Digital Integrated Circuits
(83-313)

Lecture 3:
MOSFET Modeling

27 April 2020
Lecture Content

- Basic MOS Models
- Advanced MOS Models
- Threshold Voltage Revisited
- Simulating Variation
- Leakage in NanoScaled Transistors
Basic MOS Models

1. Basic MOS Models
2. Advanced MOS Models
3. VT Revisited
4. Simulating Variation
5. Leakage
TCAD vs. Compact Models

• Technology CAD (TCAD) is a simulation environment for accurately simulating device behavior:
  • Provide a process “recipe” and device layout
  • Produce IV or CV curves through device simulator
  • Used to predict device and process physics
  • Takes 1hr-1day per IV curve and 100s MB RAM per transistor

• Compact models (a.k.a. SPICE models or ECAD) are simple models used for circuit simulation
  • Provide a set of equations that SPICE uses to calculate IV or CV curves
  • Should take <100us per IV curve and a few KB per transistor
  • Usually extracted empirically from measurements

TCAD is too slow and memory hungry to be used for circuit simulation!

SPICE uses compact models for calculating device behavior
Switch Model

- The most simple MOSFET model is the Switch Model.
The Piece-Wise Linear Model

- As we know, when the channel **pinches off**, the current **saturates**.
- This can be depicted with the simple **Piece-Wise Linear Switch Model**
Adding Channel Length Modulation

- Channel Length Modulation modeled as a finite output resistance, causes a saturation current dependence on $V_{DS}$.

\[ I_{DSAT} = \frac{1}{\lambda} I_{DSAT} \]

\[ V_{DS} \]

\[ I_{DS} = \frac{1}{R_{on}} \]

\[ I_{DSAT} - \frac{1}{\lambda} I_{DSAT} \]

\[ V_{GT} \]

\[ V_{DS} \]

The Switch Model

The Piece-Wise Linear Model

The Square Law Model

The V-Sat Model

The Unified Model

The VT* Model

The $\alpha$-Power Law

BSIM
Square Law (Shockley) Model

To get a more accurate model, we already are familiar with the Shockley or Square Law Model.

Current is just charge times velocity, so at any point, \( x \), along the channel:

\[
I_D(x) = -\nu(x)Q(x)W \, dx
\]

We found that charge can be approximated as:

\[
Q(x) = -C_{ox} \left[ V_{GS} - V_{CS}(x) - V_T \right]
\]

And the velocity is the mobility times the electrical field:

\[
\nu(x) = -\mu E(x) = \mu_n \frac{dV}{dx}
\]
Square Law (Shockley) Model

• So we get:

\[ I_D dx = \mu_n C_{ox} W (V_{GS} - V - V_T) dV \]

• And integrating from source to drain, we get

\[ I_{DS} = \int_0^L I_D dx = \int_0^{V_{DS}} \mu_n C_{ox} W (V_{GS} - V - V_T) dV = \mu_n C_{ox} \frac{W}{L} V_{DS} \left( V_{GS} - V_T - \frac{1}{2} V_{DS} \right) \]

• At pinch-off \((V_{DS}=V_{GS}-V_T)\), the voltage over the channel is constant, so we get:

\[ I_{DSAT} = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2 \]

• This is where the “Square-Law” name comes from.
Square Law (Shockley) Model

- Replacing $V_{DS}$ with $V_{DSeff} = \min(V_{GS} - V_T, V_{DS})$ we get:

$$I_{DS} = \mu_n C_{ox} \frac{W}{L} \left[ (V_{GS} - V_T) V_{DSeff} - \frac{V_{DSeff}^2}{2} \right] \left( 1 + \lambda V_{DS} \right)$$

The Switch Model
The Piece-Wise Linear Model
The Square Law Model
The V-Sat Model
The Unified Model
The VT* Model
The $\alpha$-Power Law
BSIM

$V_{DS} = V_{DD}$
$V_{DS} = V_{DD}/2$

$I_{DS}$ vs. $V_{GS}$
$I_{DS}$ vs. $V_{DS}$
The Velocity Saturation Model

- However, when looking at a short channel device, we see a linear dependence on $V_{GS}$.
- This can be attributed to Velocity Saturation.

