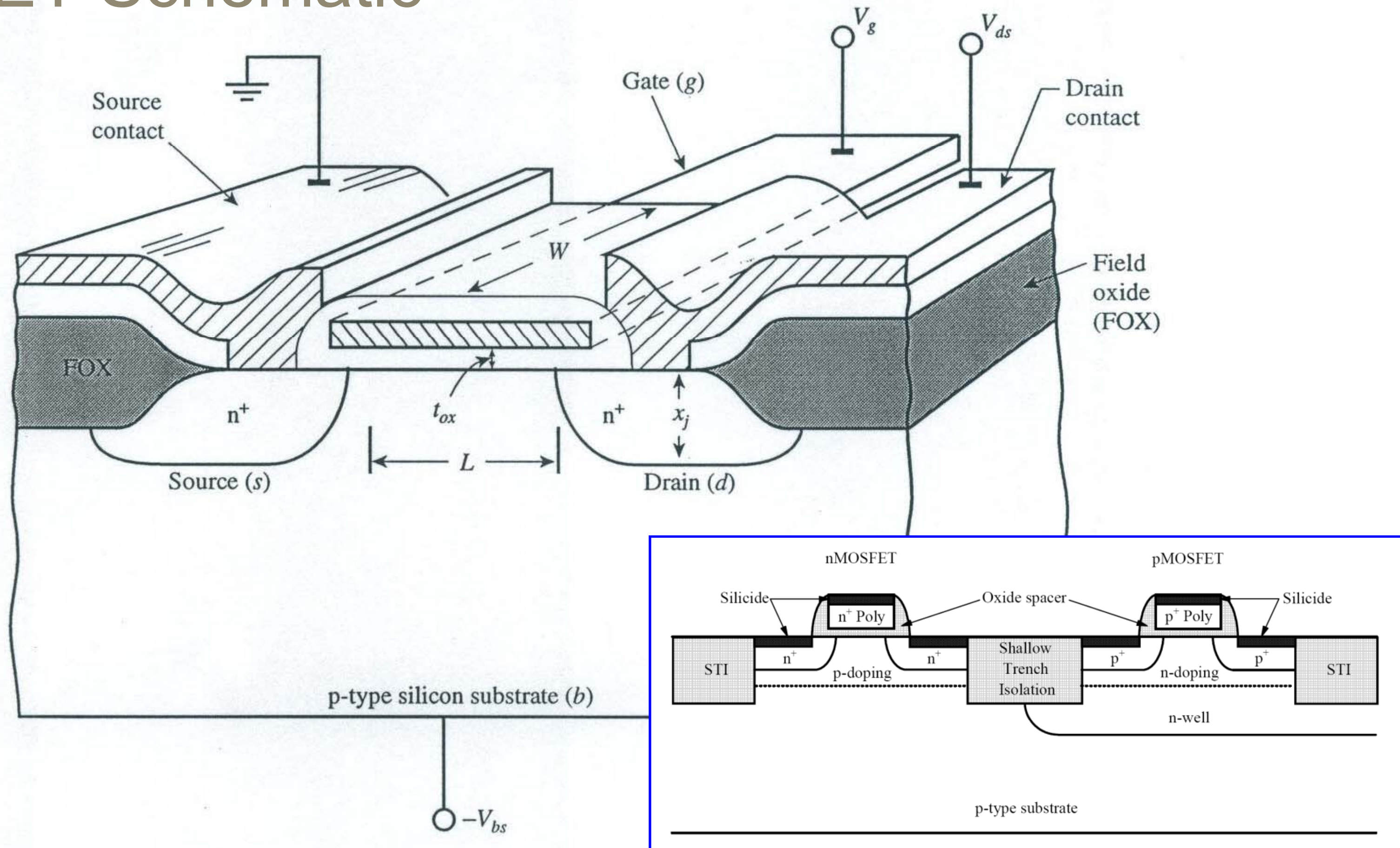
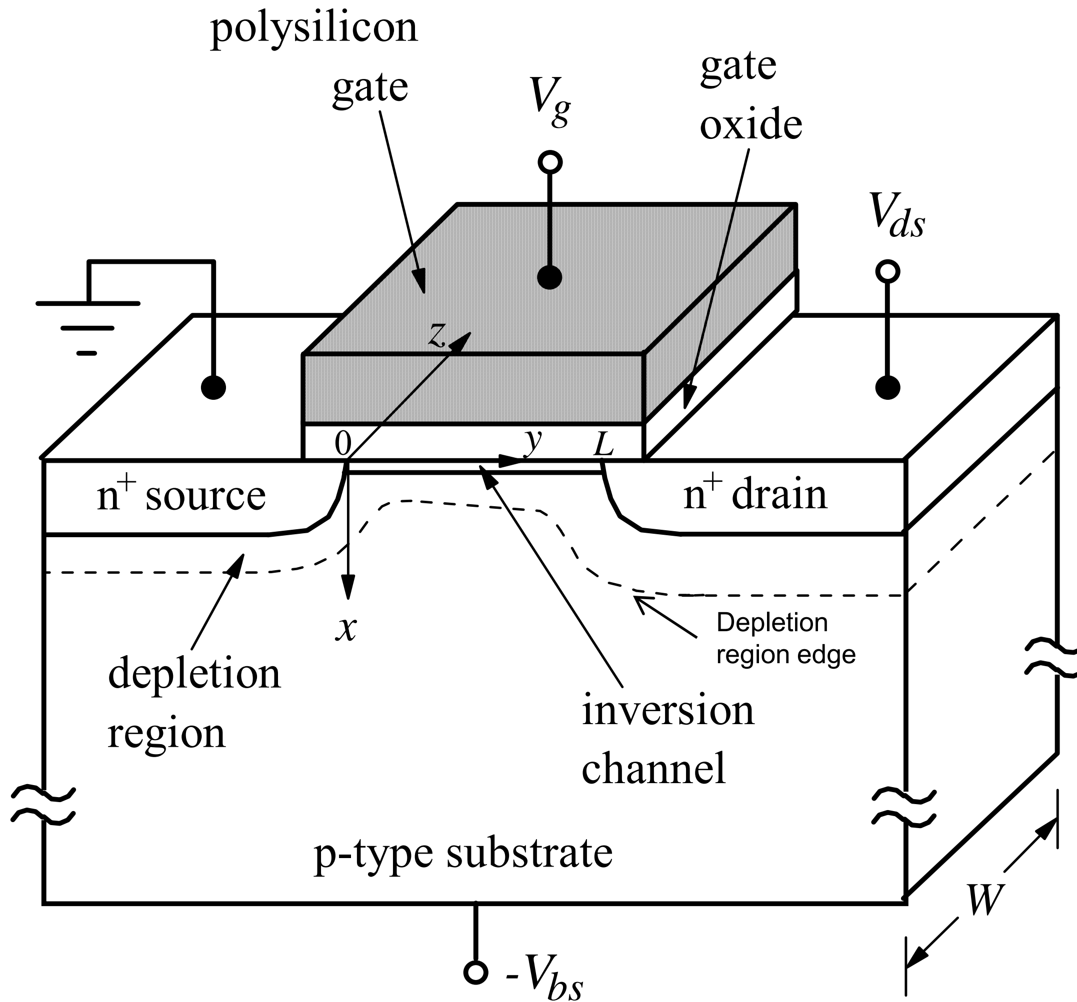
N-MOSFET Schematic



- ❑ Four structural masks: Field, Gate, Contact, Metal.
- ❑ Reverse doping polarities for pMOSFET in N-well.

