

Historically, short channel effect (SCE) is mainly used to refer to the decrease of threshold voltage with decreasing gate length. There are many other physical effects at short gate length of course. There were numerous publications on the topic at one time, I personally also published several papers on the topic back in the early 90s as a grad student.

figure 9.5 shows the linear region threshold voltage obtained from peak gm extrapolation as a function of gate length, for devices with and without halo doping (implemented using different doping levels).

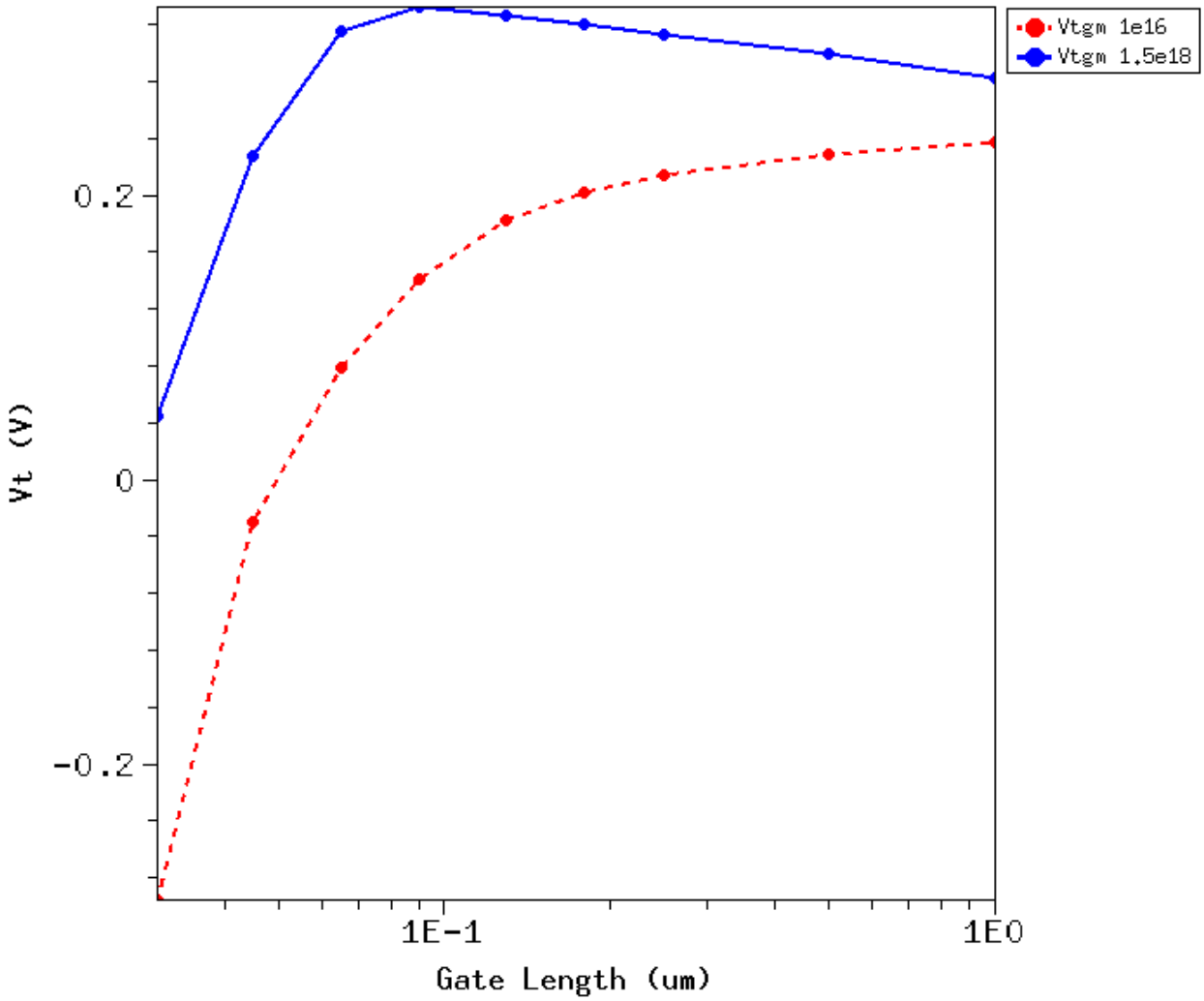


Figure 9.5: Vd=0.05V threshold voltage from peak gm extrapolation versus gate length. Both halo and non-halo devices are shown.

