Advanced MOSFET Basics

Dr. Lynn Fuller
Webpage: http://people.rit.edu/lffee
Microelectronic Engineering
Rochester Institute of Technology
82 Lomb Memorial Drive
Rochester, NY 14623-5604
Tel (585) 475-2035
Email: Lynn.Fuller@rit.edu
Department Webpage: http://www.rit.edu/kgcoe/microelectronic/
OUTLINE

Introduction
Short Channel vs Long Channel
Effective Channel Length
Sub Threshold Effects
Low Doped Drain
NMOS with N+ Poly Gate
PMOS with N+ Poly Gate
PMOS with P+ Poly Gate
References
INTRODUCTION

The idea is to design a MOSFET that is as small as possible without short channel effects compromising the device performance much.

That is we want the smallest transistor possible that exhibits long channel characteristics.
Long-channel MOSFET is defined as devices with width and length long enough so that edge effects from the four sides can be neglected.

Channel length $L$ must be much greater than the sum of the drain and source depletion widths.
LONG CHANNEL MOSFET I-V CHARACTERISTICS

Family of Curves

+Id
+Vgs
+Vds
+Vd = 0.1 Volt

Saturation Region

Vgs = Vds
(Id)

Non Saturation Region

Vd = 0.1 Volt

Id (Amps)

Subthreshold

Sub Vt Slope
(mV/dec)

Rochester Institute of Technology
Microelectronic Engineering
UNIFORMLY DOPED PN JUNCTION

- Phosphorous donor atom and electron ($\text{P}^+$)
- Ionized Immobile Phosphorous donor atom ($\text{P}_+^-$)
- Ionized Immobile Boron acceptor atom ($\text{B}^-$)
- Boron acceptor atom and hole ($\text{B}_+^-$)

Space Charge Layer

- Charge density, $\rho$
- Electric Field, $\mathbf{E}$
- Potential, $\Psi$

$$qN_A W_1 = qN_D W_2$$

$$+qN_D$$

$$-W_1$$

$$-qN_A$$

$$W_2$$

$$\varepsilon_o$$

$$\Psi_o + V_R$$

Rochester Institute of Technology
Microelectronic Engineering
UNIFORMLY DOPED PN JUNCTION

Built in Voltage: \( \Psi_o = \frac{KT}{q} \ln \left( \frac{N_A N_D}{n_i^2} \right) \)

\( n_i = 1.45 \times 10^{10} \text{ cm}^{-3} \)

Width of Space Charge Layer, \( W \): with reverse bias of \( V_R \) volts

\[ W = \left( W_1 + W_2 \right) = \left[ \left( \frac{2\varepsilon}{q} \right) \left( \Psi_o + V_R \right) \left( \frac{1}{N_A} + \frac{1}{N_D} \right) \right]^{1/2} \]

\( W_1 \) width on p-side \( W_2 \) width on n-side

\[ W_1 = W \left[ \frac{N_D}{(N_A + N_D)} \right] \quad W_2 = W \left[ \frac{N_A}{(N_A + N_D)} \right] \]

Maximum Electric Field:

\[ E_o = - \left[ \left( \frac{2q}{\varepsilon} \right) \left( \Psi_o + V_R \right) \left( \frac{N_A N_D}{(N_A + N_D)} \right) \right]^{1/2} \]

Junction Capacitance per unit area:

\[ C_j' = \varepsilon_o \varepsilon_r / W = \varepsilon_o \varepsilon_r / \left[ \left( \frac{2\varepsilon}{q} \right) \left( \Psi_o + V_R \right) \left( \frac{1}{N_A} + \frac{1}{N_D} \right) \right]^{1/2} \]

\( \varepsilon = \varepsilon_o \varepsilon_r = 8.85 \times 10^{-12} \text{ (11.7) F/m} \)

\( = 8.85 \times 10^{-14} \text{ (11.7) F/cm} \)
EXAMPLE CALCULATIONS

EXAMPLE CALCULATIONS

ROCHESTER INSTITUTE OF TECHNOLOGY
MICROELECTRONIC ENGINEERING

To use this spreadsheet change the values in the white boxes. The rest of the sheet is protected and should not be changed unless you are sure of the consequences. The calculated results are shown in the purple boxes.