$v_{sat} \approx 10^5 \text{m/s}$

$\xi = V/L$
The Velocity Saturation Model

• A good approximation of the mobility curve is:

\[ v = \begin{cases} \frac{\mu \xi}{1 + \xi} & \xi < \xi_{\text{crit}} \\ \xi_{\text{sat}} & \xi > \xi_{\text{crit}} \end{cases} \]

• For continuity:

\[ \xi_{\text{crit}} = \frac{2v_{\text{sat}}}{\mu} \]

• After integration, we get:

\[ I_{DS} = \frac{\mu_n C_{ox}}{1 + V_{DS}/\xi_{\text{crit}}L} \frac{W}{L} \left[ (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \]
The Velocity Saturation Model

• This is hard to use, but we can reach an important conclusion.

• We found that:

\[ I_{DS} = \frac{\mu_n C_{ox}}{1 + \frac{V_{DS}}{\xi_{crit} L}} W \left[ (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \]

• And we know that for a velocity saturated device:

\[ I_{DS} = W C_{ox} (V_{GS} - V_{DSAT} - V_T) v_{sat} \]

• Equating, we get:

\[ V_{DSAT} = \frac{(V_{GS} - V_T) \xi_{crit} L}{(V_{GS} - V_T) + \xi_{crit} L} \]

\[ V_{DSAT} (\xi_{crit} L \gg V_{GT}) = V_{GS} - V_T \Rightarrow \text{pinch off} \]

\[ V_{DSAT} (\xi_{crit} L \ll V_{GT}) = \xi_{crit} L \Rightarrow \text{vel sat} \]
A few simple estimations will make the V-Sat model more user-friendly:

- The mobility is piecewise linear, saturating at $\zeta > \zeta_{\text{crit}}/2$
- $V_{\text{DSAT}}$ is piecewise linear, saturating at $V_{\text{DSAT}} = \zeta_{\text{crit}} L/2$, when $V_{\text{GT}} > \zeta_{\text{crit}} L/2$
The Unified Model for Hand Analysis

• This brings us to the Unified Model:

\[ I_{DS} = \mu_n C_{ox} \frac{W}{L} \left[ (V_{GS} - V_T) V_{DSeff} - \frac{V_{DSeff}^2}{2} \right] (1 + \lambda V_{DS}) \]

\[ V_{DSeff} = \min(V_{GS} - V_T, V_{DS}, V_{DSAT}) \]

\[ V_{DSAT} = \frac{\xi_{crit} L}{2} \quad \xi_{crit} = \frac{2V_{sat}}{\mu} \]
Advanced MOS Models
VT* Model

- Sometimes we want to use a really simple model.
- We can assume that if the transistor is on, it’s velocity saturated.

\[
I_{DS} = k_n \left[ (V_{GS} - V_T) V_{DSAT} - \frac{V_{DSAT}^2}{2} \right] = \\
= k_n \left( V_{GS} - \left( V_T + \frac{V_{DSAT}}{2} \right) \right) V_{DSAT}
\]

\[
I_{DS} = \begin{cases} 
0 & V_{GS} < V_T^* \\
 k_n \left( V_{GS} - V_T^* \right) V_{DSAT} & V_{GS} > V_T^* 
\end{cases}
\]

\[
V_T^* = V_T + \frac{V_{DSAT}}{2}
\]
The Alpha Power Law Model

Sakurai found that by changing the exponent of the square law, a better fit can be found with simple calculations.

\[ I_{DSAT}(\alpha) = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^\alpha \]
BSIM and Newer Models

- BSIM (Berkeley Short-channel IGFET Model) is the primary compact model family used for SPICE simulation for the last three decades.
- These models use hundreds of parameters to achieve a good fit.
- BSIM4 is the main model for bulk CMOS
  - Takes into account most physical effects as well as many fitting parameters.
- The Compact Model Coalition (CMC) chooses, maintains and promotes new models
  - Additional models include EKV, PSP, and models for non-MOS devices.

Source: Yen Ki Lin, UC Berkeley
Threshold Voltage Revisited
Energy Band Diagrams

- To understand the threshold voltage and other secondary effects of the MOS device, we often use energy band diagrams.
- The first approach is looking in from the gate:
Energy Band Diagrams

• The second approach is looking from the source to the drain.
Threshold Voltage - Basic Theory

- The basic definition of threshold voltage is the gate voltage \((V_G)\) required to invert the channel

\[
V_{T0} = \Phi_{MS} - 2\Phi_F - \frac{Q_{OX}}{C_{ox}} - \frac{Q_{dep}}{C_{ox}} - \frac{Q_I}{C_{ox}}
\]

\[
Q_{dep} = \sqrt{2qN_A\varepsilon_{si}} | -2\Phi_F |
\]

\[
\Phi_F = -\phi_T \ln \frac{N_A}{n_i} \quad \phi_T = \frac{kT}{q}
\]
Body Effect