Without halo doping, threshold voltage decreases with gate length, as the source/drain junction lateral field penetrates into the channel, raising the surface potential. As a result, less gate voltage is required for the surface potential to

reach the same threshold value. Since 45nm is so short compared to the lateral depletion width of the source/drain to channel junctions, the whole channel is now under the influence of the source/drain junctions.

With a high drain voltage of Vdd, the situation is only worsened. If you repeat the surface plot for multiple Vd, you will find that the peak Ec point will move towards the source with increasing Vd, and the barrier will be lowered with increasing Vd too. I have shown you one such plot in class, you can generate one yourself.

This of course is a way too simplified picture of reality, which is inherently 2-D, while threshold surface potential was a concept from 1-D MOS structure analysis. When we have so much lateral surface potential variation, most of the concepts such as threshold surface potential, even threshold voltage, lose its classic meanings. However, in practice, we continue to use the long channel like equations, so we continue to use these concepts like threshold voltage. We should keep in mind that more complex nature of short channel device.

We have seen from figure 9.5 that with a high halo doping level of 1.5e18, the threshold voltage decrease with gate length scaling has been greatly reduced. In other words, halo doping is effective in suppressing short channel effect. Next let us take a deeper look at how halo helps suppressing short channel effect.

9.8 How Halo Helps

Recall from Elec6700 that PN junction depletion thickness can be reduced by increasing doping. The situation is similar. The pocket halo doping placed near the source/drain reduces the depletion thickness of the source/drain to channel junctions. Effectively, the lateral distance over which the source/drain have control is reduced, or its impact is weaker for the same distance from source/drain.

To illustrate this, 2-D potential contours of 45nm and 1um devices with and without halo doping are compared in figure 9.6. Biasing conditions are the same as above, Vg=0V and Vd=Vdd.

It takes some patience and efforts to have all the plots formatted in the right way with the right scale for x and y axes, suitable for apple-to-apple comparison. The following considerations went into formatting these plots:

1. The same scale is used for the x and y axes for the same gate length.
2. The y-axis scale for all four plots is made the same to allow comparison of depletion thickness.
3. The x-axis scale for the 1um contours is set to a larger value than for the 45nm contours to make the whole channel visible. Default is the same x-axis scale for short and long channel, which means you can only see the center portion of the longer device.

Without halo, in the 45nm device, the whole area between source and drain is flooded with high potential, and the depletion region is much thicker as indicated by the white curve, the depletion boundary. Source/drain's control over the channel is effectively reduced by the halo doping.

In the 1um device, depletion thickness is clearly reduced in the halo doping region. The depletion boundary moves upward right where the halo doping is.

Note: The increase of Vth with decreasing gate length is due to an increase of the average channel doping, as well as the reduced short channel effect. Again, this is an approximate picture, as the doping is not uniform along the whole channel laterally.

9.9 Reverse Short Channel Effect (RSCE)

An inspection of figure 9.5 immediately shows the signature of halo doping, that is, an increase of threshold voltage with decreasing channel length, known as reverse short channel effect (RSCE), as it is opposite to short channel effect, i.e. decrease of threshold voltage with decreasing gate length. For shortest gate length, Vth starts to decrease again.

