Lecture 6:
Non-ideal (real) transistors

Mark McDermott
Electrical and Computer Engineering
The University of Texas at Austin

Agenda

UT ECE

- Transistor I-V Review
- Non-ideal Transistor Behavior
  - High Field Effects
    - Mobility Degradation
    - Velocity Saturation
  - Channel Length Modulation
  - Threshold Voltage Effects
    - Body Effect
    - Drain-Induced Barrier Lowering
    - Short Channel Effect
  - Leakage
    - Subthreshold Leakage
    - Gate Leakage
    - Junction Leakage
- Process and Environmental Variations

Ideal Transistor I-V

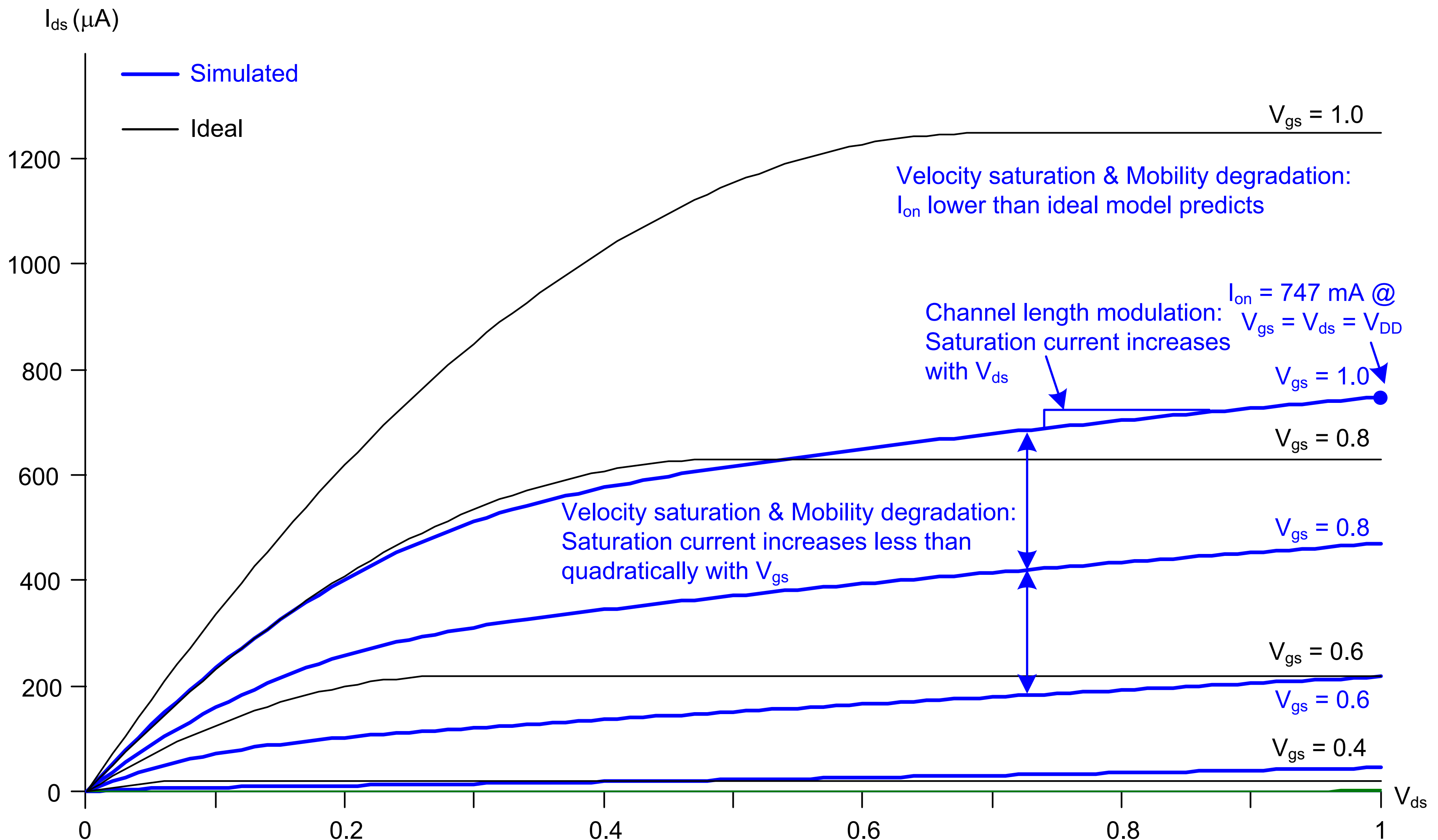
- *Shockley* 1st order transistor models

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ \beta\left(V_{gs} - V_t - \dfrac{V_{ds}}{2}\right)V_{ds} & V_{ds} < V_{dsat} & \text{linear} \\ \dfrac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} & \text{saturation} \end{cases}$$

Ideal vs. Simulated nMOS I-V Plot

UT ECE

65 nm IBM process, V_{DD} = 1.0 V

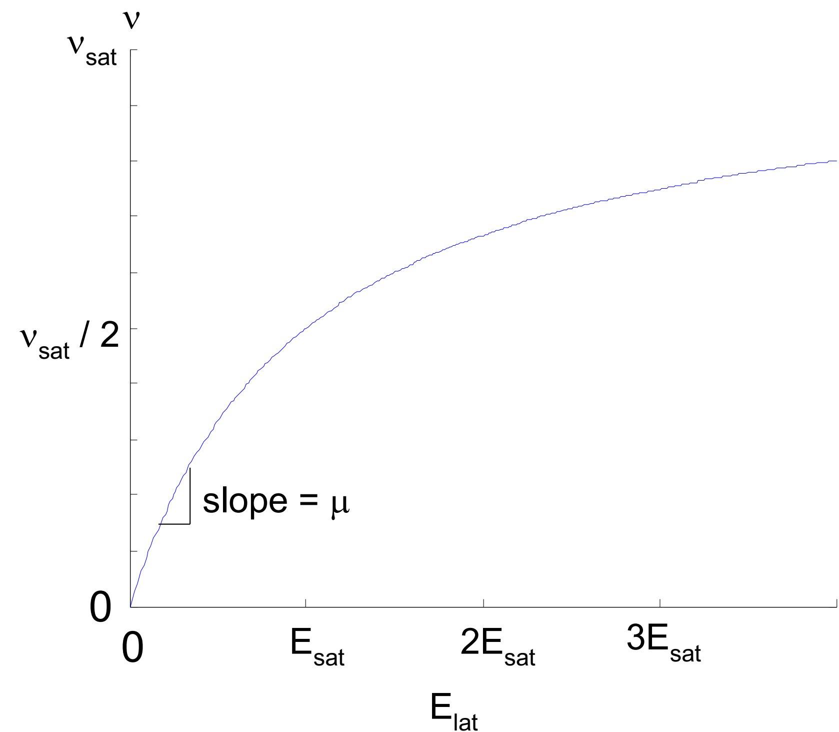


Velocity Saturation

- **We assumed carrier velocity / E-field**
  - $v = \mu E_{lat} = \mu V_{ds}/L$
- **At high fields, this ceases to be true**
  - Carriers scatter off atoms
  - Velocity reaches v_{sat}
    - Electrons: $6\text{-}10 \times 10^6$ cm/s
    - Holes: $4\text{-}8 \times 10^6$ cm/s

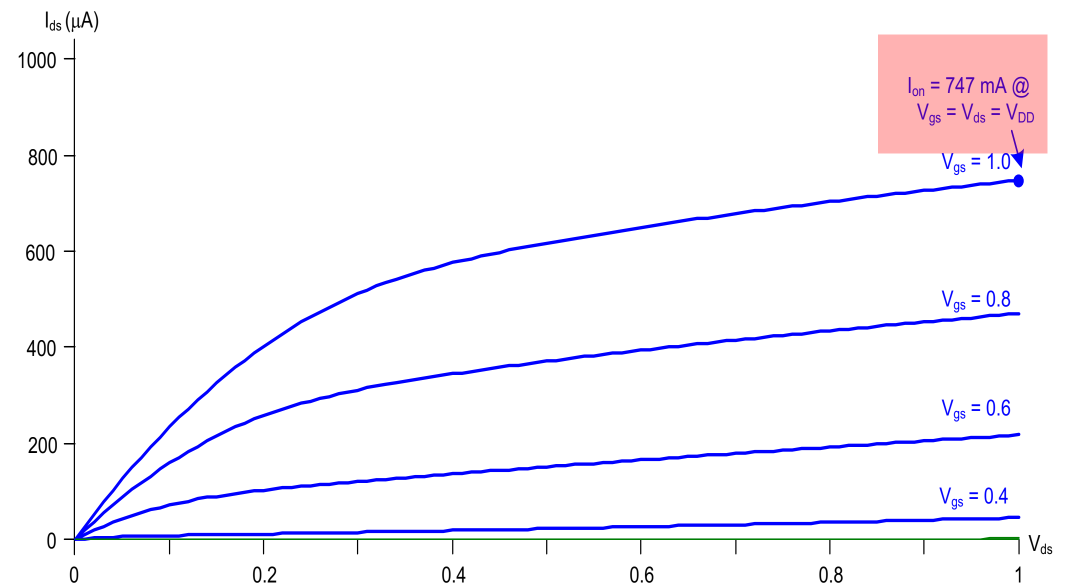
Better model is:

$$v = \frac{\mu E_{lat}}{1 + \frac{E_{lat}}{E_{sat}}} \Rightarrow v_{sat} = \mu E_{sat}$$