- The appearance of a voltage difference between the source and body ($V_{SB}$) is known as “The Body Effect”
- This can be modeled by the additional charge that needs to be depleted.

\[
Q_{dep} = \sqrt{2qN_A \varepsilon_{si} \left( | -2\Phi_F + V_{SB} | \right)}
\]

\[
V_T = V_{T0} + \gamma \left( \sqrt{| -2\Phi_F + V_{SB} |} - \sqrt{| -2\Phi_F |} \right)
\]

\[
V_{T0} \equiv \Phi_{MS} - 2\Phi_F - \frac{Q_{ox}}{C_{ox}} - \frac{Q_{dep0}}{C_{ox}} \quad \gamma = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C_{ox}}
\]
Modern Body Effect

- A different approach is to look at the capacitive voltage divider between the gate and body \((C_{GB})\)

\[
Q_{inv} = -C_{oxe} (V_{GS} - V_{CS} - V_{T0}) + C_{dep} (V_{SB} + V_{CS})
\]

\[
= -C_{oxe} (V_{GS} - nV_{CS} - V_{T0})
\]

\[
n \triangleq 1 + \frac{C_{dep}}{C_{oxe}} = 1 + \frac{3T_{oxe}}{W_{d max}}
\]

\[
C_{dep} = \frac{\varepsilon_s}{W_{d max}}
\]
Modern Body Effect

• This can be shown to redefine $V_T$ as:

$$V_T(V_{SB}) = V_{T0} + \frac{C_{dep}}{C_{oxe}} V_{SB}$$

• In modern technologies, $\frac{C_{dep}}{C_{oxe}}$ is a constant, so $V_T$ is **linearly dependent** on $V_{SB}$!
Poly Depletion and Channel Depth

The threshold voltage is affected by two additional factors that we have disregarded until now:

• Polysilicon Depletion
  • Since polysilicon is, itself, a semiconductor, the depletion layer into the poly effectively increases the oxide thickness.

• Channel Depth
  • Since the channel is not a 2-dimensional line along the surface, the oxide thickness is essentially increased.

\[ n \Delta = 1 + \frac{C_{\text{dep}}}{C_{\text{oxe}}} = 1 + \frac{3T_{\text{oxe}}}{W_{d \text{ max}}} \]
Hot Carrier Effects

- Electrons can get so fast that they can tunnel into the gate oxide and increase the threshold voltage.

\[ V_{T0} = \Phi_{MS} - 2\Phi_F - \frac{Q_{OX}}{C_{ox}} - \frac{Q_{dep}}{C_{ox}} - \frac{Q_I}{C_{ox}} \]

- This is a reliability issue as it happens over time.
$V_T$ Roll Off (Short Channel Effect)

- As channel length is reduced, effective channel length is reduced by depletion regions.
- A trapezoid is created under the gate, dividing the channel into the region controlled by the gates and by the drain.
- In essence, $V_T$ is reduced.

$V_T$ Roll-Off

$V_T$ vs Lg (um) for 65nm technology:
- $V_{ds} = 50mV$
- $V_{ds} = 1.0V$

Source/Channel barrier

Long channel

Short channel

K. Goto et al., IEDM 2003
DIBL (Drain Induced Barrier Lowering)

• In short channels, the barrier of the channel is essentially lowered, as the drain causes the energy band to drop closer to the source.

• This is exponentially dependent on $V_{DS}$.

\[ V_T = V_{T,\text{long}} - \left( V_{DS} + 0.4 \right) \cdot \frac{C_d}{C_{\text{oxe}}} \]
Roll Off / DIBL combined

Source: Chris Kim, U. Minn
Reverse Short Channel Effect (RSCE)

- \( V_T \) actually \textit{increases} at channel lengths a bit higher than minimum...
How to Measure VT

- There are various ways to measure $V_T$.
- One classic way takes a small $V_{DS}$ and sweeps $V_{GS}$.

\[ I_d = k \left[ (V_{GS} - V_T)V_{DS} - 0.5V_{DS} \right] \propto V_{GS} - V_T \]

- So we can find the $V_{GS}$ at which the linear part crosses $I_{ds} = 0$. 