<table>
<thead>
<tr>
<th>CONSTANTS</th>
<th>VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>1.38E-23 JK</td>
</tr>
<tr>
<td>q</td>
<td>1.60E-19 Coul</td>
</tr>
<tr>
<td>Ego</td>
<td>1.12 eV</td>
</tr>
<tr>
<td>z0</td>
<td>8.85E-14 F/cm</td>
</tr>
<tr>
<td>z1</td>
<td>11.7</td>
</tr>
<tr>
<td>ni</td>
<td>1.45E+10 cm^-3</td>
</tr>
<tr>
<td>Breakdown E</td>
<td>3.00E+05 V/cm</td>
</tr>
</tbody>
</table>

Exceeds E_{max}!!!

CALCULATIONS:

\[ E_g = E_{go} - (sT^2)/(T+B) \]
\[ n_i^2 = A T^3 e^{-\left(E_g/KT/q\right)} \]
\[ KT/q = 0.0259 \text{ Volts} \]
\[ V_{bi} = (KT/q) \ln\left(ND/N_i^2\right) \]
\[ W = \sqrt{[E_{bi}V_{bi}+V_y]/[1+K+(1+2)/2]} \]
\[ W_1 = W_{0} \left[N_D/(N_D+N_i)\right] \]
\[ W_2 = W_{0} \left[N_i/(N_D+N_i)\right] \]
\[ E_0 = -\left[(E_{bi}+V_y)/(N_D+N_i)\right]^{0.5} \]
\[ V_y = \sqrt{W} \]

\[ I_D = I_s (e V/V_K - 1) \]
\[ I_s = C T^2 \exp\left(qE_0/kT\right) \]

\[ I_s = 3.73E-13 \]

<table>
<thead>
<tr>
<th>V</th>
<th>I_d</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.9</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.8</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.7</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.6</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.5</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.4</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.3</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.2</td>
<td>-3.73E-13</td>
</tr>
<tr>
<td>-0.1</td>
<td>-3.52E-13</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>4.53E-11</td>
</tr>
<tr>
<td>0.2</td>
<td>2.20E-09</td>
</tr>
<tr>
<td>0.3</td>
<td>1.04E-07</td>
</tr>
<tr>
<td>0.4</td>
<td>7.24E-07</td>
</tr>
<tr>
<td>0.5</td>
<td>4.93E-06</td>
</tr>
<tr>
<td>0.6</td>
<td>3.55E-06</td>
</tr>
<tr>
<td>0.7</td>
<td>2.34E-05</td>
</tr>
<tr>
<td>0.8</td>
<td>1.62E-05</td>
</tr>
<tr>
<td>0.9</td>
<td>1.53E-05</td>
</tr>
<tr>
<td>1</td>
<td>8.44E-06</td>
</tr>
<tr>
<td>1.1</td>
<td>5.87E-06</td>
</tr>
<tr>
<td>1.2</td>
<td>3.83E-06</td>
</tr>
<tr>
<td>1.3</td>
<td>2.63E-06</td>
</tr>
<tr>
<td>1.4</td>
<td>1.72E-06</td>
</tr>
<tr>
<td>1.5</td>
<td>9.88E-07</td>
</tr>
<tr>
<td>1.6</td>
<td>5.92E-07</td>
</tr>
<tr>
<td>1.7</td>
<td>3.61E-07</td>
</tr>
<tr>
<td>1.8</td>
<td>2.17E-07</td>
</tr>
<tr>
<td>1.9</td>
<td>1.31E-07</td>
</tr>
<tr>
<td>2</td>
<td>7.98E-08</td>
</tr>
<tr>
<td>2.1</td>
<td>5.26E-08</td>
</tr>
<tr>
<td>2.2</td>
<td>3.27E-08</td>
</tr>
<tr>
<td>2.3</td>
<td>2.04E-08</td>
</tr>
<tr>
<td>2.4</td>
<td>1.29E-08</td>
</tr>
<tr>
<td>2.5</td>
<td>8.29E-09</td>
</tr>
<tr>
<td>2.6</td>
<td>5.46E-09</td>
</tr>
<tr>
<td>2.7</td>
<td>3.59E-09</td>
</tr>
<tr>
<td>2.8</td>
<td>2.43E-09</td>
</tr>
<tr>
<td>2.9</td>
<td>1.62E-09</td>
</tr>
<tr>
<td>3</td>
<td>1.11E-09</td>
</tr>
</tbody>
</table>