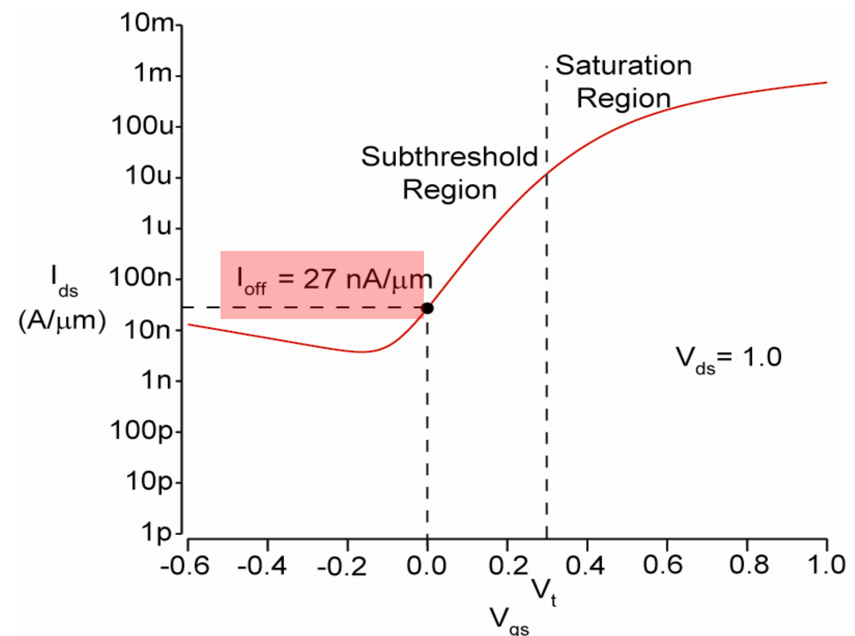


ON and OFF Current

- $I_{on} = I_{ds}$ @ $V_{gs} = V_{ds} = V_{DD}$
  - Saturation



- $I_{off} = I_{ds}$ @ $V_{gs} = 0$, $V_{ds} = V_{DD}$
  - Cutoff



Electric Fields Effects

- **Vertical electric field: E_{vert} = Vgs / tox**
  - Attracts carriers into channel
  - Long channel: $Q_{channel} \propto E_{vert}$

- **Lateral electric field: E_{lat} = Vds / L**
  - Accelerates carriers from drain to source
  - Long channel: $v = \mu E_{lat}$

Mobility Degradation

- High E_{vert} effectively reduces mobility
  - Collisions with oxide interface

$$\mu_{\text{eff}-n} = \frac{540 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \left(\frac{V_{gs} + V_t}{0.54 \frac{\text{V}}{\text{nm}} t_{\text{ox}}} \right)^{1.85}} \qquad \mu_{\text{eff}-p} = \frac{185 \frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \frac{\left| V_{gs} + 1.5 V_t \right|}{0.338 \frac{\text{V}}{\text{nm}} t_{\text{ox}}}}$$

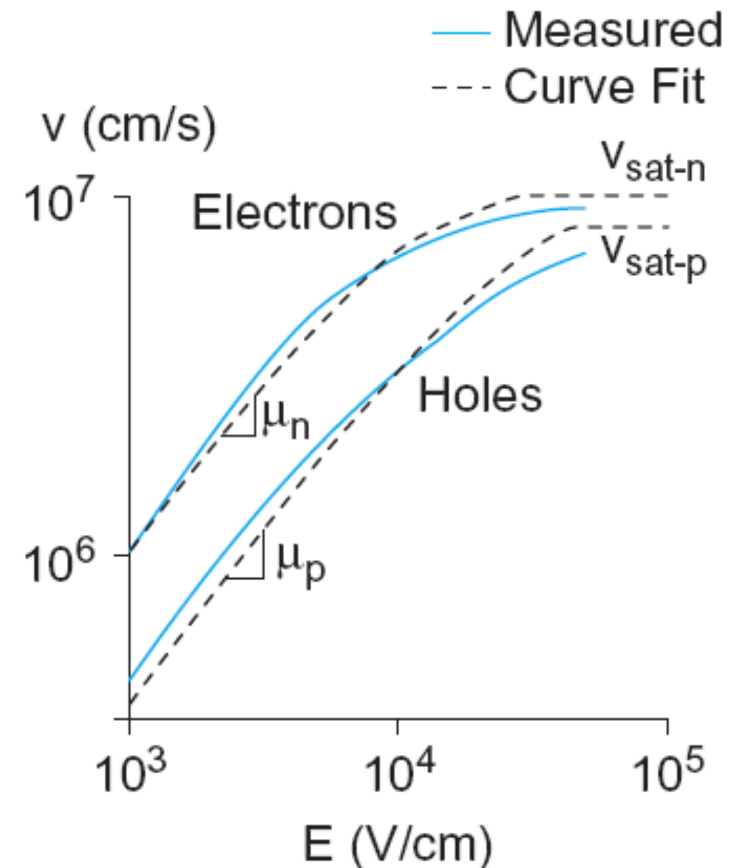
Velocity Saturation

- **At high E_{lat}, carrier velocity rolls off**
  - Carriers scatter off atoms in silicon lattice
  - Velocity reaches v_{sat}
    - Electrons: 10^7 cm/s
    - Holes: 8×10^6 cm/s
  - Better model

$$v = \begin{cases} \dfrac{\mu_{\text{eff}} E}{1 + \dfrac{E}{E_c}} & E < E_c \\ v_{\text{sat}} & E \geq E_c \end{cases}$$

where, by continuity, the *critical electric field* is

$$E_c = \frac{2 v_{\text{sat}}}{\mu_{\text{eff}}}$$



Velocity Saturation I-V Effects

- **Ideal transistor ON current increases with V_{DD}^2**

$$I_{ds} = \mu C_{\text{ox}} \frac{W}{L} \frac{\left(V_{gs} - V_t\right)^2}{2} = \frac{\beta}{2}\left(V_{gs} - V_t\right)^2$$

- **Velocity-saturated ON current increases with V_{DD}**

$$I_{ds} = C_{\text{ox}} W \left(V_{gs} - V_t\right) v_{\max}$$

- **Real transistors are partially velocity saturated**
  - Approximate with α-power law model
  - $I_{ds} \propto V_{DD}^{\alpha}$
  - $1 < \alpha < 2$ determined empirically

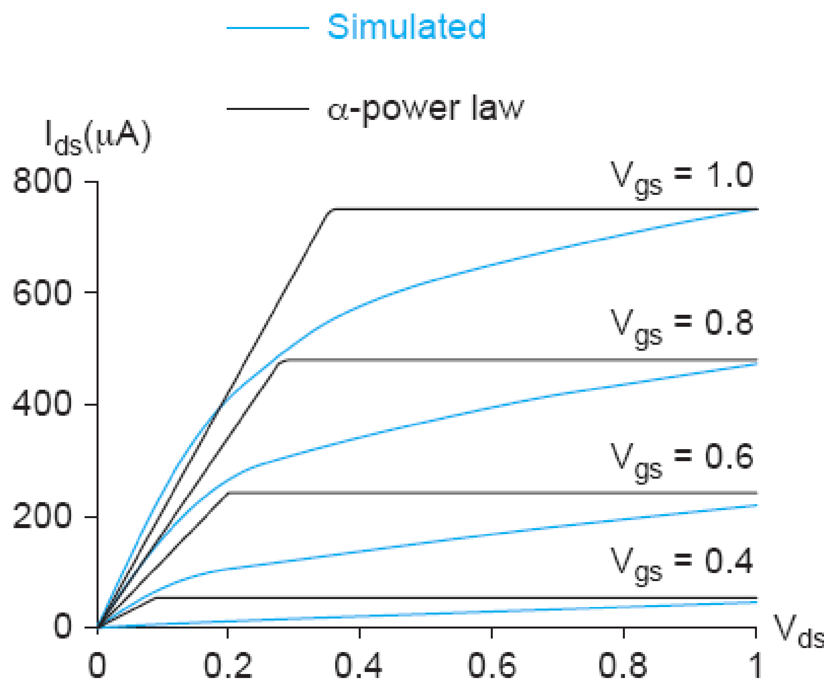
α-Power Model

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ I_{dsat}\dfrac{V_{ds}}{V_{dsat}} & V_{ds} < V_{dsat} & \text{linear} \\ I_{dsat} & V_{ds} > V_{dsat} & \text{saturation} \end{cases}$$