N-MOSFET Schematic



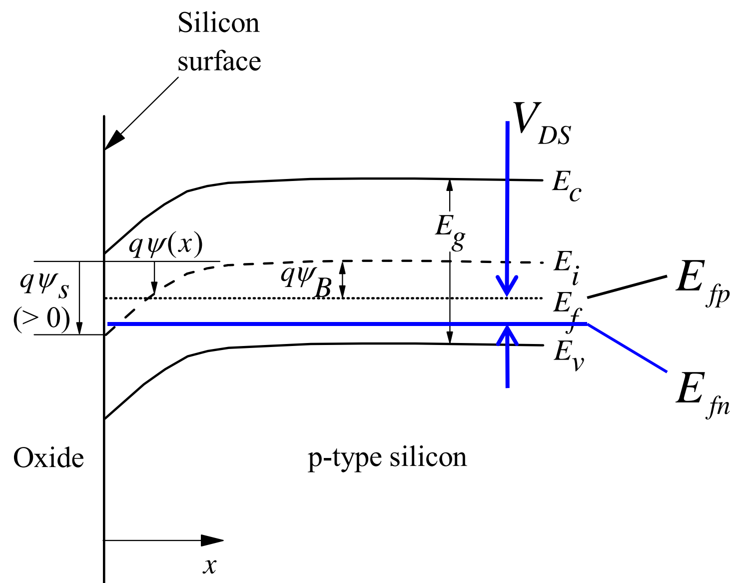
- Source terminal: Ground potential.
- Gate voltage: V_{gs}
- Drain voltage: V_{ds}
- Substrate bias voltage: $-V_{bs}$ (V_{sb})

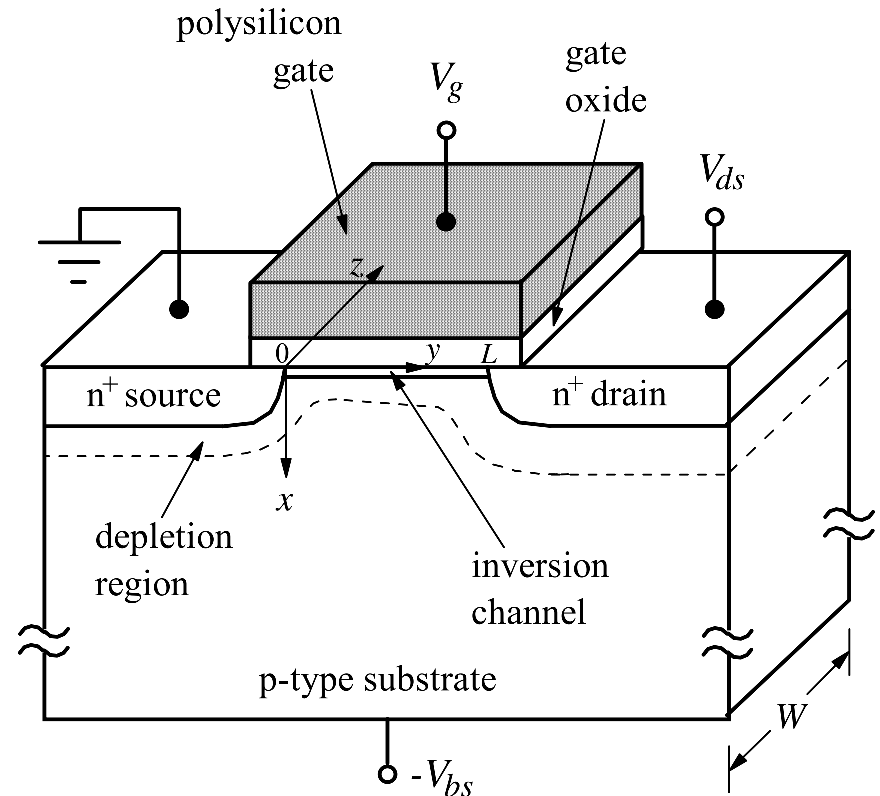
Do you remember what is quasi-Fermi level?

- $\psi(x,y)$: Band bending at any point (x,y).
- $V(y)$: Quasi-Fermi potential along the channel.
- ***Boundary conditions:*** $V(y=0) = 0$, $V(y=L) = V_{ds}$.

Long Channel Behavior

- **The electric field in the channel is essentially one-dimensional (normal to the semiconductor surface)**
- **Mathematically: $E_x >> E_y$**





Drain Current Model

Electron concentration:

$$n(x) = \frac{n_i^2}{N_a} e^{q(\psi - V)/kT}$$

Electric field:

$$\mathrm{E}^2(x,y) = \left(\frac{d\psi}{dx}\right)^2 = \frac{2kTN_a}{\varepsilon_{si}}\left[\left(e^{-q\psi/kT} + \frac{q\psi}{kT} - 1\right) + \frac{n_i^2}{N_a^2}\left(e^{-qV/kT}(e^{q\psi/kT} - 1) - \frac{q\psi}{kT}\right)\right]$$

Condition for surface inversion:

$$\psi(0,y) = V(y) + 2\psi_B$$

$V(y)$ plays the role of the reverse bias in a MOS capacitor under non-equilibrium

Maximum depletion layer width at inversion:

$$W_{dm}(y) = \sqrt{\frac{2\varepsilon_{si}[V(y) + 2\psi_B]}{qN_a}}$$

Current Density Equation

drift diffusion

$$J_n = q\mu_n nE + qD_n \frac{dn}{dy}$$

$$n = \frac{n_i^2}{N_A} e^{\frac{q(\psi - V)}{kT}} \quad ; \frac{kT}{q}\mu_n = D_n$$

$$\frac{dn}{dy} = \frac{q}{kT} n\left(\frac{d\psi}{dy} - \frac{dV}{dy}\right)$$

Current density equation (both drift and diffusion):

$$J_n(x,y) = -q\mu_n n(x,y)\frac{dV(y)}{dy}$$

Quasi-Fermi level

Gradual Channel Approximation

Assumes that vertical field (E_x) is stronger than lateral field (E_y) in the channel region, thus 2-D Poisson's equation can be solved in terms of 1-D vertical slices.

Current density equation (both drift and diffusion):

$$J_n(x,y) = -q\mu_n n(x,y)\frac{dV(y)}{dy}$$

← Quasi-Fermi level

Integrate in *x*- and *z*-directions,

$$I_{ds}(y) = -\mu_{eff} W \frac{dV}{dy} Q_i(y) = -\mu_{eff} W \frac{dV}{dy} Q_i(V)$$

where $Q_i(y) = -q\int_0^{x_i} n(x,y)dx$ is the inversion charge per unit area.