© January 29, 2017 Dr. Lynn Fuller, Professor
Sort channel MOSFET is defined as devices with width and length short enough such that the edge effects can not be neglected.

Channel length $L$ is comparable to the depletion widths associated with the drain and source.
CHANNEL LENGTH MODULATION

Channel Length Modulation Parameter $\lambda$

$\lambda = \frac{\text{Slope}}{I_{\text{dsat}}}$

$I_{\text{dsat}} = \frac{\mu W \text{Cox'} (V_g - V_t)^2 (1 + \lambda V_{ds})}{2L}$

NMOS Transistor in Saturation Region
DC Model, $\lambda$ is the channel length modulation parameter and is different for each channel length, $L$. Typical value might be 0.02
TERADA-MUTA METHOD FOR EXTRACTING Leff and Rds

Terada-Muta Method for Leff and Rds

In the linear region ($V_D$ is small):

$$I_D = \frac{\mu W \cdot Cox' \cdot (V_{gs} - V_{t} - V_d/2)}{Leff} \cdot V_D$$

$$Leff = L_m - \Delta L$$

where $\Delta L$ is correction due to processing

$L_m$ is the mask length

$$R_m = \frac{V_D}{I_D} = \text{measured resistance}$$

$$= R_{ds} + \frac{(L_m - \Delta L)}{\mu W \cdot Cox' \cdot (V_{gs} - V_{t})}$$

so measure $R_m$ for different channel length transistors and plot $R_m$ vs $L_m$

where $R_m$ = intersect find value for $\Delta L$ and $R_{ds}$

Then $Leff$ can be calculated for each different length transistor from $Leff = L_m - \Delta L$
### LAMBDA VERSUS CHANNEL LENGTH

<table>
<thead>
<tr>
<th>µA UNIT</th>
<th>SLOPE</th>
<th>IDSAT</th>
<th>W</th>
<th>L</th>
<th>LAMBDA PMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>205</td>
<td>4.9</td>
<td>6.8</td>
<td>32</td>
<td>2</td>
<td>0.144118</td>
</tr>
<tr>
<td>71</td>
<td>2</td>
<td>7.1</td>
<td>32</td>
<td>4</td>
<td>0.056338</td>
</tr>
<tr>
<td>56</td>
<td>1.8</td>
<td>7.3</td>
<td>32</td>
<td>6</td>
<td>0.049315</td>
</tr>
<tr>
<td>34</td>
<td>1.2</td>
<td>7.5</td>
<td>32</td>
<td>8</td>
<td>0.032</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>7</td>
<td>32</td>
<td>16</td>
<td>0.057143</td>
</tr>
<tr>
<td>8.8</td>
<td>0.3</td>
<td>7.6</td>
<td>32</td>
<td>32</td>
<td>0.007895</td>
</tr>
<tr>
<td>415</td>
<td>4.3</td>
<td>6.5</td>
<td>32</td>
<td>2</td>
<td>0.132308</td>
</tr>
<tr>
<td>137</td>
<td>0.95</td>
<td>7.1</td>
<td>32</td>
<td>4</td>
<td>0.026761</td>
</tr>
<tr>
<td>91</td>
<td>0.4</td>
<td>7</td>
<td>32</td>
<td>6</td>
<td>0.011429</td>
</tr>
<tr>
<td>137</td>
<td>0.8</td>
<td>7.2</td>
<td>32</td>
<td>8</td>
<td>0.022222</td>
</tr>
<tr>
<td>27</td>
<td>0.2</td>
<td>7.2</td>
<td>32</td>
<td>16</td>
<td>0.005556</td>
</tr>
<tr>
<td>15</td>
<td>0.15</td>
<td>7.15</td>
<td>32</td>
<td>32</td>
<td>0.004196</td>
</tr>
</tbody>
</table>