$$I_{dsat} = P_c \frac{\beta}{2}\left(V_{gs} - V_t\right)^{\alpha}$$

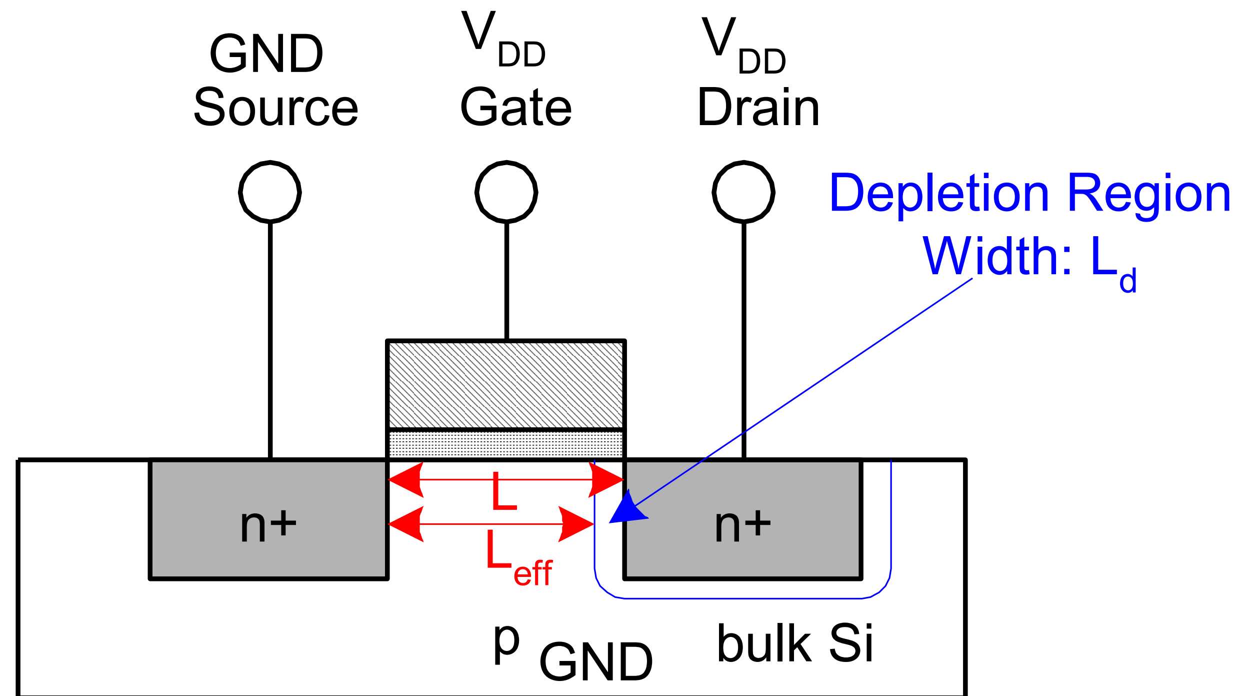
$$V_{dsat} = P_v\left(V_{gs} - V_t\right)^{\alpha/2}$$

where α, βP_c, and P_v are empirically derived



Channel Length Modulation

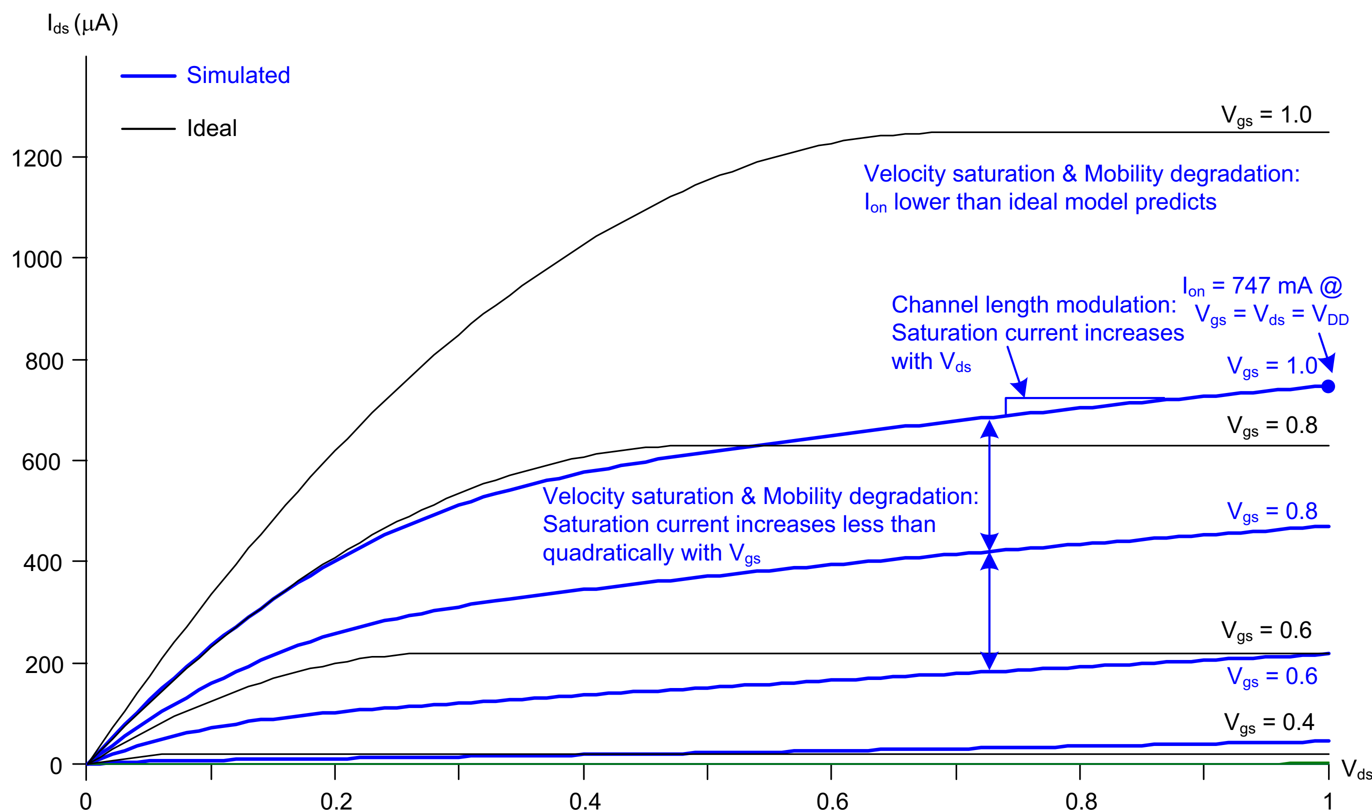
- **Reverse-biased p-n junctions form a *depletion region***
  - Region between n and p with no carriers
  - Width of depletion L_d region grows with reverse bias
  - $L_{eff} = L - L_d$
- **Shorter L_{eff} = more current**
  - I_{ds} increases with V_{ds}
  - Even in saturation



Channel Length Modulation I-V Curve

- **λ = *channel length modulation coefficient***
  - Empirically fit to I-V characteristics

$$I_{ds} = \frac{\beta}{2}\left(V_{gs} - V_t\right)^2\left(1 + \lambda V_{ds}\right)$$



Threshold Voltage Effects

- Vt is Vgs for which the channel starts to invert
- Ideal models assumed Vt is constant
- Really depends (weakly) on almost everything else:
  - Body voltage: Body Effect
  - Drain voltage: Drain-Induced Barrier Lowering
  - Channel length: Short Channel Effect

Body Effect

- **Body is a fourth transistor terminal**
- **V_{sb} affects the charge required to invert the channel**
  - **Increasing V_s or decreasing V_b increases V_t**

$$V_t = V_{t0} + \gamma\left(\sqrt{\phi_s + V_{sb}} - \sqrt{\phi_s}\right)$$

- **ϕ_s = *surface potential* at threshold**

$$\phi_s = 2v_T \ln \frac{N_A}{n_i}$$

  - **Depends on doping level N_A**
  - **And intrinsic carrier concentration n_i**

- **γ = *body effect coefficient***

$$\gamma = \frac{t_{\text{ox}}}{\varepsilon_{\text{ox}}}\sqrt{2\text{q}\varepsilon_{\text{si}}N_A} = \frac{\sqrt{2\text{q}\varepsilon_{\text{si}}N_A}}{C_{\text{ox}}}$$

Body Effect (Cont)

- **For small source-to-body voltage, treat as linear**

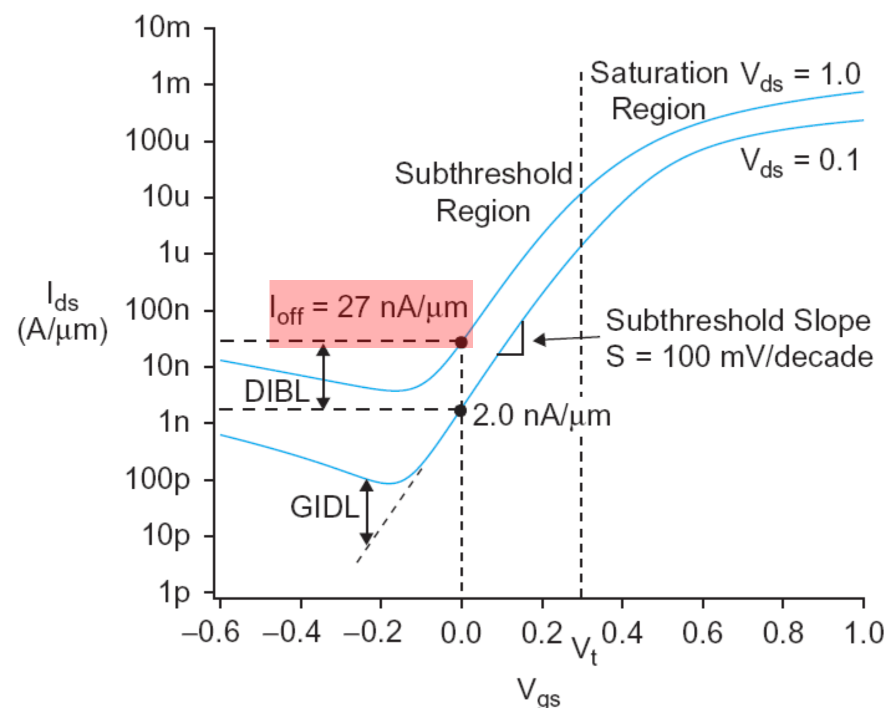
$$V_t = V_{t0} + k_\gamma V_{sb}$$

$$k_\gamma = \frac{\gamma}{2\sqrt{\phi_s}} = \frac{\sqrt{\dfrac{q\varepsilon_{si} N_A}{v_T \ln \frac{N_A}{n_i}}}}{2C_{ox}}$$