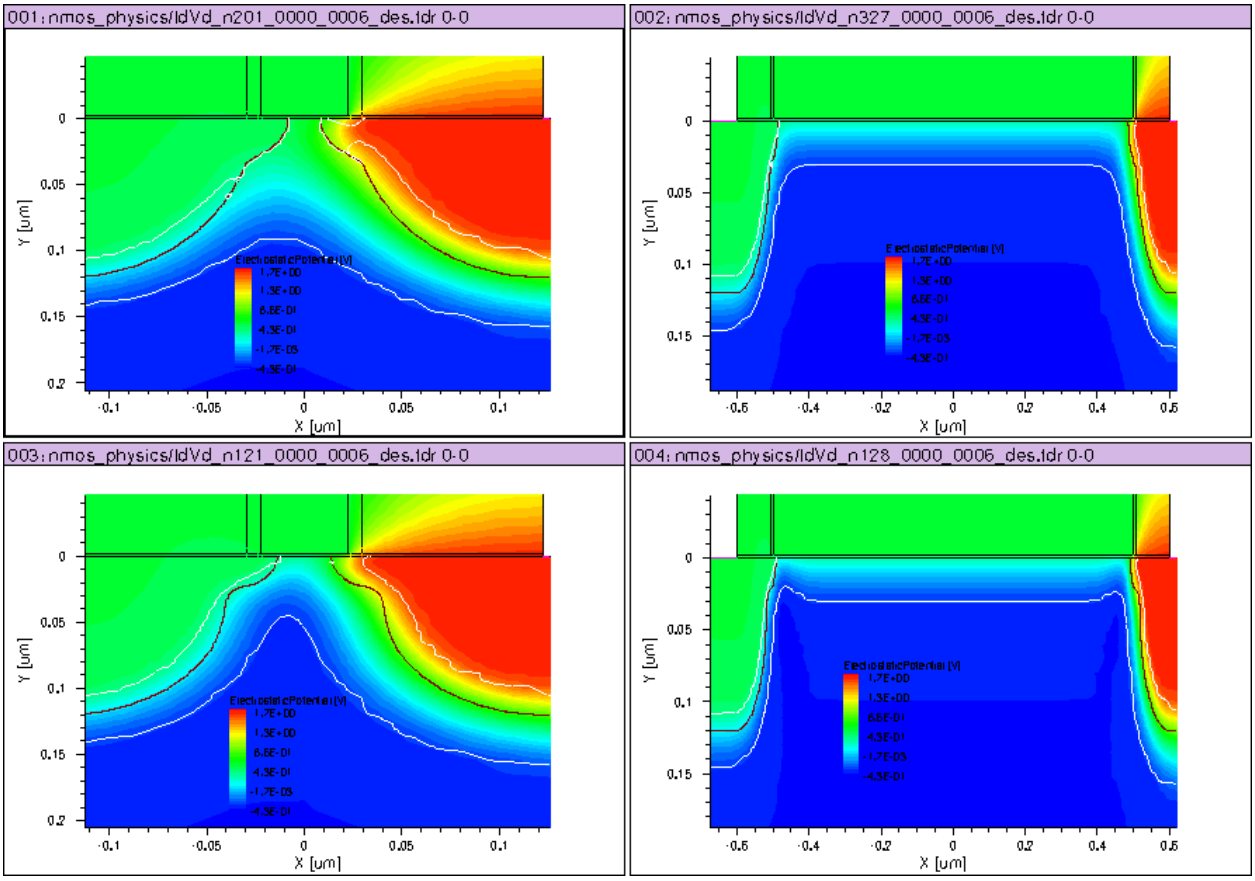


Figure 9.6: 2-D potential contour at Vg=0V and Vd=Vdd, for 45nm and 1um MOSFETs, with and without halo doping.

9.10 VTI and DIBL effects

figure 9.7 shows current defined threshold voltages and DIBL versus gate length. Both halo and non-halo devices are shown. For short lengths, DIBL is stronger without halo.

However, with halo, DIBL is strong *with* halo in the long channel devices. This is not an error. It is real. The exact physics on this is beyond the scope of this class. I can recommend some recent research papers on this topic if interested.

DIBL is fortunately weak in long channel devices to begin with, so some degradation due to halo does not present a big problem. It is mainly a concern for analog circuits that often use long channel devices. The consequence is an increase of the output conductance.