Some MOSFET models use lambda but you would need a different lambda for each different length transistor. More advanced models use an equation to find a lambda as a function of the length $L_{eff}$. 

![Graph showing lambda versus channel length](image)
**SHORT CHANNEL VT ROLL OFF**

As the channel length decreases the channel depletion region becomes smaller and the Vt needed to turn on the channel appears to decrease.

A similar effect occurs for increasing $V_{DS}$ which causes an increase in the drain space charge layer. Called drain induced barrier lowering or DIBL.
A Test Chip is used that includes nMOS and pMOS transistors of various lengths from 0.1 µm to 5.0 µm and the threshold voltage is plotted versus channel length. The threshold voltage needs to be high enough so that when the input is zero or +Vsupply the transistor current is many decades lower than when it is on. $V_t$ and sub-$V_t$ slope interact.
**NARROW GATE WIDTH EFFECTS**

Fringing field causes channel depletion region to extend beyond the gate in the width direction. Thus, additional gate charge is required causing an apparent increase in threshold voltage. In wide channel devices, this can be neglected, but as the channel becomes smaller, it becomes more important.

In NMOS devices, encroachment of the channel stop impurity atoms under the gate edges causes the edges to be heavier doped, requiring more charge on the gate to turn on the entire channel width. In PMOSFETs, the phosphorous pile up at the surface under the field region causes a similar apparent increase in doping at the edges of the channel width.
V\textsubscript{t} initially increases with decrease in channel length then decreases. This is caused by various effects that result in lateral dopant nonuniformity in the channel.

Example: Oxidation Enhanced Diffusion or enhanced diffusion due to implant damage causing the dopant concentration to be higher in the channel near the drain and source edges of the poly gate.
The subthreshold characteristics are important in VLSI circuits because when the transistors are off they should not carry much current since there are so many transistors. (typical value about 100 mV/decade). Thinner gate oxide makes subthreshold slope larger. Surface channel has larger slope than buried channel.
DRAIN INDUCED BARRIER LOWERING

DIBL = change in VG / change in VD at ID=1E-9 amps/µm or 1.6E-8 amps for this size transistor

\[ \approx \frac{1.1957 - 1.1463}{5 - 0.1} \]

\[ \approx 10 \text{mV/V} \]
As the voltage on the drain increases the space charge associated with the drain pn junction increases. Current flow through the transistor increases as the source and drain space charge layers approach each other. The first indication is an increase in the sub threshold current and a decrease in the subthreshold slope.
PUNCHTHROUGH

Long channel behavior

Short channel behavior

Punchthrough

Rochester Institute of Technology
Microelectronic Engineering
MEASURED $I_d-V_{ds}$ Family

Punchthrough
Punch through implant increases the well doping below the drain and source depth making the space charge layer smaller helping to reduce punch through.
PUNCHTHROUGH HALO IMPLANT

Boron Implant at High Angle

Source

Gate

Drain

P-type well
RETROGRADE WELL TO REDUCE PUNCHTHROUGH

Fig. 1. SIMS measurement of the channel implant for nMOS and pMOS (a), and the source-drain extension/halo profiles for nMOS (b) and pMOS (c).
Sketch the three space charge layers
The Channel Space Charge
The Drain Space Charge
The Source Space Charge