DIBL

- **Electric field from drain affects channel**
- **More pronounced in small transistors where the drain is closer to the channel**
- **Drain-Induced Barrier Lowering**
  - **Drain voltage also affect Vt**

$$V_t' = V_t - \eta V_{ds}$$



- **High drain voltage causes current to increase.**

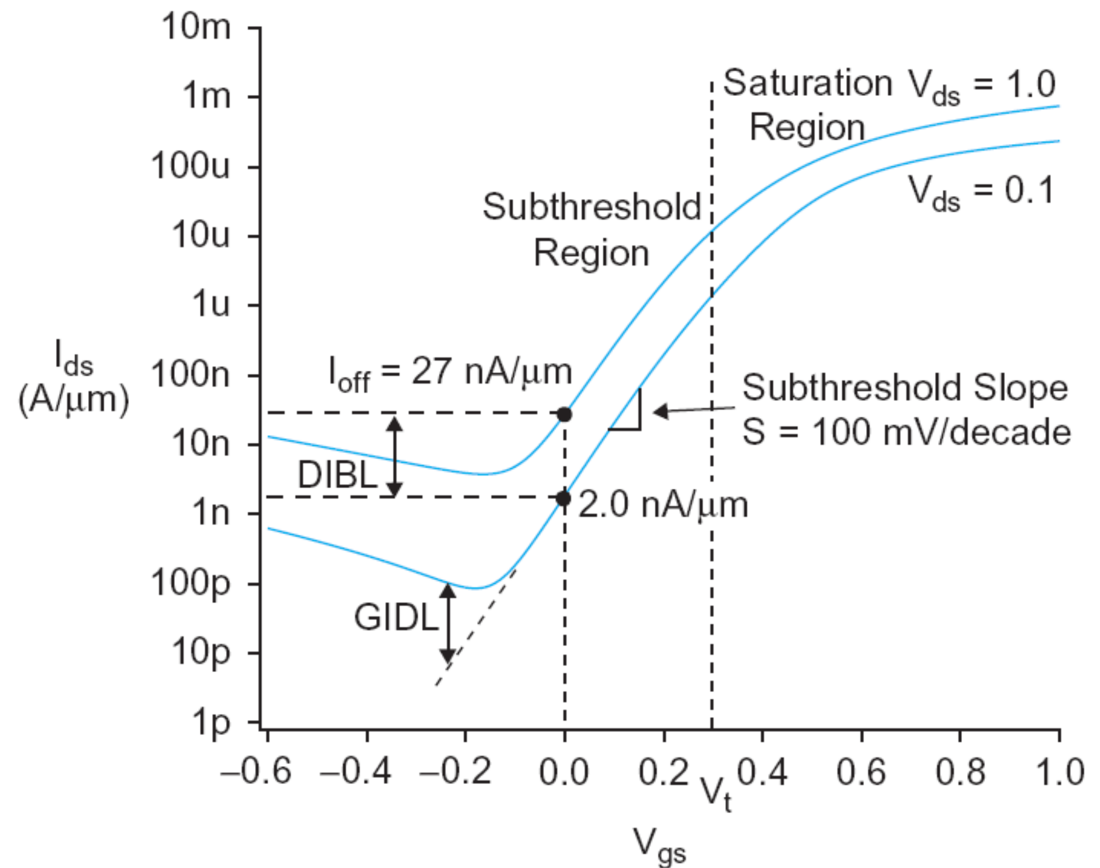
Short Channel Effect

- In small transistors, source/drain depletion regions extend into the channel
  - Impacts the amount of charge required to invert the channel
  - And thus makes Vt a function of channel length
- Short channel effect: Vt increases with L
  - Some processes exhibit a reverse short channel effect in which Vt decreases with L

Leakage

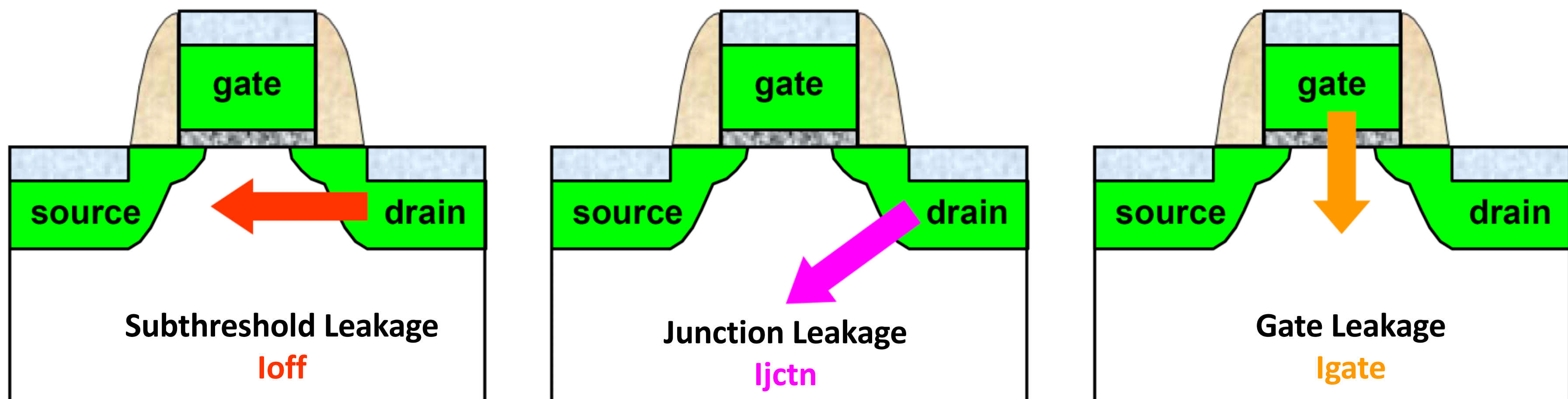
- What about current in cutoff?
- Simulated results
- What differs?

  Current doesn't go to 0 in cutoff



Leakage Sources

- **Subthreshold conduction**
  - Transistors don't turn completely OFF
- **Junction leakage**
  - Reverse-biased PN junction diode current
- **Gate leakage**
  - Tunneling through ultra-thin gate dielectric



Subthreshold leakage is the biggest source in modern transistors

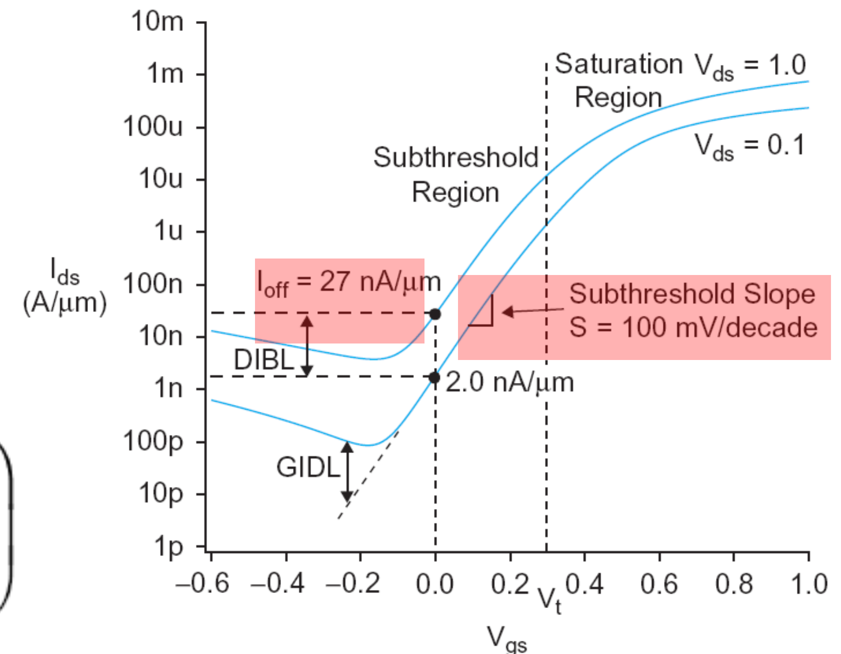
Subthreshold Leakage

- **Subthreshold leakage exponential with V_{gs}**

$$I_{ds} = I_{ds0} e^{\frac{V_{gs} - V_{t0} + \eta V_{ds} - k_\gamma V_{sb}}{n v_T}} \left(1 - e^{\frac{-V_{ds}}{v_T}}\right)$$