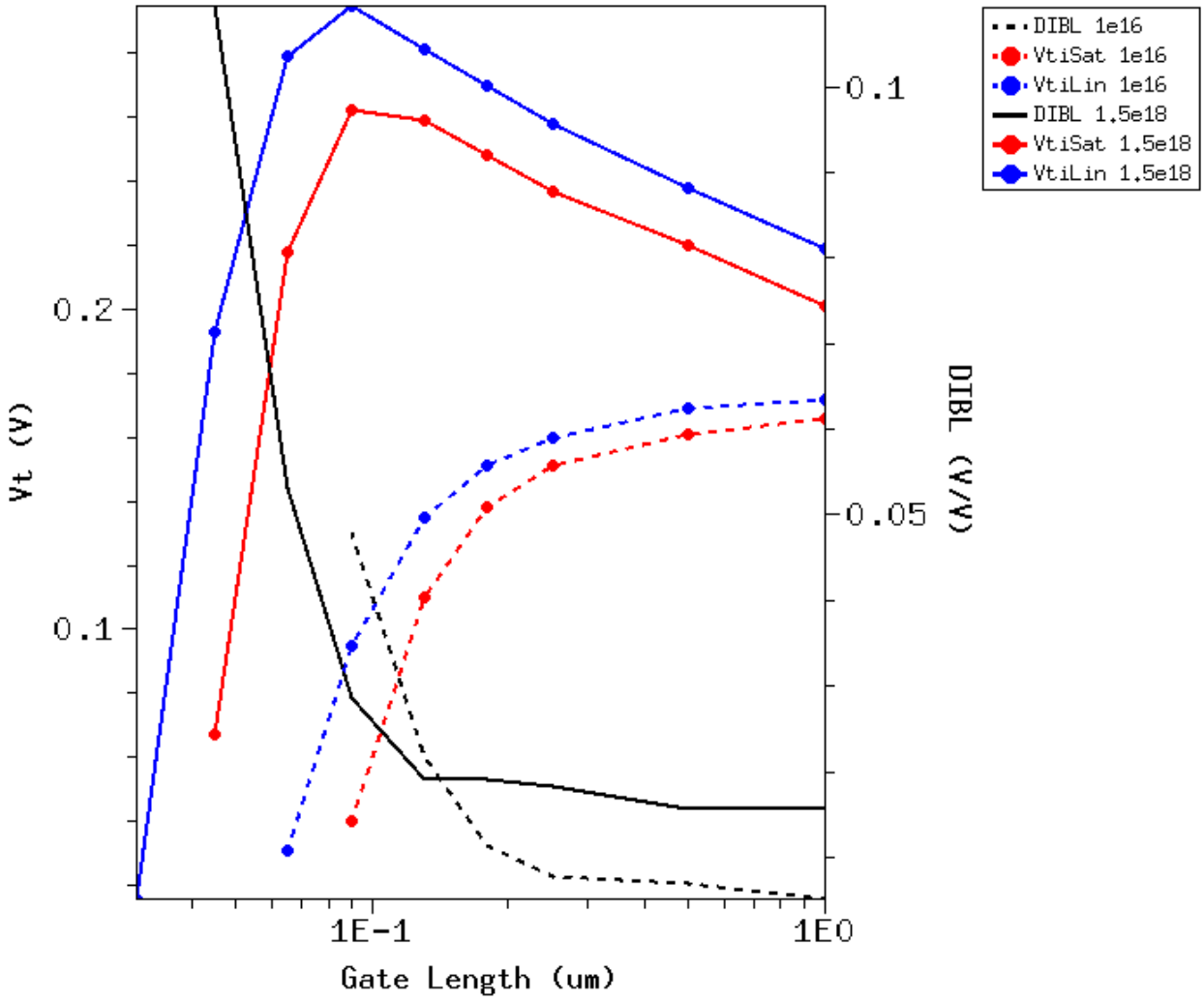


Figure 9.7: Current defined threshold voltages and DIBL versus gate length. Both halo and non-halo devices are shown.

9.11 Subthreshold Swing

The benefits of halo also show up as much reduced subthreshold swing at shorter gate length, as can be seen from figure 9.8. The improvement for Vd=Vdd is even more obvious.