Current continuity requires I_{ds} independent of *y*, integration with respect to *y* from 0 to *L* yields

$$I_{ds} = \mu_{eff}\frac{W}{L}\int_0^{V_{ds}}\left(-Q_i(V)\right)dV$$

Pao-Sah's Double Integral

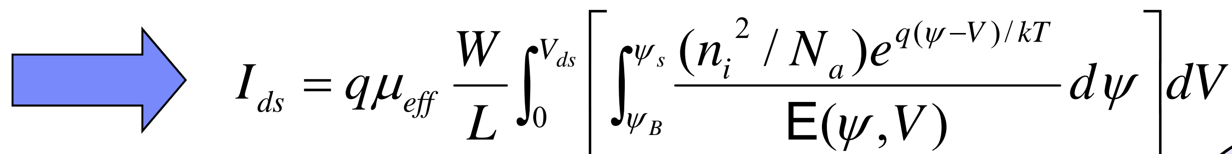
H. C. Pao and C. T. Sah, "Effects of diffusion current on characteristics of metal-oxide (insulator)-semiconductor transistors," *Solid-State Electron.,* vol. 9, no. 10, pp. 927–937, Oct. 1966.

Change variable from (x,y) to (ψ,V),

$$n(x,y) = n(\psi,V) = \frac{n_i^2}{N_a} e^{q(\psi - V)/kT}$$

$$Q_i(V) = -q\int_{\psi_s}^{\psi_B} n(\psi,V)\frac{dx}{d\psi}d\psi = -q\int_{\psi_B}^{\psi_s} \frac{(n_i^2/N_a)e^{q(\psi-V)/kT}}{\mathrm{E}(\psi,V)}d\psi$$

Substituting into the current expression,



$$I_{ds} = q\mu_{eff}\frac{W}{L}\int_0^{V_{ds}}\left[\int_{\psi_B}^{\psi_s}\frac{(n_i^2/N_a)e^{q(\psi-V)/kT}}{\mathrm{E}(\psi,V)}d\psi\right]dV$$

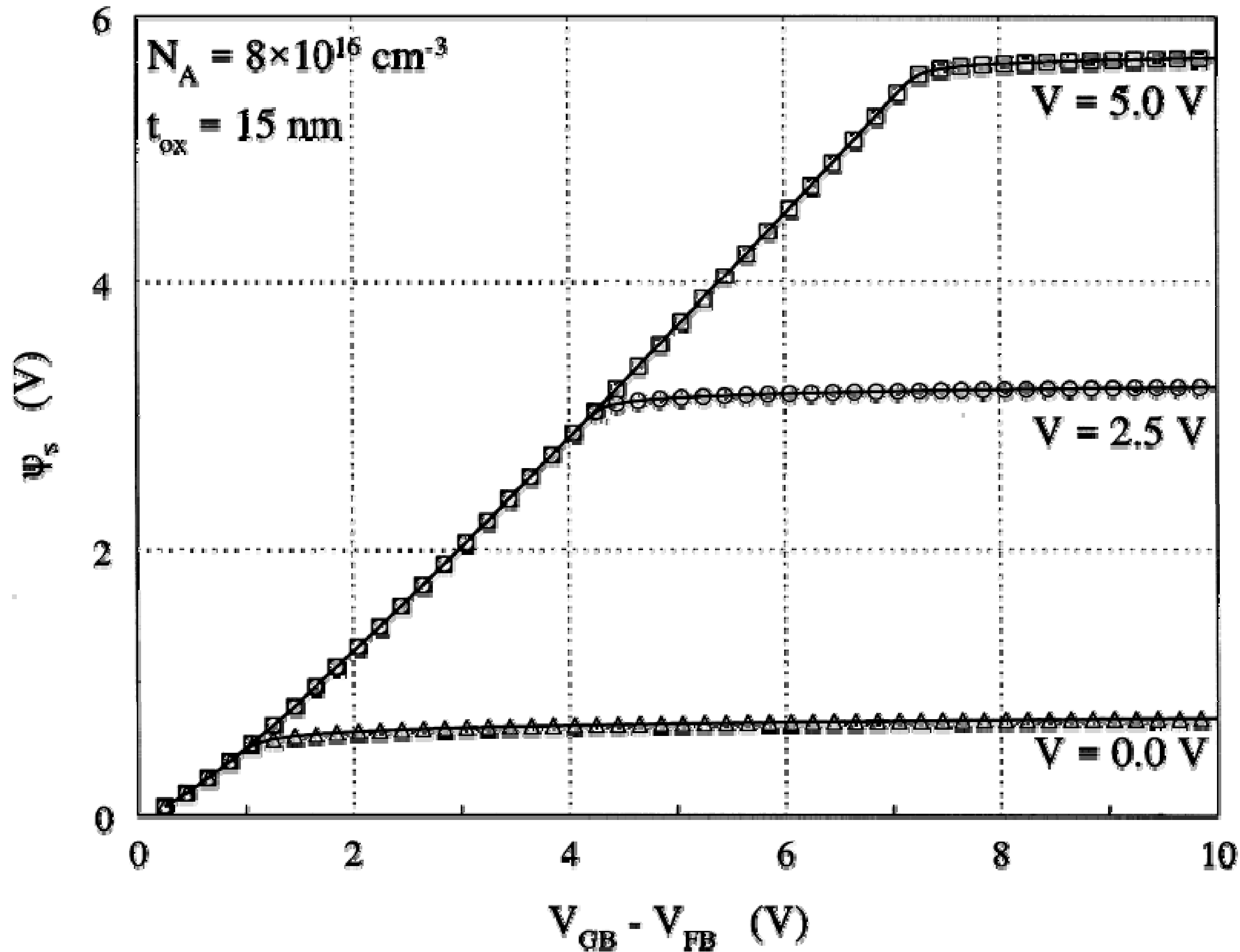
where $\psi_s(V)$ is solved by the gate voltage eq. for a vertical slice of the MOSFET:

$$V_g = V_{fb} + \psi_s - \frac{Q_s}{C_{ox}} = V_{fb} + \psi_s + \frac{\sqrt{2\varepsilon_{si}kTN_a}}{C_{ox}}\left[\frac{q\psi_s}{kT} + \frac{n_i^2}{N_a^2}e^{q(\psi_s - V)/kT}\right]^{1/2}$$

How do you get this approximation?

(see Lecture Notes p. 2-12 and 2-22 and make approximations)

Example of ψ_s vs V_G Relationship



R. van Langevelde, F.M. Klaassen / Solid-State Electronics 44 (2000) 409 - 418

Charge Sheet Approximation

Assumes that all the inversion charges are located at the silicon surface like a sheet of charge and that there is no potential drop across the inversion layer.

After the onset of inversion, the surface potential is pinned at $\psi_s = 2\psi_B + V(y)$.