- **n is process dependent**
  - typically 1.3-1.7
- **Rewrite relative to I_{off} on log scale**

$$I_{ds} = I_{off} 10^{\frac{V_{gs} + \eta\left(V_{ds} - V_{dd}\right) - k_\gamma V_{sb}}{S}} \left(1 - e^{\frac{-V_{ds}}{v_t}}\right)$$



- **S ≈ 100 mV/decade @ room temperature**

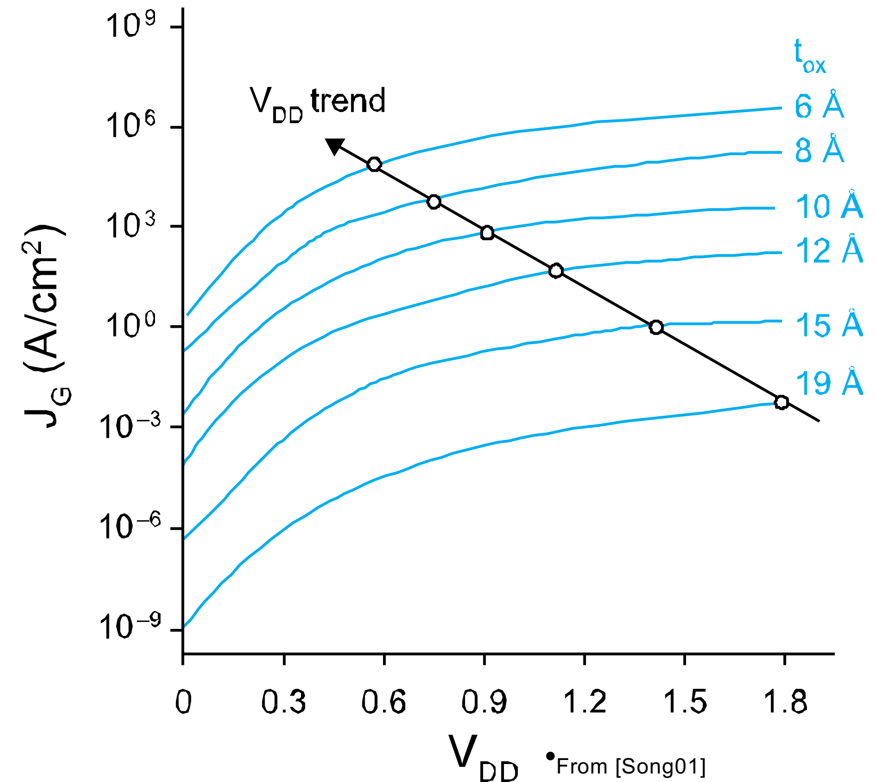
$$S = \left[\frac{d\left(\log_{10} I_{ds}\right)}{dV_{gs}}\right]^{-1} = n v_T \ln 10$$

Gate Leakage

- Carriers tunnel thorough very thin gate oxides
- Exponentially sensitive to tox and VDD

$$I_{gate} = WA\left(\frac{V_{DD}}{t_{ox}}\right)^2 e^{-B\frac{t_{ox}}{V_{DD}}}$$

  - A and B are tech constants
  - Greater for electrons
    - nMOS gates leak more



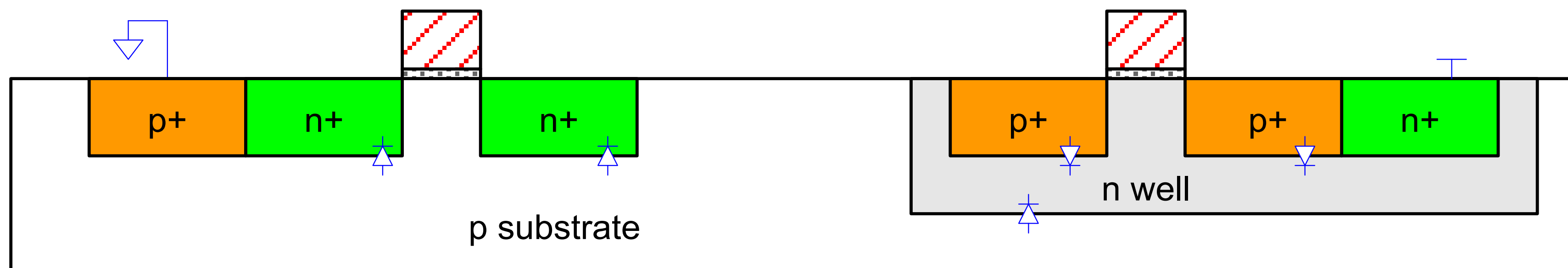
- Negligible for older processes (tox > 20 Å)
- Critically important at 65 nm and below (tox ≈ 10.5 Å)

Junction Leakage

- **Reverse-biased p-n junctions have some leakage**
  - Ordinary diode leakage
  - Band-to-band tunneling (BTBT)
  - Gate-induced drain leakage (GIDL)

$$I_D = I_S\left(e^{\frac{V_D}{v_T}} - 1\right)$$

- **At any significant negative diode voltage, $I_D = -I_s$**
- **I_s depends on doping levels**
  - And area and perimeter of diffusion regions
  - Typically < 1 fA/mm2 for > 60nm technologies

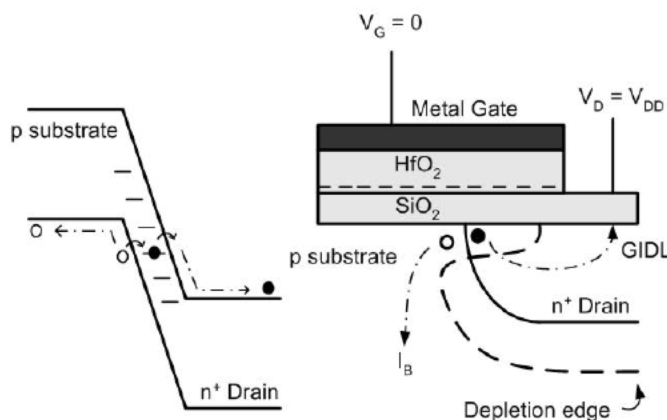


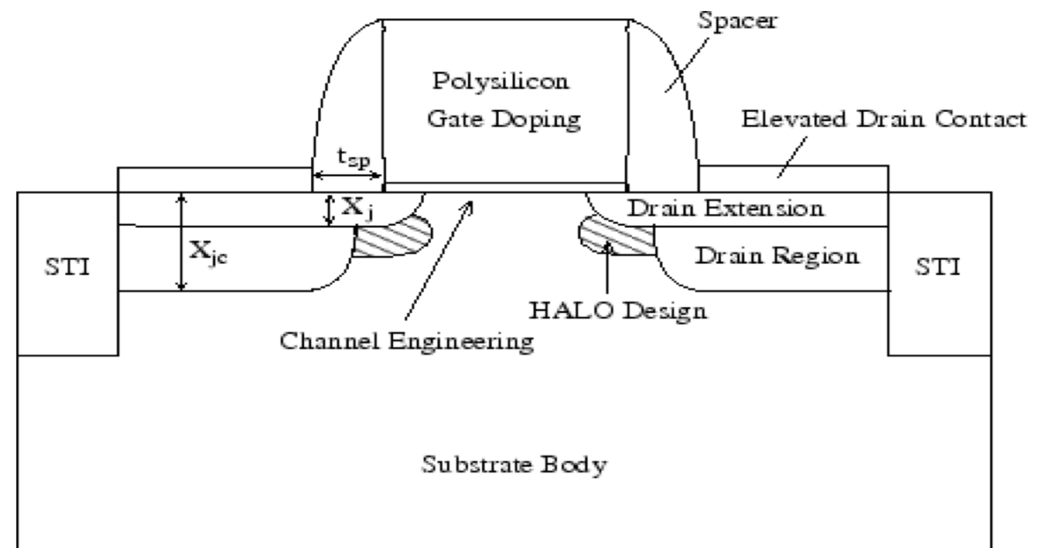
Band-to-Band Tunneling

- **Tunneling across heavily doped p-n junctions**
  - Especially sidewall between drain & channel when halo doping is used to increase V_t
- **Increases junction leakage to significant levels**

$$I_{BTBT} = W X_j A \frac{E_j}{E_g^{0.5}} V_{dd}\, e^{-B\frac{E_g^{1.5}}{E_j}} \qquad E_j = \sqrt{\frac{2qN_{halo}N_{sd}}{\varepsilon\left(N_{halo}+N_{sd}\right)}\left(V_{DD} + v_T \ln\frac{N_{halo}N_{sd}}{n_i^2}\right)}$$

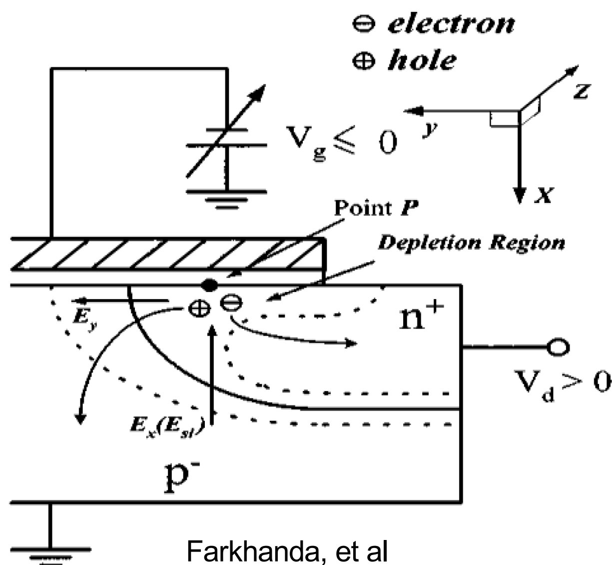
  - X_j: sidewall junction depth
  - E_g: bandgap voltage
  - A, B: tech constants