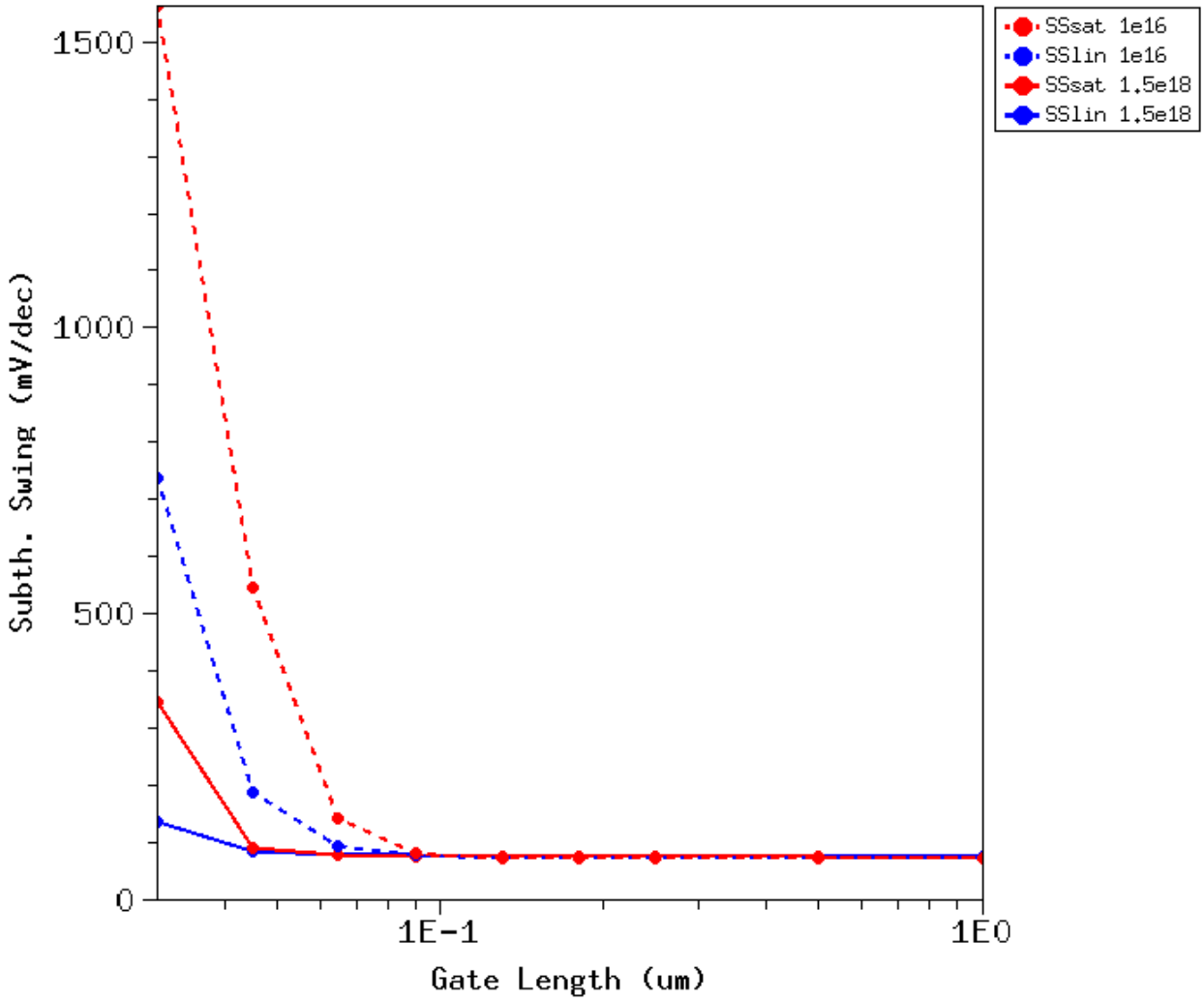


Figure 9.8: Subthreshold swing (SS) versus gate length for linear and saturation operation. Both halo and non-halo devices are shown.

The physics is similar to how halo reduced short channel effect.

9.12 Homework

As we have just had a midterm project, we do not have any official new homework this week. Think a bit more about what you did, the doubts you have, and how you can better organize your data and plots to support your conclusions. Use the suggestions I made during today (Nov 4)'s lecture.

I also encourage you to reproduce figure 9.5 yourself. I will address issues you may experience next class.

Read section 3.2, Short Channel MOSFET, on page 175.