Look at Punchthrough

Punchthrough will occur at lower drain voltages in the device with deeper D/S
Electron and hole mobilities in silicon at 300 K as functions of the total dopant concentration (N). The values plotted are the results of the curve fitting measurements from several sources. The mobility curves can be generated using the equation below with the parameters shown:

\[ \mu(N) = \mu_{\text{min}} + \frac{(\mu_{\text{max}} - \mu_{\text{min}})}{1 + \left(\frac{N}{N_{\text{ref}}}\right)^\alpha} \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Arsenic</th>
<th>Phosphorous</th>
<th>Boron</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu_{\text{min}})</td>
<td>52.2</td>
<td>68.5</td>
<td>44.9</td>
</tr>
<tr>
<td>(\mu_{\text{max}})</td>
<td>1417</td>
<td>1414</td>
<td>470.5</td>
</tr>
<tr>
<td>(N_{\text{ref}})</td>
<td>9.68X10^{16}</td>
<td>9.20X10^{16}</td>
<td>2.23X10^{17}</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.680</td>
<td>0.711</td>
<td>0.719</td>
</tr>
</tbody>
</table>

From Muller and Kamins, 3rd Ed., pg 33
CURRENT DRIVE - MOBILITY

\[ I_D = \frac{\mu W C_{ox}' (V_g - V_t - V_d/2) V_d}{L} \]

Non Saturation Region

\[ I_{Dsat} = \frac{\mu W C_{ox}' (V_g - V_t)^2}{2L} \]

Saturation Region

Mobility decreases with increase in doping concentration
MOBILITY DEGRADATION

In a MOSFET the mobility is lower than the bulk mobility because of the scattering with the Si-SiO2 interface. The vertical electric field causes the carriers to keep bumping into the interface causing the mobility to degrade. The electric fields can be 1E5 or 1E6 V/cm and at that level the collisions with the interface reduce the mobility even more. The vertical electrical field is higher for heavier doped substrates and when Vt adjust implants are used.

Mobility (cm²/volt·sec)

Ex (V/cm)
MOBILITY DEGRADATION

Note: Id should follow green line in long channel devices.

short channel

long channel

Note: Id should follow green line in long channel devices.
Carriers in semiconductors typically move in response to an applied electric field. The carrier velocity is proportional to the applied electric field. The proportionality constant is the mobility.

\[ \text{Velocity} = \mu \times \text{electric field} = \mu E \]

At very high electric fields this relationship ceases to be accurate. The carrier velocity stops increasing (or we say saturates). In a one micrometer channel length device with one volt across it the electric field is 1E4 V/cm.
VELOCITY SATURATION

Short channel

long channel

Note: \( I_d \) should increase with \((V_{gs} - V_t)^2\) in long channel devices
LOW DOPED DRAIN REDUCES LATERAL FIELD

- Low Doped Drain
- Source
- Gate
- Silicide
- Drain
- Side wall Spacer
- Field Oxide
- Stop
- P-type Punch Through Implant
- P-type well
- P-type well

Rochester Institute of Technology
Microelectronic Engineering
**NMOS WITH N⁺ POLY GATE**

- Vt Is Typically Negative Or If Positive Near Zero
- Vt Adjust Implant Is Boron In A P-type Substrate Making The Nmos Transistor A Surface Channel Device

![Graph showing Boron Vt Implant in a p-type wafer](image)

- $N_A (\text{cm}^{-3})$
- 1E16
- Depth into Wafer, $\mu$m
PMOS WITH N+ POLY GATE

- $V_t$ cannot be positive because all the contributors to the $V_t$ are negative. Even making $Q_{ss}=0$ and $N_d = 0$ does not make $V_t$ positive.
- $V_t$ is typically more negative than desired like -2 Volts.
- $V_t$ adjust implant is boron in an N-type substrate making the PMOS transistor a buried channel device (charge carriers move between drain and source at some distance away from the gate oxide/silicon interface).
**PMOS WITH N+ POLY GATE**