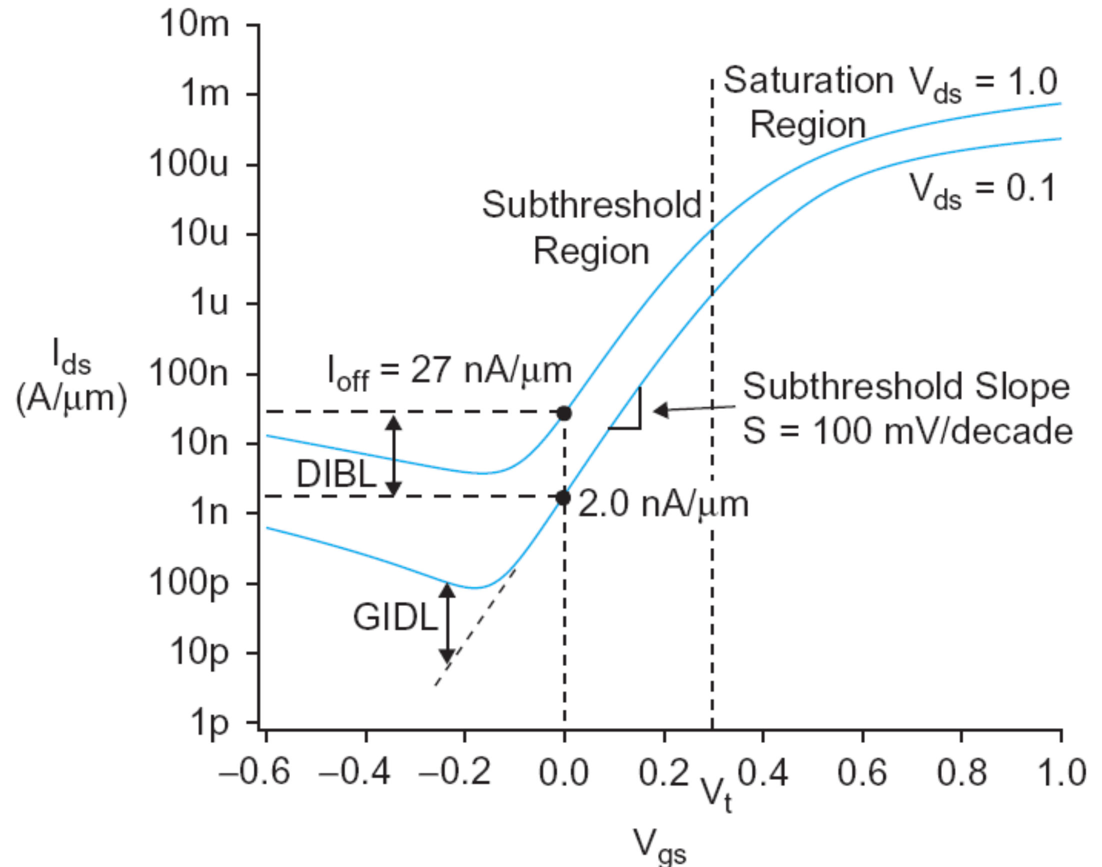


Gate-Induced Drain Leakage

- **Occurs at overlap between gate and drain**
  - **Most pronounced when drain is at V_{DD}, gate is at a negative voltage**
  - **Negates efforts to reduce subthreshold leakage using a negative gate voltage**

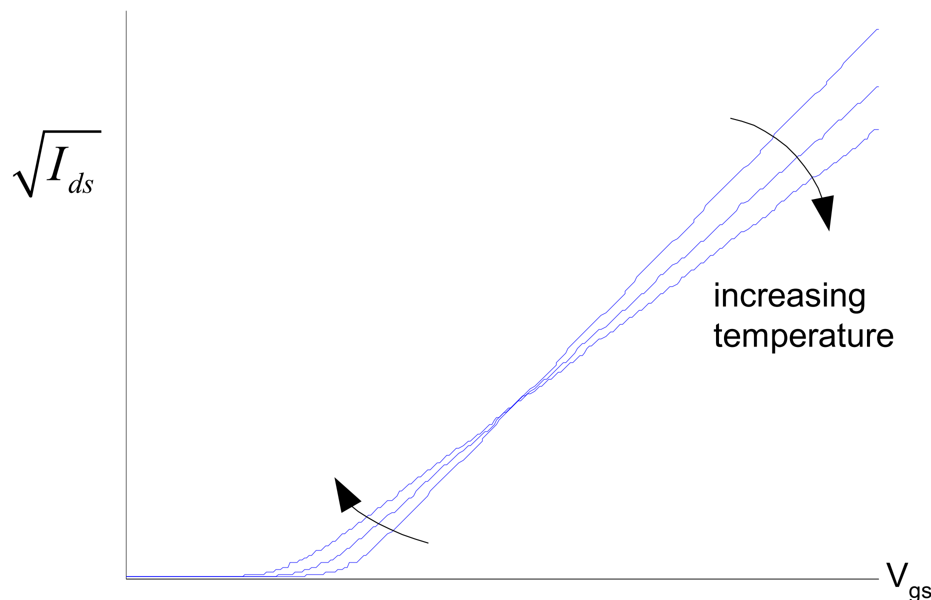


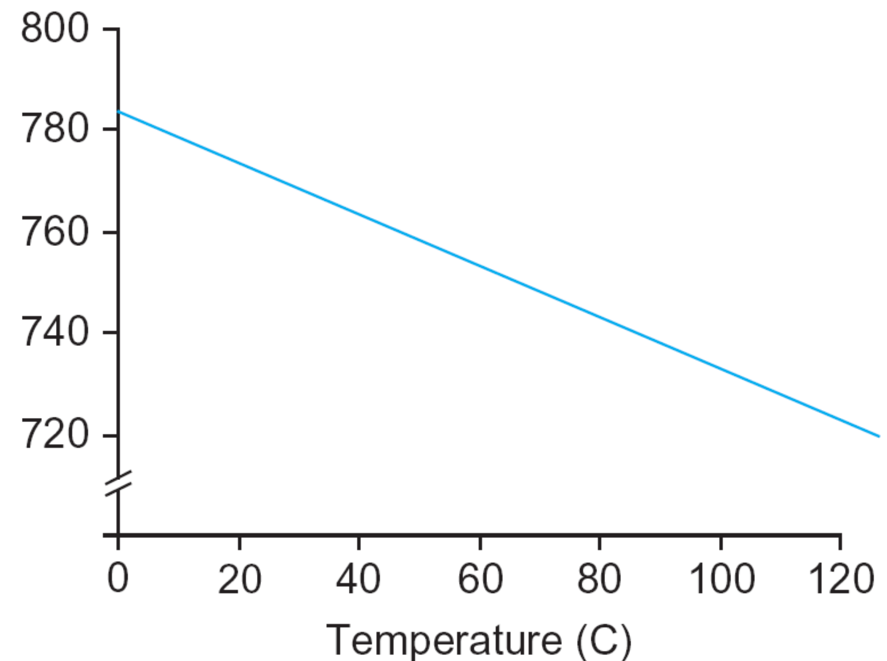
Farkhanda, et al



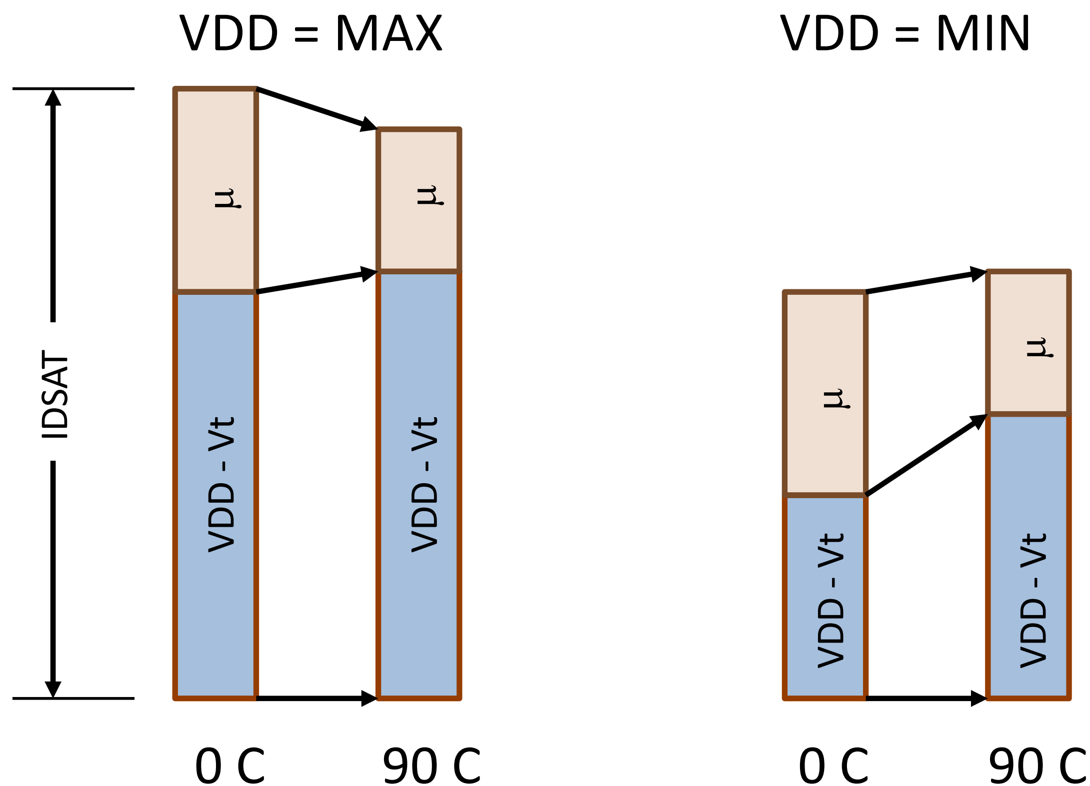
Temperature Sensitivity

- **Increasing temperature**
  - Reduces mobility
  - Reduces Vt
- **ION decreases with temperature**
- **IOFF increases with temperature**