- Depletion charge: $Q_d = -qN_a W_{d\max} = -\sqrt{2\varepsilon_{si} q N_a (2\psi_B + V)}$
- Total charge: $Q_s = -C_{ox}(V_g - V_{fb} - \psi_s) = -C_{ox}(V_g - V_{fb} - 2\psi_B - V)$
- Inv. charge: $Q_i = Q_s - Q_d = -C_{ox}(V_g - V_{fb} - 2\psi_B - V) + \sqrt{2\varepsilon_{si} q N_a (2\psi_B + V)}$

Substituting in $I_{ds} = \mu_{eff} \frac{W}{L} \int_0^{V_{ds}} \left(-Q_i(V)\right) dV$ and integrate:

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L} \left\{ \left(V_g - V_{fb} - 2\psi_B - \frac{V_{ds}}{2} \right) V_{ds} - \frac{2\sqrt{2\varepsilon_{si} q N_a}}{3C_{ox}} \left[(2\psi_B + V_{ds})^{3/2} - (2\psi_B)^{3/2} \right] \right\}$$

J.R. Brews, “A charge-sheet model of the MOSFET,” Solid-State Electronics, Volume 21, Issue 2, February 1978, Pages 345-355

Linear Region I-V Characteristics

For $V_{ds} << V_g$,

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L}\left(V_g - V_{fb} - 2\psi_B - \frac{\sqrt{4\varepsilon_{si} q N_a \psi_B}}{C_{ox}}\right)V_{ds} = \mu_{eff} C_{ox} \frac{W}{L}(V_g - V_t)V_{ds}$$

where $V_t = V_{fb} + 2\psi_B + \frac{\sqrt{4\varepsilon_{si} q N_a \psi_B}}{C_{ox}}$ is the MOSFET ***threshold voltage***.

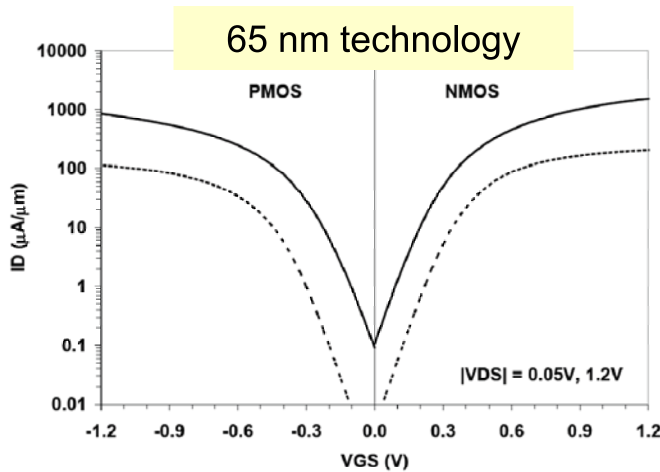
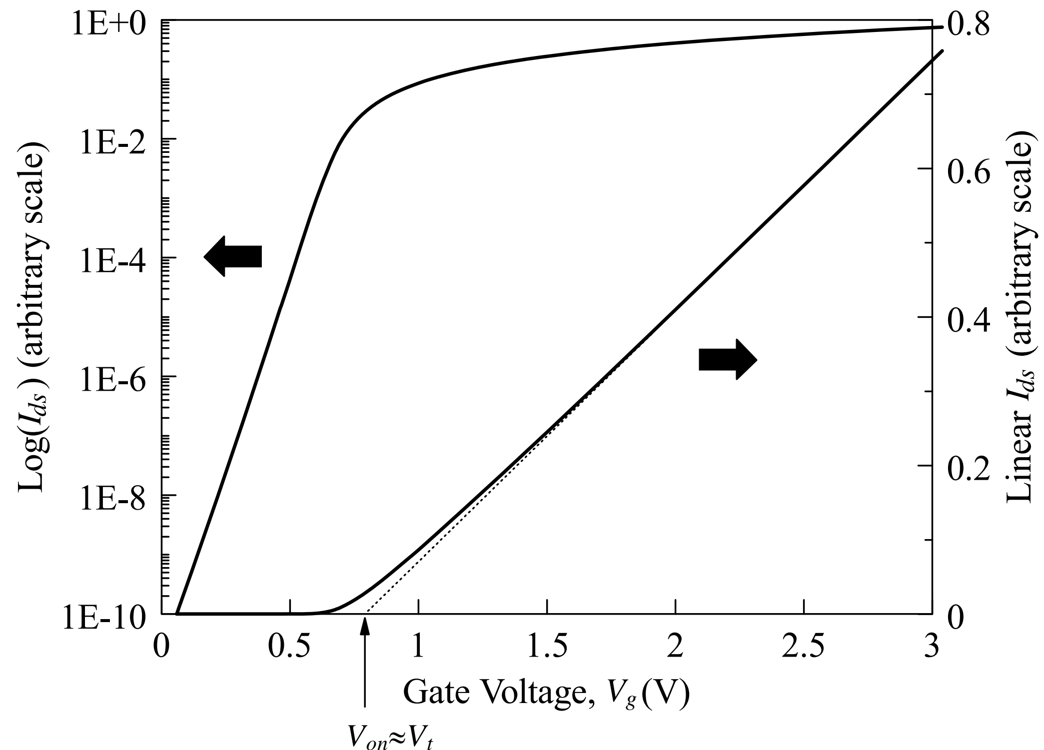


Figure 6: Sub-threshold curves for 35nm Lgate devices



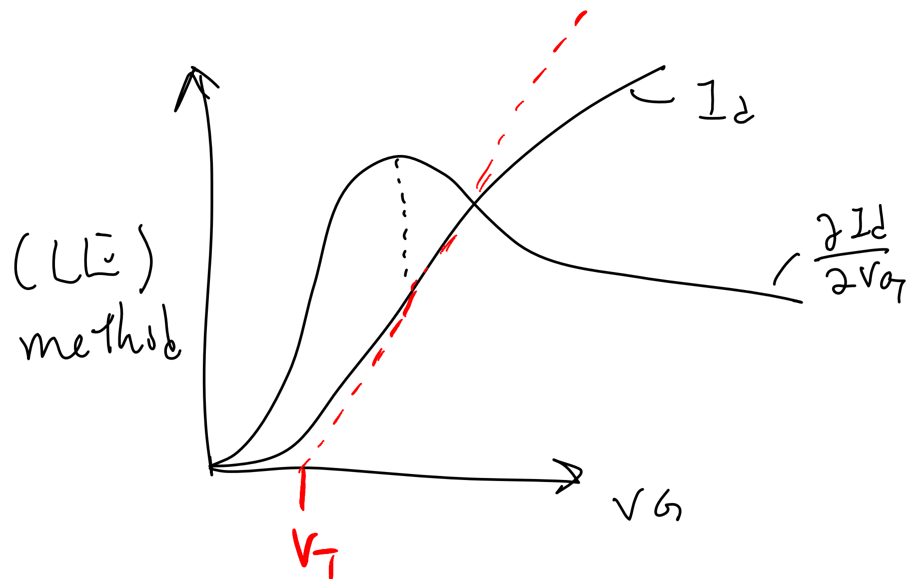
P. Bai et al., "A 65nm Logic Technology Featuring 35nm Gate Lengths, Enhanced Channel 8 Cu Interconnect Layers, Low-k ILD and 0.57 µm² SRAM Cell," *IEDM,* p. 657 (2004).