- **N (cm⁻³)**
  - 1E16

- **Depth into Wafer, µm**
  - 0.0
  - 0.2
  - 0.4
  - 0.6

- **Boron Vt Implant**
- **Phosphorous n-type wafer**
PMOS WITH P+ POLY GATE

- Changes Work Function Of The Metal
- Thus Metal-semiconductor Workfunction Difference Becomes About +1 Volt Rather Than ~0 Volts.
- This Makes Vt More Positive Than Desired So An Ion Implant Of N-type Impurity Is Needed Making The Device A Surface Channel Device Rather Than A Buried Channel Device.
**PMOS WITH P+ POLY GATE**

**Phosphorous Vt Implant**

N_D (cm⁻³)  
1E16  

Depth into Wafer, μm  
0.0 0.2 0.4 0.6  

Phosphorous n-type wafer
**SURFACE CHANNEL VS BURIED CHANNEL**

- Surface Channel Devices Exhibit Higher Subthreshold Slope
- Surface Channel Devices Are Less Sensitive To Punch Through
- Surface Channel Devices Have Less Severe Threshold Voltage Rolloff
- Surface Channel Devices Have Higher Transconductance
- Surface Channel Devices Have About 15% Lower Carrier Mobility
### SCALING OF MICROCHIPS

<table>
<thead>
<tr>
<th>Micron, Boise ID</th>
<th>L min</th>
<th>Chip Area</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>16 Meg DRAM</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.5 µm</td>
<td>140.1 mm²</td>
</tr>
<tr>
<td>1993</td>
<td>0.43</td>
<td>96.2</td>
</tr>
<tr>
<td>1994</td>
<td>0.35</td>
<td>57.0</td>
</tr>
<tr>
<td><strong>single level metal</strong></td>
<td>1995</td>
<td>59.6</td>
</tr>
<tr>
<td>1996</td>
<td>0.35</td>
<td>43.6</td>
</tr>
<tr>
<td>1996</td>
<td>0.30</td>
<td>38.3</td>
</tr>
<tr>
<td>1996</td>
<td>0.25</td>
<td>30.6</td>
</tr>
<tr>
<td>1997</td>
<td>0.30</td>
<td>29.2</td>
</tr>
<tr>
<td><strong>64 Meg DRAM</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>0.35</td>
<td>191.0</td>
</tr>
<tr>
<td>1996</td>
<td>0.30</td>
<td>123.3</td>
</tr>
<tr>
<td>1997</td>
<td>0.25</td>
<td>93.2</td>
</tr>
</tbody>
</table>
Let the scaling factor $K$ be: $K = \frac{\text{SIZE OLD}}{\text{SIZE NEW}}$

Example: to go from $1.0 \, \mu m$ to $0.8 \, \mu m$

$$K = \frac{1.0}{0.8} = 1.25$$

To reduce the gate length we also need to reduce the width of the D/S space charge layers. This can done by increasing the substrate doping. Now that the substrate doping is increased the MOSFET $V_t$ is harder to turn on; this can be corrected by decreasing the oxide thickness. Scaling a device in such a way as to keep the internal electric fields constant is called constant-field scaling.
CONSTANT FIELD SCALING

Quantity in Scaled Device = old Quantity times Scaling Factor

Dimensions (L’, W’, Xox’, Xj’) \( \frac{1}{K} \)
Area \( \frac{1}{K^2} \)
Packing Density \( K^2 \)
Doping Concentrations \( K \)
Bias Voltages and Vt \( \frac{1}{K} \)
Bias Currents \( \frac{1}{K} \)
Power dissipation \( \frac{1}{K^2} \)
Capacitance \( \frac{1}{K^2} \)
Electric Field Intensity \( 1 \)
Body Effect Coefficient \( \frac{1}{K^{0.5}} \)
Transistor Transit Time \( \frac{1}{K} \)
Transistor Power Delay Product \( \frac{1}{K^3} \)
### OTHER SCALING RULES