![Diagram showing measurement of $V_T$](image-url)
How to Measure VT

• One of the more common ways is to find the \( V_{GS} \) at which \( I_{DS} = 100 \text{ nA} \times \frac{W}{L} \).

• For \( V_{T,\text{lin}} \), set a low \( V_{DS} \) (\( V_{DS} = 50 \text{ mV} \))

• For \( V_{T,\text{sat}} \), set a high \( V_{DS} \) (\( V_{DS} = V_{DD} \))
Note about Simulation

• The first step of all Spice (Spectre) simulations is a DC Operating Point calculation.
Simulation tip: OP and MP in Spectre

• So we saw that the threshold voltage is dependent on the operating point.
• How do you know what the the $V_T$ of a transistor is in a given simulation?
• To find $V_{T0}$, use the “MP” option.

• To find $V_{T,\text{lin}}$, $V_{T,\text{sat}}$, $V_{T,\text{gm}}$, use the “OP” option.

• How would you go about plotting roll-off (SCE) and RSCE?
The Computer Hall of Fame

• We generally consider the ENIAC to be the first computer, but the official first fully electronic computer was the ACE

  • Atansoff-Berry Computer
    Conceived in 1937, operational in 1942

  • Built at Iowa State University by Prof. John Atansoff and his student Clifford Berry.
  • Not programmable nor Turing-complete, but included binary arithmetic and electronic switching elements.
  • A patent dispute over the first electronic computer was settled in 1973, when the patent of the ENIAC was invalidated.
Simulating Variation
Process Corners and Monte Carlo Simulation
Reminder: Impact of Process Variations

How do we take this into consideration during simulation?

Source: Rabaey, et. al.
Remove the Rust: Probability Basics

• Properties of Random Variables
  • The probability distribution function (PDF) \( f(x) \) specifies the probability that a value of a continuous random variable \( X \) falls in a particular interval:
  \[
P[a < X \leq b] = \int_{a}^{b} f(x) \, dx
  \]
  • The cumulative distribution function (CDF) \( F(x) \) specifies the probability that \( X \) is less than some value \( x \):
  \[
  F(x) = P[X < x] = \int_{-\infty}^{x} f(u) \, du
  \]
  \[
  f(x) = \frac{d}{dx} F(x)
  \]
  • The mean (\( \mu \)) and variance (\( \sigma^2 \)) are defined as:
  \[
  \mu(X) = \bar{X} = E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx
  \]
  \[
  \sigma^2(X) = E\left[ (x - \bar{X})^2 \right] = \int_{-\infty}^{\infty} (x - \bar{X})^2 \, f(x) \, dx
  \]
• Normal Random Variables

A normal (Gaussian) random variable, shifted to have a zero mean ($\mu=0$) and a normalized standard variation ($\sigma^2=1$) has:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

$$F(x) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x}{\sqrt{2}} \right) \right]$$

- $3\sigma = 99.8\%$
- $6\sigma = 1 \text{ in a billion}$
Global Variation Modeling: Process Corners

• Global Variation assumption:
  • If a certain process step is skewed, the entire chip is affected equivalently.
  • We will define “corner cases” of fabrication, i.e., $3\sigma$ from the mean.
  • We also assume the voltage and temperature are globally affected.
• Devices are modeled for fast, slow and nominal corners.
  • Changes in $V_T$, $W$, $L$, $t_{ox}$
• Devices are tested at various temperatures
  • Temperature affects mobility and $V_T$.
  • Typically 0°C – 85°C or -40°C – 125°C
• Devices are tested at various supply voltages
  • Higher voltages cause increased currents
  • Typically $\pm10\%V_{DD}$
Process Corners

• What are the PVT (Process, Voltage, Temperature) settings for each simulation corner?

<table>
<thead>
<tr>
<th>Corner</th>
<th>$V_T$</th>
<th>$L_{eff}$</th>
<th>$t_{ox}$</th>
<th>$V_{DD}$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Typical</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Slow</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
</tr>
</tbody>
</table>
What about Local Variation?

• Often there are *too many* parameters to think about and setting a specific corner case is insufficient.

• For example: Pelgrom’s Law
  
  • $V_T$ variance is inversely proportional to transistor area

\[
\sigma(V_T) = \frac{K}{\sqrt{W \cdot L}}
\]

How do we deal with this?
Monte Carlo Simulation

- The basic approach is to “roll the dice” for each parameter and run a simulation.
  - These are called *Monte Carlo* Statistical Simulations.
- The result is a distribution plot of design constraints, e.g., delay or noise margin
- Both Global and Local Variations can be taken into consideration.