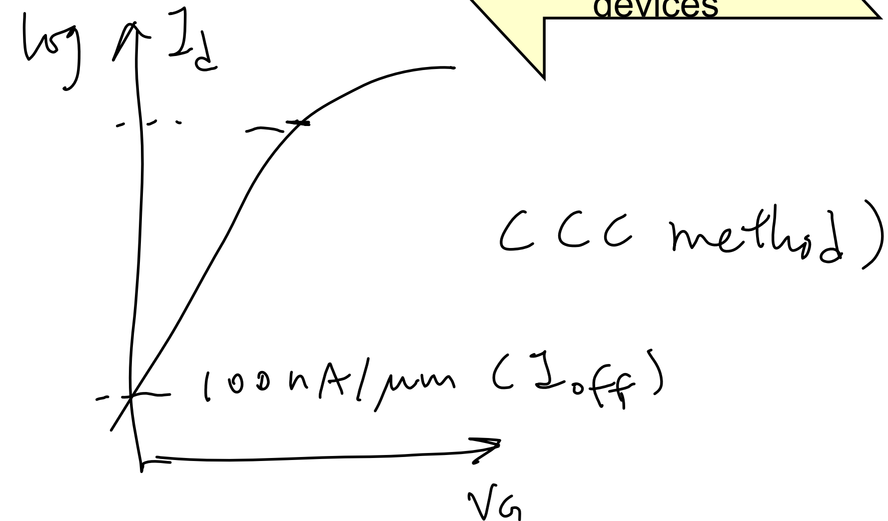
Experimental Determination of the Threshold Voltage

- **Linear extrapolation (LE) at the maximum G_m point**
- **Constant current (CC) method**
- **Transconductance change (TC) method**