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Constant Field</th>
<th>Constant Voltage</th>
<th>Quasi-Constant Voltage</th>
<th>Generalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>W, L</td>
<td>1/K</td>
<td>1/K</td>
<td>1/K</td>
<td>1/K</td>
</tr>
<tr>
<td>Xox</td>
<td>1/K</td>
<td>1/β</td>
<td>1/K</td>
<td>1/K</td>
</tr>
<tr>
<td>N</td>
<td>K</td>
<td>K</td>
<td>K</td>
<td>K²/β</td>
</tr>
<tr>
<td>V, Vt</td>
<td>1/K</td>
<td>1</td>
<td>1/β</td>
<td>1/β</td>
</tr>
</tbody>
</table>

1 < β < K
SCALING EXAMPLES

Example: 5 Volt, L=1.0 μm NMOS, Na = 5E16, Xox=250 Å Scale to 0.8 μm NMOS. Constant Field Scaling

K = 1.0/0.8 = 1.25

Xox = 250/1.25 = 200 Å

N = 5E16 (1.25) = 2.5E17 cm⁻³

Vsupply = 5 Volts / 1.25 = 4 Volts and Vt = 1/1.25 = 0.8 Volts
The gate should be as thin as possible to reduce the short channel effects. In addition there is a limit imposed by considerations that affect the long term reliability of the gate oxide. This requirement imposes a maximum allowed electric field in the oxide under the long term normal operating conditions. This limit is chosen as 80% of the oxide field value at the on-set of Fowler-Nordheim (F-N) tunneling through the oxide. Since the latter is 5 MV/cm, a 4 MV/cm oxide field is considered as the maximum allowed for long term, reliable operation. For example:

For 2.5 volt operation, $X_{ox}$ is set at: $X_{ox} = \frac{V_{dd}}{E_{max}}$

$$= \frac{2.5 \text{ V}}{4 \text{ MV/cm}} = 65 \text{Å}$$
Ti Salicide will reduce the sheet resistance of the poly and the drain and source regions. Salicide is an acronym for Self Aligned Silicide and Silicide is a material that is a combination of silicon and metal such as Ti, W or Co. These materials are formed by depositing a thin film of the metal on the wafer and then heating to form a Silicide. The Silicide forms only where the metal is in contact with the Silicon or poly. Etchants can remove the metal and leave the Silicide thus the term Self Aligned Silicide or SALICIDE.
RIT’s FIRST SUB MICRON TRANSISTOR

Mark Klare 7/22/94 Electron beam direct write on wafer, n-well process 5E12 dose, P+ Poly Gate PMOS, shallow BF2 D/S implant, no Vt adjust implant.

- L=0.75 µm
- Xox=300 Å
- D/S Xj = 0.25 µm
- P+ poly
- Nd well ~3E16

Vt = -0.15
Sub Vt Slope=130 mV/dec

![Graph showing Ids (mA) vs Vds Volts]
RIT NMOS Transistor with Leffective = 0.4 µm

\[ I_D \] vs \[ V_D \] for NMOS Transistor

- \( V_G = 3.5 \text{V} \)
- \( V_G = 2.92 \text{V} \)
- \( V_G = 2.33 \text{V} \)
- \( V_G = 1.75 \text{V} \)
- \( V_G = 1.17 \text{V} \)
- \( V_G = 0.58 \text{V} \)

- \( L_{\text{mask drawn}} = 0.6 \text{µm} \)
- \( L_{\text{effective}} = 0.4 \text{µm} \)

*This is RIT’s first sub-0.5 µm Transistor*

Mike Aquilino May 2004
**RIT NMOS Transistor with $L_{eff} = 0.4 \ \mu m$**

- $L_{eff} = 0.4 \ \mu m$
- $I_D @ (V_G=V_D=3.5V) = 140 \ \mu A/\mu m$
- $V_t = 0.75V$
- $SS = 103 \ \text{mV/decade}$
- $\log (I_{on}/I_{off}) = 7.5 \ \text{Orders of Magnitude}$