Source: Stanford, EE380
Simulation questions

• How do we plot threshold voltage variation?
Leakage in NanoScaled Transistors
Leakage in Nanoscaled Transistors

- Transistors that are supposed to be off actually leak!
Main Types of Leakage

**Polysilicon Gate**

- **Source**
- **Oxide**
- **Drain**

**P-substrate**

- **n+**

**Diode**

**Punchthrough**

**Gate**

**Subthreshold**

**GIDL**

**DIBL**

\[
I_{sub} \propto e^{V_{GS}-V_T}
\]

\[
I_{gate} \propto V_{GC}, \frac{1}{t_{ox}}
\]

\[
I_{GIDL} \propto V_{DG}
\]

\[
DIBL \propto e^{V_{DS}}
\]
Subthreshold Leakage

• When $V_{GS} < V_T$, there is still a finite carrier concentration at the surface:

$\frac{n_s}{e^{\phi_s/\phi_T}} \implies I_{sub} \propto e^{V_{GS}-V_T/n\phi_T}$

• As we saw with the body effect, due to bulk to channel capacitance, the surface voltage isn’t only controlled by the gate:

$n \equiv 1 + \frac{C_{dep}}{C_{oxe}}$

$d\phi_s = \frac{C_{oxe}}{C_{oxe} + C_{dep}} \frac{1}{n}$
Subthreshold Leakage

- Let’s make things easier:
  - Remember that:
  \[ I_{DS} \text{(sub)} = Const \frac{W}{L} \cdot e^{\frac{V_{GS} - V_T}{n\Phi_T}} \]
  - And we now defined the threshold voltage according to current:
  \[ I_{DS} (V_{GS} = V_T) = 100 \text{nA} \cdot \frac{W}{L} \]
  - So the boundary condition requires:
  \[ I_{DS} \text{(sub)} \bigg|_{V_{GS} = V_T} = Const \frac{W}{L} \cdot e^{\frac{0}{n\Phi_T}} = Const \frac{W}{L} = 100 \text{nA} \frac{W}{L} \]
  - And we can now calculate subthreshold current as:
  \[ I_{sub} [\text{nA}] = 100 \frac{W}{L} e^{\frac{V_{GS} - V_T}{n\Phi_T}} \]
Subthreshold Leakage

• An even easier way is to look at the plot of $\log(I_{DS})=f(V_{GS})$:
  • The slope of this curve is called the “Subthreshold Slope”
  • The inverse of this slope is known as the “Subthreshold Swing” ($S$):

$$S \equiv \ln(10) \frac{kT}{q} \left(1 + \frac{C_{dep}}{C_{ox}}\right) = 2.3 \cdot n \cdot \phi_T$$

$$I_{sub} [\text{nA}] = 100 \frac{W}{L} e^{\frac{V_{GS} - V_T}{n\phi_T}} = 100 \frac{W}{L} 10^{\frac{V_{GS} - V_T}{S}}$$

• And $I_{off}$, which is defined as the current when $V_{GS}=0$ is:

$$I_{off} [\text{nA}] = 100 \frac{W}{L} 10^{-\frac{V_T}{S}}$$
Subthreshold Leakage

- So subthreshold leakage is:
  - Exponentially dependent on $V_{GS}$.
  - Exponentially dependent on $V_T$.

- $S$ is the subthreshold swing coefficient.
  - Optimally, $S_{\text{opt}} = 60$ mV/dec
  - Realistically $S \approx 100$ mV/dec

$$S \equiv n\phi_T \ln 10 \geq 0.06 \quad n \equiv 1 + \frac{C_{\text{dep}}}{C_{\text{oxe}}}$$
Subthreshold Leakage

Example:

• We want to design a transistor with:

\[ \frac{I_{on}}{I_{off}} \geq 10^4 \]

\[ S = 60 \text{ mV/dec} \]

• What is the minimum \( V_T \)?
Impact of DIBL

- **DIBL** causes an additional *exponential* increase in subthreshold leakage with $V_{DS}$.

\[
I_{sub} = I_0 e^{\left(\frac{V_{GS} - V_T}{n\phi_T}\right)} \times \left(1 - e^{\frac{-V_{DS}}{\phi_T}}\right) \times e^{\frac{\eta V_{DS}}{n\phi_T}}
\]

90nm technology. Gate length: 45nm

Source: Intel, T. Ghani et al., IEDM 2003
Subthreshold Dependence on Temperature

- This is rather complex, as mobility degrades with temperature and other device values (such as flatband voltage) are temperature dependent.
- Altogether, subthreshold leakage rises exponentially with temperature*.

\[ I_{\text{sub}} \propto e^{\frac{V_{GS} - V_T}{\eta \phi_T}} \]

\[ \phi_T = \frac{kT}{q} \]

* Without considering temperature inversion
Temperature Inversion

- Classic approach to temperature effect on delay:
  \[
  I \propto \mu \\
  \mu \propto \frac{1}{T} \\
  \rightarrow t_{pd} \propto T
  \]
  So speed decreases with temperature.