Often used in industry

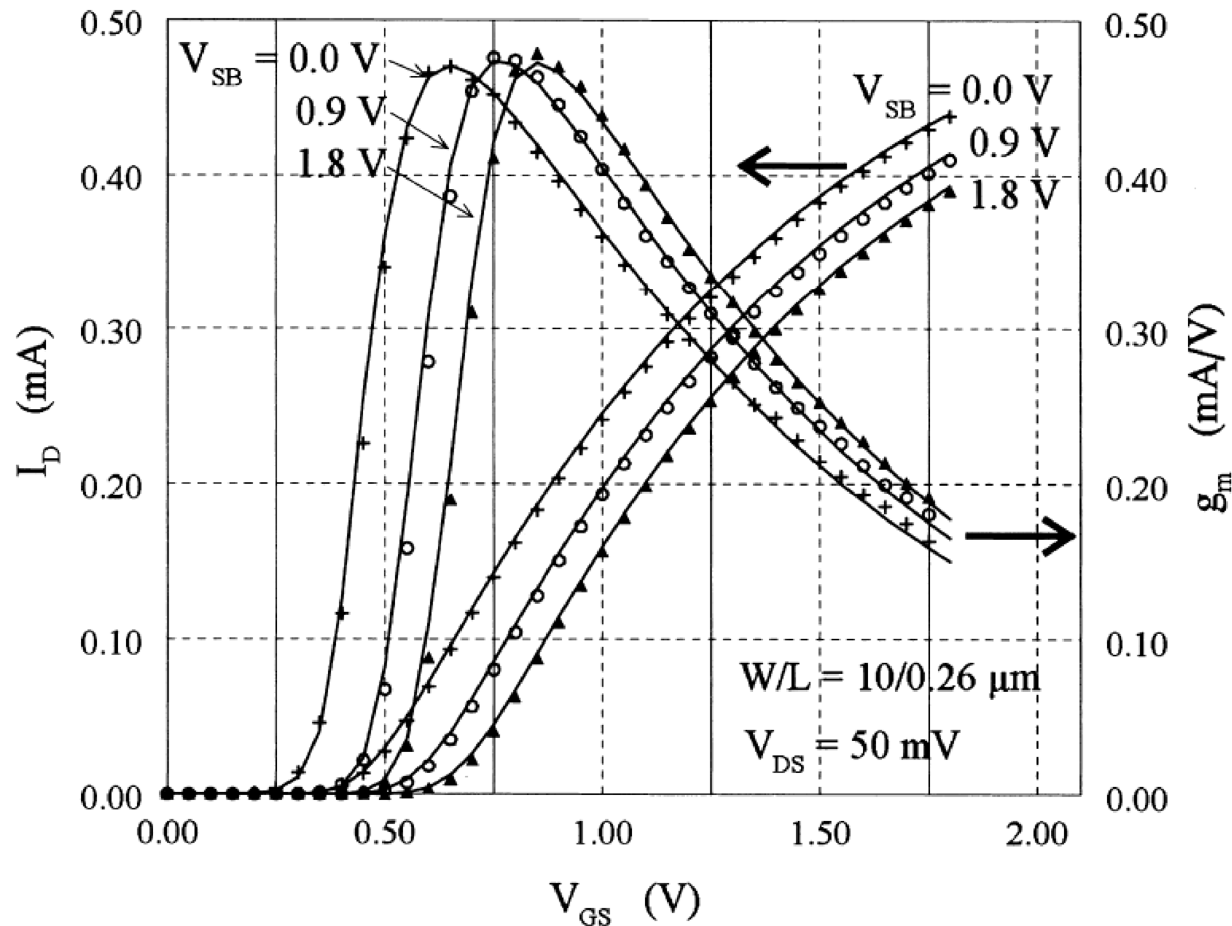
Used for modeling of devices





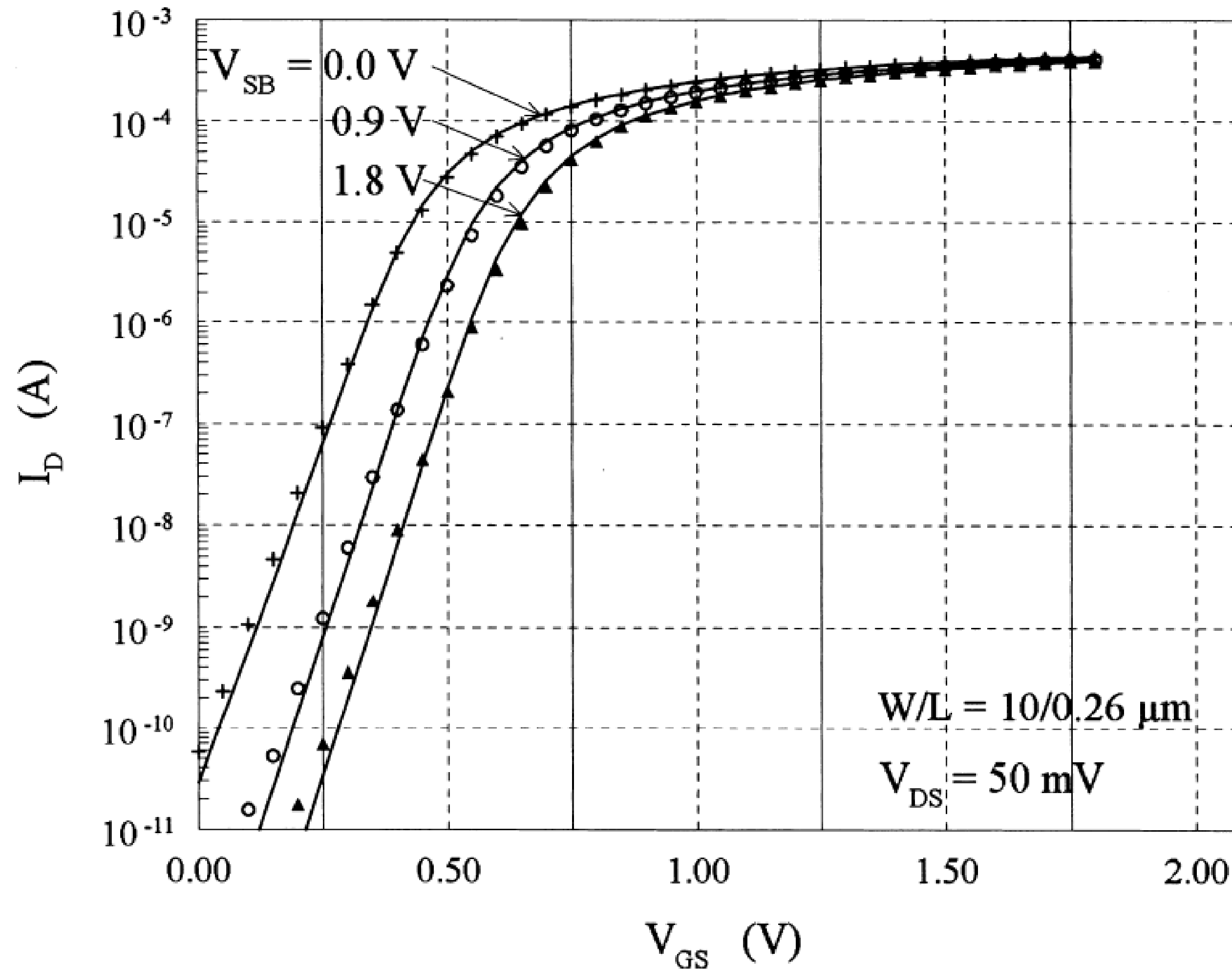
D.K. Schroeder *Semiconductor material and device characterization, 2nd ed*, Wiley, New York (1998).

Example I_d (and G_m) vs V_{GS} curves



R. van Langevelde, F.M. Klaassen / Solid-State Electronics 44 (2000) 409 - 418

Example log(I_d) vs V_{GS} curves



R. van Langevelde, F.M. Klaassen / Solid-State Electronics 44 (2000) 409 - 418

Threshold Voltage Extraction Method Illustration

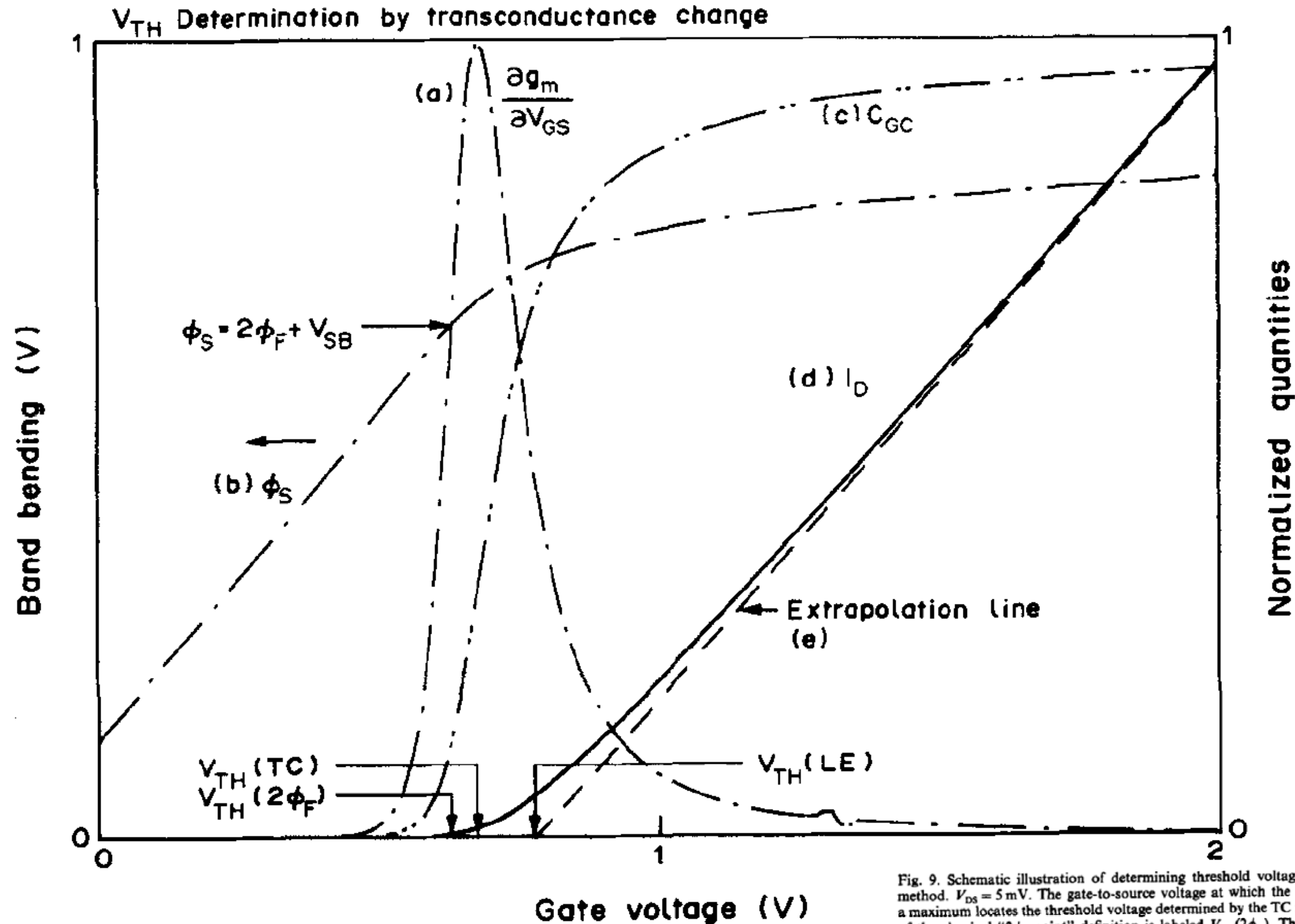
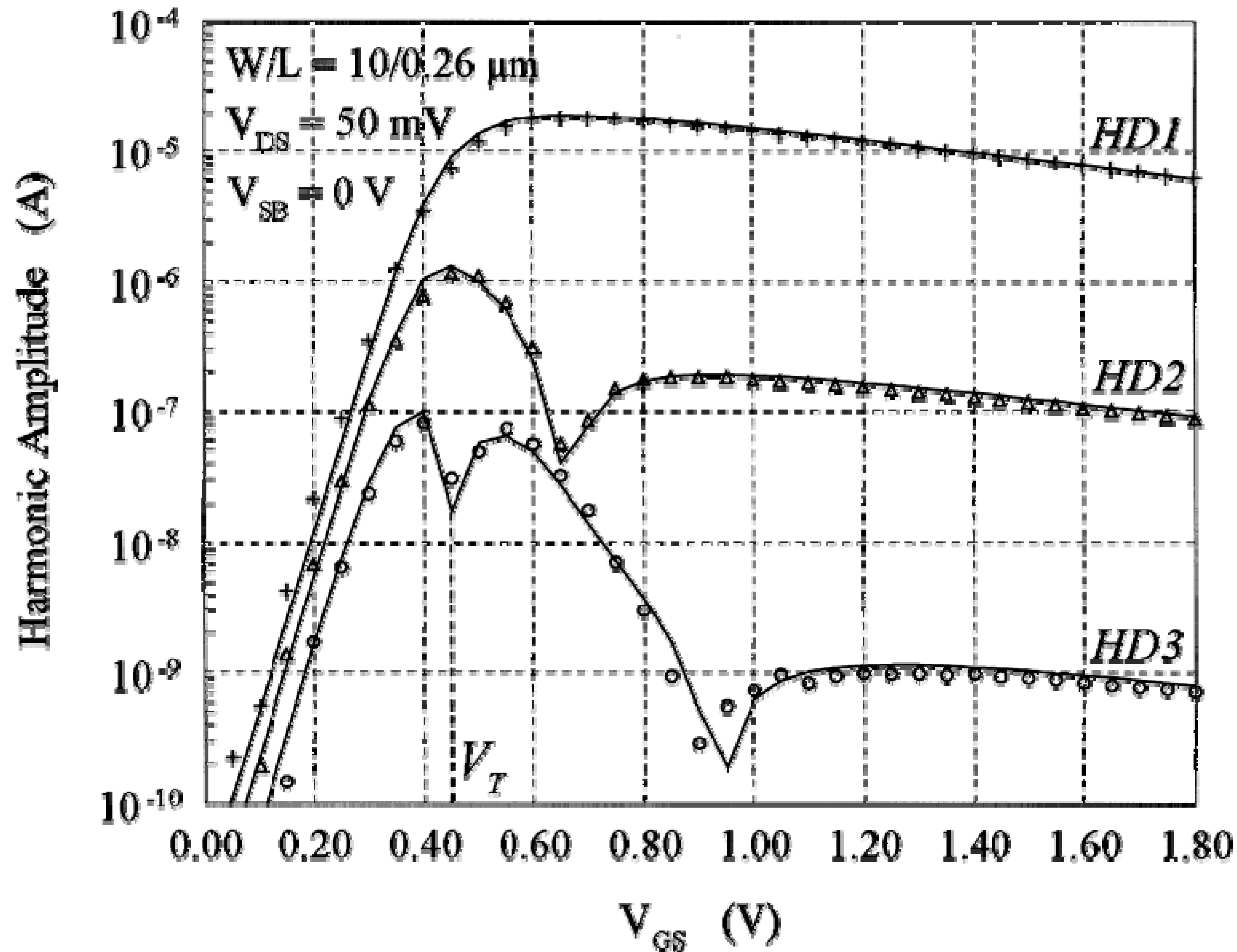