---

Mike Aquilino May 2004
RIT NMOS Transistor with $L_{\text{effective}} = 0.4 \, \mu\text{m}$

- $0.5 \, \mu\text{m}$ exhibits well controlled short channel effects
- $0.4 \, \mu\text{m}$ device can be used depending on off-state current requirements
- $33\%$ Increase in Drive Current compared to $0.5 \, \mu\text{m}$ device

Mike Aquilino May 2004
**SUB 0.25μm NMOSFET**

- \( L_{mask} = 0.5 \ \mu m \)
- \( L_{poly} = 0.25 \ \mu m \)
- \( L_{effective} = 0.2 \ \mu m \)

Mike Aquilino  
May 2006

Figure 28: ID-VD for 0.25 μm NMOS Transistor

- ID = 177 μA/μm @ \( V_G = V_D = 2.5 \ \text{V} \)
- \( V_T = 1.0 \ \text{V} \)

*This is RIT’s Smallest NMOS Transistor*
**SUB 0.25\(\mu\)m PMOSFET**

- \(L_{\text{mask}} = 0.6 \ \mu\text{m}\)
- \(L_{\text{poly}} = 0.25 \ \mu\text{m}\)
- \(L_{\text{effective}} = 0.2 \ \mu\text{m}\)

<table>
<thead>
<tr>
<th>ID-VD 0.25 (\mu)m PMOS Transistor</th>
</tr>
</thead>
<tbody>
<tr>
<td>VD (volts)</td>
</tr>
<tr>
<td>ABS(ID) ((\mu)A/(\mu)m)</td>
</tr>
</tbody>
</table>

Figure 31: ID-VD for 0.25 \(\mu\)m PMOS Transistor

- \(|ID| = 131 \ \mu\text{A}/\mu\text{m} @ VG=VD=-2.5 \ \text{V}\)
- \(V_T = -0.75 \ \text{V}\)

*This is RIT’s Smallest PMOS Transistor*

Mike Aquilino
May 2006
0.25μm Leff NMOSFET

- $I_{off} = 13 \text{ pA}/\mu\text{m} @ VD=0.1 \text{ V}$ (with drain diode leakage removed)
- $I_{off} = 11 \text{ nA}/\mu\text{m} @ VD=2.5 \text{ V}$ (with drain diode leakage removed)
- $\log(I_{on}/I_{off}) = 4.2 \text{ decades}$
- $SS = 119 \text{ mV/decade} @ VD=0.1 \text{ V}$

0.25μm Leff PMOSFET

- $I_{off} = -20 \text{ fA}/\mu\text{m} @ VD=-0.1 \text{ V}$
- $I_{off} = -4.9 \text{ pA}/\mu\text{m} @ VD=-2.5 \text{ V}$
- $\log(I_{on}/I_{off}) = 7.4 \text{ decades}$
- $SS = 75 \text{ mV/decade} @ VD=-0.1 \text{ V}$
- $SS = 85 \text{ mV/decade} @ VD=-2.5 \text{ V}$
- $DIBL = 8.3 \text{ mV/V} @ ID=-1 \text{ nA}/\mu\text{m}$

ID-VG for 0.25 μm PMOS Transistor
REFERENCES

HOMEWORK – SHORT CHANNEL MOSFETs

1. In short channel devices the threshold voltage becomes less than expected for long channel devices. Why.
2. Explain reverse short channel effect.
3. What is the effect of narrow channel width on transistor device characteristics.
4. What is the purpose of low doped drain structures?
5. How does mobility degradation and velocity saturation effect transistor device characteristics?
6. Why is P+ doped poly used for PMOS transistors.
7. What is the difference between mask channel length and effective channel length.
8. What is punchthrough? What processing changes can be made to compensate for punchthrough?
9. When scaling from 2 um to 1.5 um give new values for: device dimensions W,L,Xox, doping concentration, bias voltages, bias currents, power dissipation, transit time.
10. What is SALICIDE process. Why is it used?