CHAPTER

TEN

CLASS FINAL PROJECT AND EXAM

10.1 Final Exam

The class final exam will be open book, open notes, open computer, and open Internet. The problems will mainly come from our notes and homework assignments. For instance, you may be asked to explain how halo doping affects the threshold voltage versus gate length curve, or how threshold voltage choice affects the Ion and Ioff of a MOS transistor, or how multiple cores help with performance/power tradeoff.

10.2 Final Project

You need to do some literature search and find a topic in the general area of semiconductor devices, including, but not limited to, physics, modeling, fabrication, measurement or simulation of either field-effect or heterojunction bipolar transistors.

Your project should not be just about literature survey, and should have a practice component. It can be either Sentaurus TCAD simulation, or experimental measurement, or both.

You could also choose to expand your mid term project on scaled CMOS physics and characterization, and address several additional requirements.

Below are a few "new" topics we have touched on recently you may consider:

1. Quantum Effect on CMOS Performance. You can simulate CMOS performance with and without quantum effects, analyze simulations details to support your observations, as we have done throughout the semester.

   There are multiple options for simulating quantum effects. Look for *QuantumPotential* and *Density Gradient* in the sdevice user manual.

2. FinFET or Tri-gate CMOS. There are some examples in the Application Library folder under the installation directory. You may choose to use 2-D double-gate MOSFET for simplicity.

3. Nano wire MOSFETs. I had sent you a couple of links earlier by email.

4. Low-frequency Noise Measurement. This can be done with a time domain sampling measurement with a semiconductor parameter analyzer.

5. Temperature dependence of CMOS transistors. In my lab we can take on-wafer measurements from -10C to 140C.

You need to let me know your topic of choice and discuss with me your written project plan by Friday the 18th, 3:00pm. It can be just a short list of simulation or measurement. I will give you feedback.

10.3 Written Report

Your final report is due 3:00pm, Dec 2nd. Please take into account my feedback to you on the mid term report in writing your final report.

Please use IEEE Transactions on Electron Device paper format. Your report should be no longer than 10 pages.

Your Sentaurus simulation files should also be submitted, with a readme file that explains how your files should be run.

CHAPTER

ELEVEN

FINAL EXAM SCOPE

The class final exam will be open book, open notes, open computer, and open Internet. The problems will mainly come from our notes and homework assignments. For instance, you may be asked to explain how halo doping affects the threshold voltage versus gate length curve, or how threshold voltage choice affects the Ion and Ioff of a MOS transistor, or how multiple cores help with performance/power trade-off.

A list of topics are given below for you to prepare for the final exam. You can refer to the journal and HTML notes (posted on blackbaord and this HTML site respectively), our text book, and use IEEE explorer for exam preparation.

We will address all of these issues in more depths next semester, Spring 2012, in ELEC7710, Field Effect Transistor, where we will also address physics based compact modeling, TCAD (in 6710, I have basically provided existing codes for you to use, you will learn how to write your own codes, including building 2d and 3d devices in 7710), RF and microwave properties, including noise and linearity, CMOS low-noise amplifier design, as well as on-wafer characterization using automatic probe station and ICCAP which you have seen in my lab towards the end of the semester, using state-of-the-art CMOS and SiGe BiCMOS wafers. As you may have noted from taking 6710, I generally emphasize hands-on experience in my teaching and will work closely with you on derivation, Matlab/Python coding, TCAD coding / demonstration, as well as Spice/Ads/Cadence tools.

11.1 Band diagrams

Draw PN junction as well as MOS capacitor band diagrams. Should be able to identify accumulation, depletion, inversion, strong inversion. Reviewing relevant notes, journal as well as HTML, will be sufficient.

11.2 Field Effect

Draw surface potential, total charge, and inversion charge as a function of gate voltage. Identify the different operation regions on all of these curves, e.g. accumulation, depletion, inversion, or strong inversion.

You can use existing homework matlab codes to draw these graphs instead of by hand.

11.3 Threshold voltage

Calculate both gate to body and gate to channel (or source) threshold voltage for given gate material, substrate doping and oxide thickness, at different channel to body bias (Vcb).

Reviewing the notes on 3 terminal MOS should be sufficient.