Fig. 9. Schematic illustration of determining threshold voltage by the transconductance change (TC) method. $V_{DS} = 5$ mV. The gate-to-source voltage at which the derivative of the transconductance (a) is a maximum locates the threshold voltage determined by the TC method [V_{TH}(TC)]. The threshold voltage of the classical "$2\phi_F$ point" definition is labeled $V_{TH}(2\phi_F)$. The linearly extrapolated threshold voltage is labeled V_{TH}(LE). The extrapolation line is drawn through the point of maximum slope (transconductance) of the I_D curve with slope equal to the peak transconductance. The "knee" point of the band-bending ϕ_s (b) corresponds to the point where the derivative of the transconductance goes through a maximum. This figure also shows the gate-channel capacitance (c), the drain current (d), and the linear extrapolation line (e) (dashed line) for reference.

H.S. Wong, M.H. White, T.J. Krutsick and R.V. Booth, Modeling of transconductance degradation and extraction of threshold voltage in thin oxide MOSFET's. *Solid State Electron* **30** (1987), p. 953.

Example $\partial I_d / \partial V_{GS}$ and $\partial^2 I_d / \partial V_{GS}{}^2$ vs V_{GS} curves



R. van Langevelde, F.M. Klaassen / Solid-State Electronics 44 (2000) 409 - 418

Saturation Region I-V Characteristics

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L}\left\{\left(V_g - V_{fb} - 2\psi_B - \frac{V_{ds}}{2}\right)V_{ds} - \frac{2\sqrt{2\varepsilon_{si} q N_a}}{3C_{ox}}\left[(2\psi_B + V_{ds})^{3/2} - (2\psi_B)^{3/2}\right]\right\}$$

Keeping the 2nd order terms in V_{ds}: $I_{ds} = \mu_{eff} C_{ox} \frac{W}{L}\left[(V_g - V_t)V_{ds} - \frac{m}{2} V_{ds}^{\ 2}\right]$

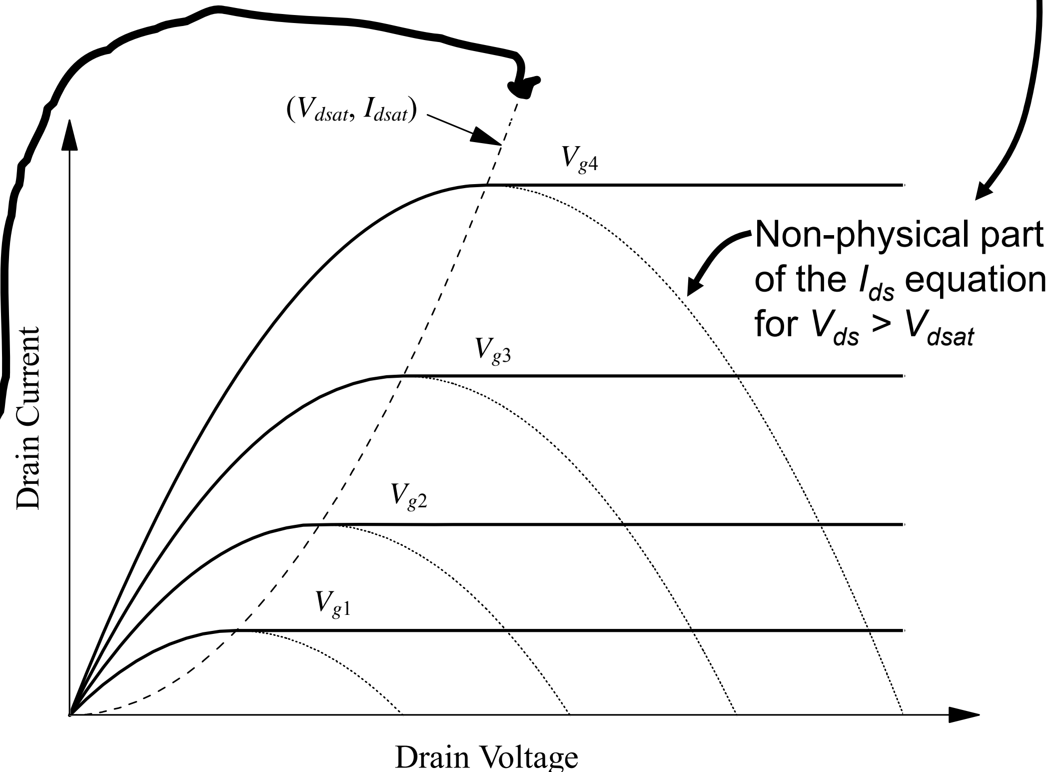
where $m = 1 + \frac{\sqrt{\varepsilon_{si} q N_a / 4\psi_B}}{C_{ox}} = 1 + \frac{C_{dm}}{C_{ox}} = 1 + \frac{3t_{ox}}{W_{dm}}$ is the body-effect coefficient