Reverse Temperature Effect



OBSERVATIONS:

Mobility (μ) and **Vt** both drop at high temperature.

The (VDD-Vt) quantity increases at high temperature, but it is not enough to compensate for the μ drop; as a result, IDSAT drops for high VDD values.

The (VDD-Vt) quantity increases at a faster rate for low VDD; it is enough to compensate the μ drop. As a result, IDSAT increases for low VDD and high temperatures; this is called the REVERSE TEMPERATURE EFFECT.

So What?

UT ECE

- So what if transistors are not ideal?
  - They still behave like switches.

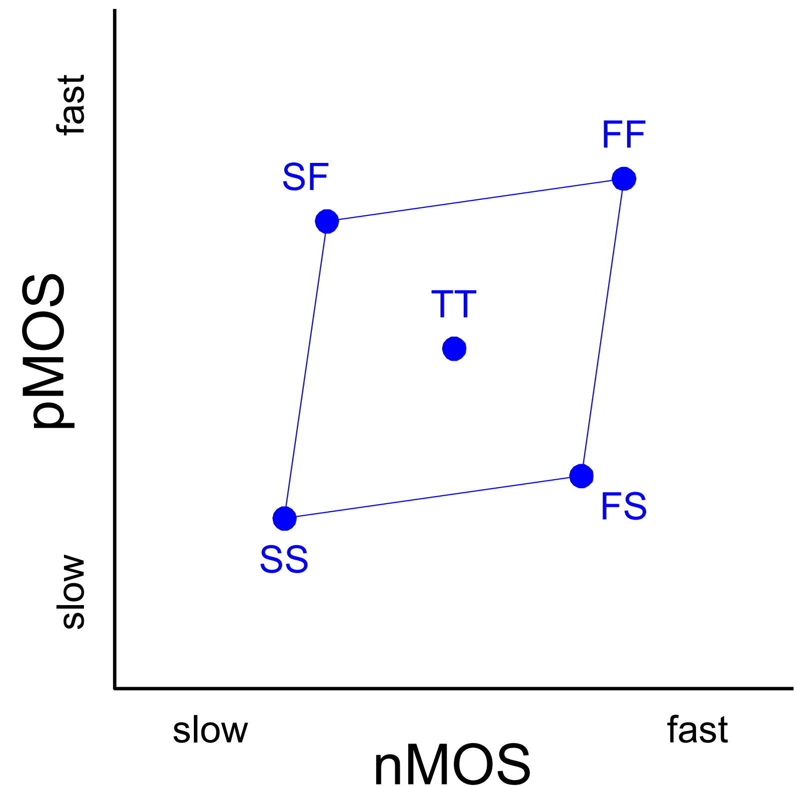
- But these effects matter for…
  - Supply voltage choice
  - Logical effort
  - Quiescent power consumption
  - Pass transistors
  - Temperature of operation

Process Corners

- **Process corners describe worst case variations**
  - **If a design works in all corners, it will probably work for any variation.**
- **Describe corners with 3 letters (T, F, S) – typical, fast, slow**
  - **Transistor speed (NMOS, PMOS)**
  - **Voltage**
  - **Temperature**

Parameter Variation

- **Transistors have uncertainty in parameters**
  - Process: Leff, Vt, tox of nMOS and pMOS
  - Vary around typical (T) values
- **Fast (F)**
  - Leff: short
  - Vt: low
  - tox: thin
- **Slow (S): opposite**
- **Not all parameters are independent for nMOS and pMOS**



Environmental Variation

- VDD and Temp also vary in time and space
- Fast corner:
  - VDD: high
  - Temp: low

Corner	Voltage	Temperature
F	1.98	0 C
T	1.8	70 C
S	1.62	125 C

Important Corners

UT ECE

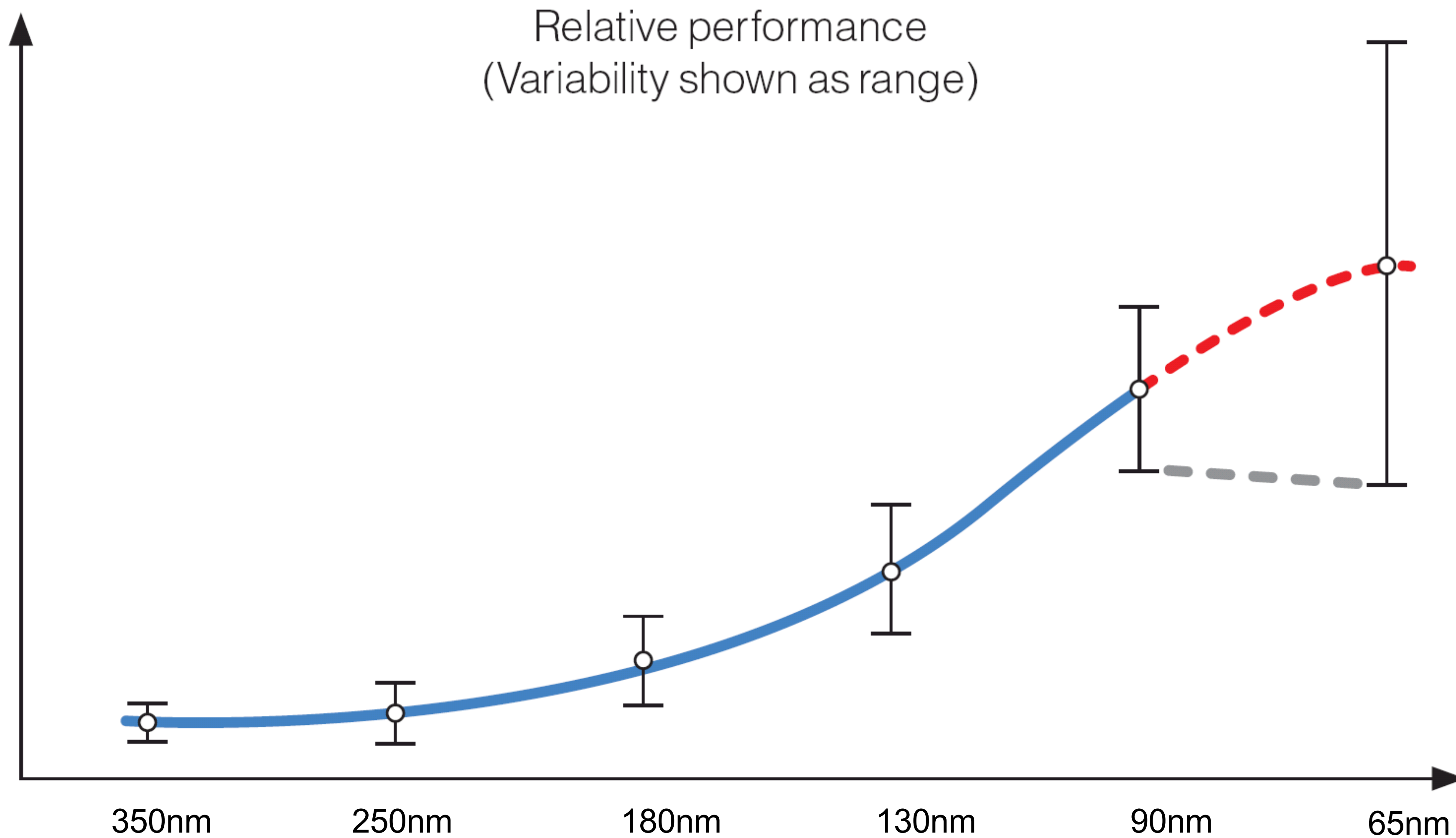
- **Some critical simulation corners include**

Purpose	nMOS	pMOS	V_{DD}	Temp
Cycle time	S	S	S	S
Power	F	F	F	F
Subthreshold leakage	F	F	F	S