- BUT!
  - \( V_T \) decreases by as much as \(-3\, \text{mV/°C}\)
  - The point of temperature inversion is the voltage at which speed increases with temperature (~\( V_{\text{DD}} = 1\, \text{V} \))!

\[
V_T \frac{1}{T} \\
\rightarrow t_{pd} \propto \frac{1}{T}
\]
So speed actually increases with temperature!
Gate Leakage

- Two mechanisms:
  - Direct tunneling (dominant)
  - Fowler Nordheim tunneling

- Exponentially Dependent on:
  - Gate Voltage ($V_G$)
  - Oxide Thickness ($t_{ox}$).

- Non-dependent on temperature.
- Much stronger in nMOS than pMOS (higher barrier for holes)

- Minimum $t_{ox} = 1.2 \text{nm}!!!$

\[ I_{gate} = A E_{ox}^2 e^{-B/E_{ox}} \]
\[ E_{ox} = \frac{V_{DD} - V_T}{t_{ox}} \]
Gate Induced Drain Leakage

- **GIDL** current flows from the drain to the substrate.
- Caused by high electric field **under the gate/drain overlap**, causing **e-h pair creation**.
- Main phenomena is **Band-to-Band Tunneling**
Diode Leakage

- $J_S = 10$-$100 \text{pA/\mu m}^2$ @ $25^\circ \text{C}$ for $0.25 \mu \text{m}$ CMOS.
- $J_S$ doubles for every $9^\circ \text{C}$
  \[ I_{DL} = J_S \times A \]
- Much smaller than other leakages in deep sub-micron.
  - But a bigger factor in low subthreshold leakage processes, like FinFET.
Punchthrough

- As $V_{DS}$ grows, so does the drain depletion region, and the channel length decreases.
- In severe cases, the source and drain are connected causing non-controllable leakage current.
Leakage Summary

- **Subthreshold Leakage:**
  - $I_{DS} > 0$ when $V_{GS} < V_T$ due to weak inversion.
  - Grows with $V_{GS}$, $V_{DS}$, lower $V_T$

- **Gate Leakage:**
  - $I_G > 0$ due to direct tunneling through the oxide.
  - Grows with $V_{GB}$, $t_{ox}$

- **Gate Induced Drain Leakage (GIDL):**
  - $I_{DB} > 0$ due to high electric field in the GD overlap region ($V_{GD}$).

- **Reverse Biased Diode Leakage**
  - $I_{SB}$, $I_{DB}$ due to diffusion and thermal generation.

- **Punchthrough:**
  - $I_{DS}$ due to drain and source depletion layers touching.
Process Corners - Revisited

- Should we now redefine the PVT settings?

<table>
<thead>
<tr>
<th>Corner</th>
<th>$V_T$</th>
<th>$L_{eff}$</th>
<th>$t_{ox}$</th>
<th>$V_{DD}$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Typical</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Slow</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Max Leakage</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
</tr>
</tbody>
</table>

But what about Temperature Inversion?
Further Reading

• J. Rabaey, “Digital Integrated Circuits” 2003, Chapters 2.5, 3.3-3.5
• Weste, Harris “CMOS VLSI Design”, Chapter 7
• C. Hu, “Modern Semiconductor Devices for Integrated Circuits”, 2010, Chapters 4-7
• Tzividis, et al. “Operation and Modeling of MOS Transistor” Chapters 1-5
• E. Alon, Berkeley EE-141, Lecture 9 (Fall 2009)
• M. Alam, Purdue ECE-606 – lectures 32-38 (2009) nanohub.org
• A. B. Bhattacharyya “Compact MOSFET models for VLSI design”, 2009,
• Managing Process Variation in Intel’s 45nm CMOS Technology, Intel Technology Journal, 2008
• Berkeley “BSIM 4.6.4 User’s Manual”