11.4 Internal potential and field

For a long channel MOSFET, sketch how surface potential, inversion charge density, depletion thickness, lateral electric field and electron velocity vary from source to drain at the following biases:

1. Vg << Vth, Vds=0
2. Vg > Vth, Vds < Vdsat
3. Vg > Vth, Vds > Vdsat

You only need to sketch the region from source to the "pinch-off" point. Velocity saturation can be neglected.

11.5 Velocity saturation

How does it affect electron velocity along the channel? How does it impact Id-Vd curves?

Why is velocity saturation more important in shorter channel devices?

11.6 Subthreshold swing (SS)

Understand its definition and how it affects the choice of threshold voltage to achieve a given off current (e.g. Id at Vg=0 has to be less than 1nA, if S=100 mV/decade, what value of Vth should you need in transistor design).

What is the smallest SS one can achieve?

For an ideal long channel MOSFET, if the surface potential increases by 0.9mV per 1mV gate voltage increase, what will the SS be?

11.7 Ion and Ioff

Understand what determines Ion and Ioff.

For instance, for a transistor with Vt=0.3V, SS = 90mV/decade, W=90nm, L=45nm, estimate its 1) Ioff, defined as Id at Vg=0V, and 2) subthreshold Id at Vg=0.15V.

Reviewing the journal note on subthreshold current should be sufficient.

11.8 Short channel effect

Explain how scaling gate length affects threshold voltage.

11.9 DIBL

Understand DIBL effect and its measure. Use tecplot_sv to illustrate how DIBL effect varies with channel length and how it is affected by halo doping in short channel devices.

11.10 Effect of Halo Doping

Explain the effect of halo doping on transistor Id-Vg, Vth-lgate using tecplot_sv. Simulated potential contours with and without halo need to be well understood.

11.11 Channel length scaling limit

Minimum channel length one can achieve highly depends on gate oxide thickness or equivalent gate oxide thickness and device structure, e.g. a tri-gate vs traditional bulk.

Many of the Id-Vg curves you see in textbooks show real bad short channel behavior at gate length of 0.25um. While at the same gate length, in the simulations you have been running, the Id-Vg is ideal and long channel like.

This is mainly because of the difference in oxide thickness.

Historically, the short channel length limit was thought to be 500nm in the 70s, 250nm in the 80s, 100nm in the 90s when I was in grad school, and now 22nm is in production.

So the channel length scaling potential was vastly underestimated, primarily because our ability to make transistor quality thin oxide was underestimated.

11.12 Metal gate, high K

Why metal gates and high K gate dielectric are used in modern CMOS.

11.13 FinFETs, Tri-gate, Nanowire FETs

Such structures are all aimed at increasing gate control over the whole Si channel volume which help keeping sub-threshold swing (SS) close to ideal. DIBL is minimized and sub-surface leakage found in bulk even regular SOI MOSFETs are suppressed.

11.14 Vdd Scaling and Multi-core Processing

Power consumption in CMOS is primarily due to switching. P = C Vdd^2 f. A key point to take away from this equation we derived is that power consumed is not linearly proportional to Vdd, rather, it is proportional to the square of Vdd. This may sound counterintuitive at first, but is satisfying if you think about the capacitive nature of the load in CMOS logic circuits. The charging current dQ/dt on average is also proportional to Q, which is CVdd, leading to the square of Vdd dependence.

Instead of running 1 core at full Vdd and full speed (f), we could run multiple cores at a reduced Vdd and a correspondingly reduced frequency. You should be able to quantify the advantages of such approach in power consumption as well as throughput. Multi-core processing also facilitates power management as individual cores can be turned on and off. See the journal notes for a numerical example.

11.15 CMOS Layout Basics

Multi-finger layout for reducing gate resistance and signal delay.

11.16 Analog Amplifiers

How does the gm * Rout product change with scaling? What is the typical gm*Rout number in 45nm and 32nm CMOS?

11.17 Drain Current Thermal Noise

What is the PSD of drain current noise of a MOSFET in sub-threshold region? How does it compare to 2qIDS with IDS being the dc biasing current?

Above threshold voltage, in saturation region, is the drain current noise smaller or larger than 2qIDS?