THE UNIVERSITY OF TEXAS AT AUSTIN UT ECE ELECTRICAL & COMPUTER ENGINEERING

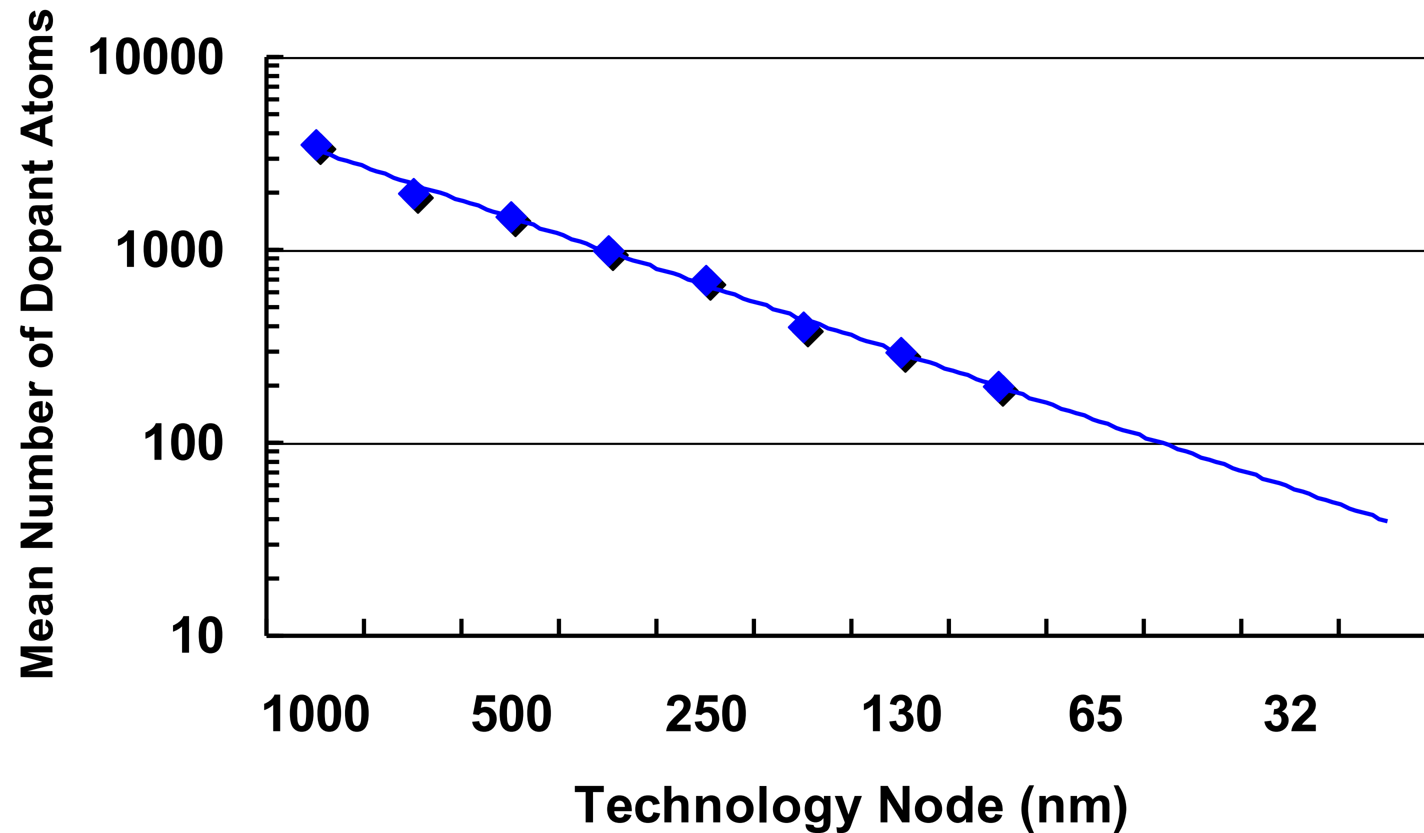
Backup

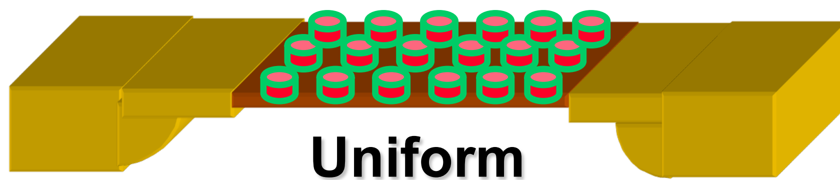
Variability per technology node

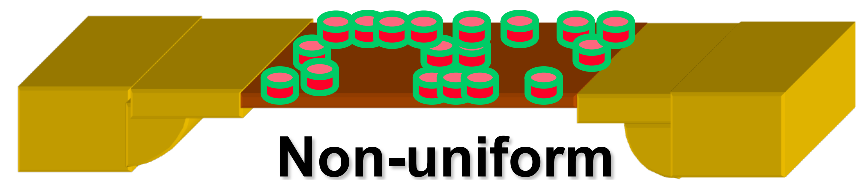


Chandu Visweswariah, *IBM Thomas J. Watson Research Center*

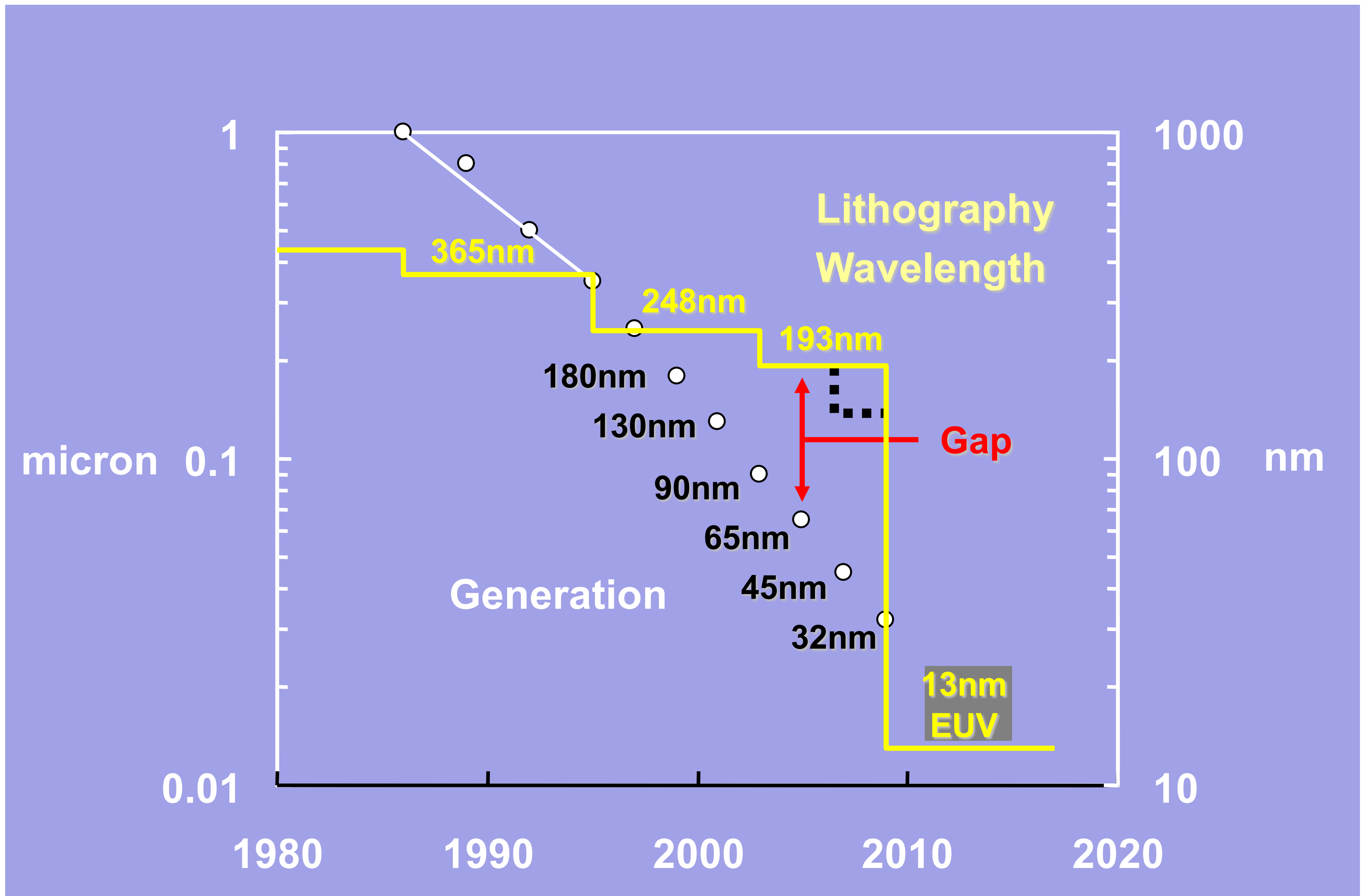
Random Dopant Fluctuations







Sub-wavelength Lithography Adds Variations



Sources of variation

- **Process variations**
  - Ion implantation, photolithography, gate etch, thermal processing, tolerance in the size of the mask image, mask alignment, angle of the etched poly-silicon, temperature, polishing, planarization
- **Subtle layout dependencies**
  - Transistor orientation
  - Post OPC effects
- **Tool errors**
  - Gate delay model, waveform model, etc.
- **Human errors**
  - Designer overwriting the tool results

F. Dartu: Variation Impact on High Performance Circuits