$$I_{ds} = I_{dsat} = \mu_{eff} C_{ox} \frac{W}{L}\frac{(V_g - V_t)^2}{2m}$$

when

$V_{ds} = V_{dsat} = (V_g - V_t)/m$

Typically, $m \approx 1.2$



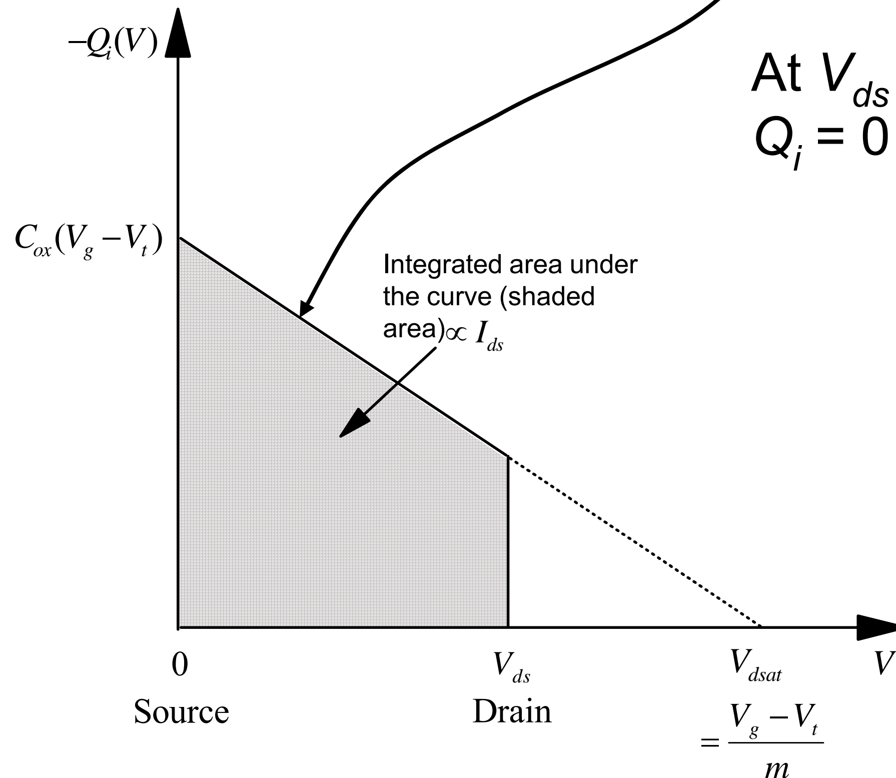
Pinch-off Condition

From inversion charge density point of view,

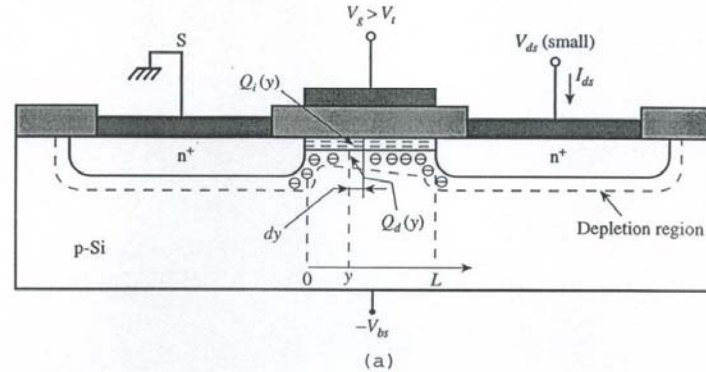
$$Q_i(V) = -C_{ox}(V_g - V_t - mV)$$

while $I_{ds} = \mu_{eff}\frac{W}{L}\int_0^{V_{ds}}(-Q_i(V))dV$

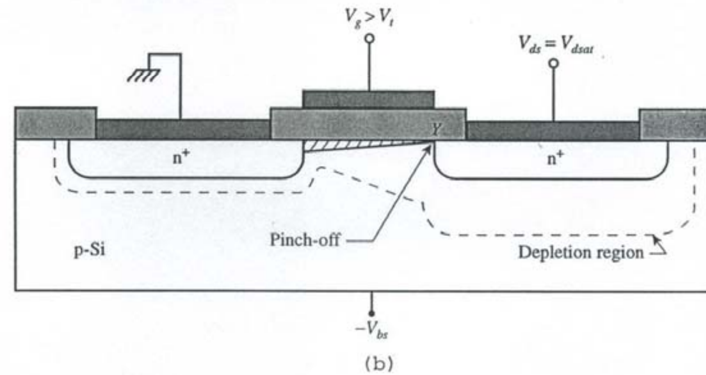
At $V_{ds} = V_{dsat} = (V_g - V_t)/m$,
$Q_i = 0$ and I_{ds} = max.



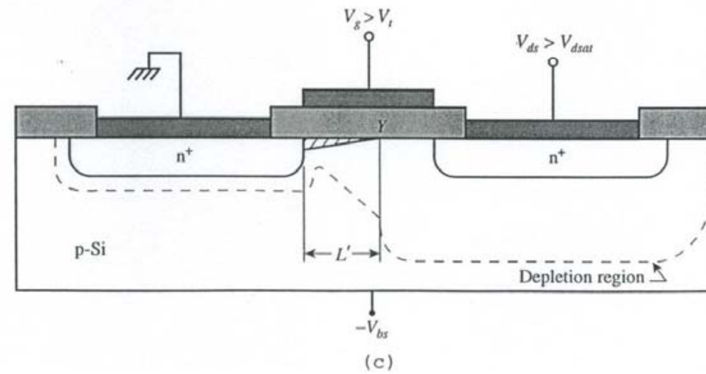
Beyond Pinch-off



(a)



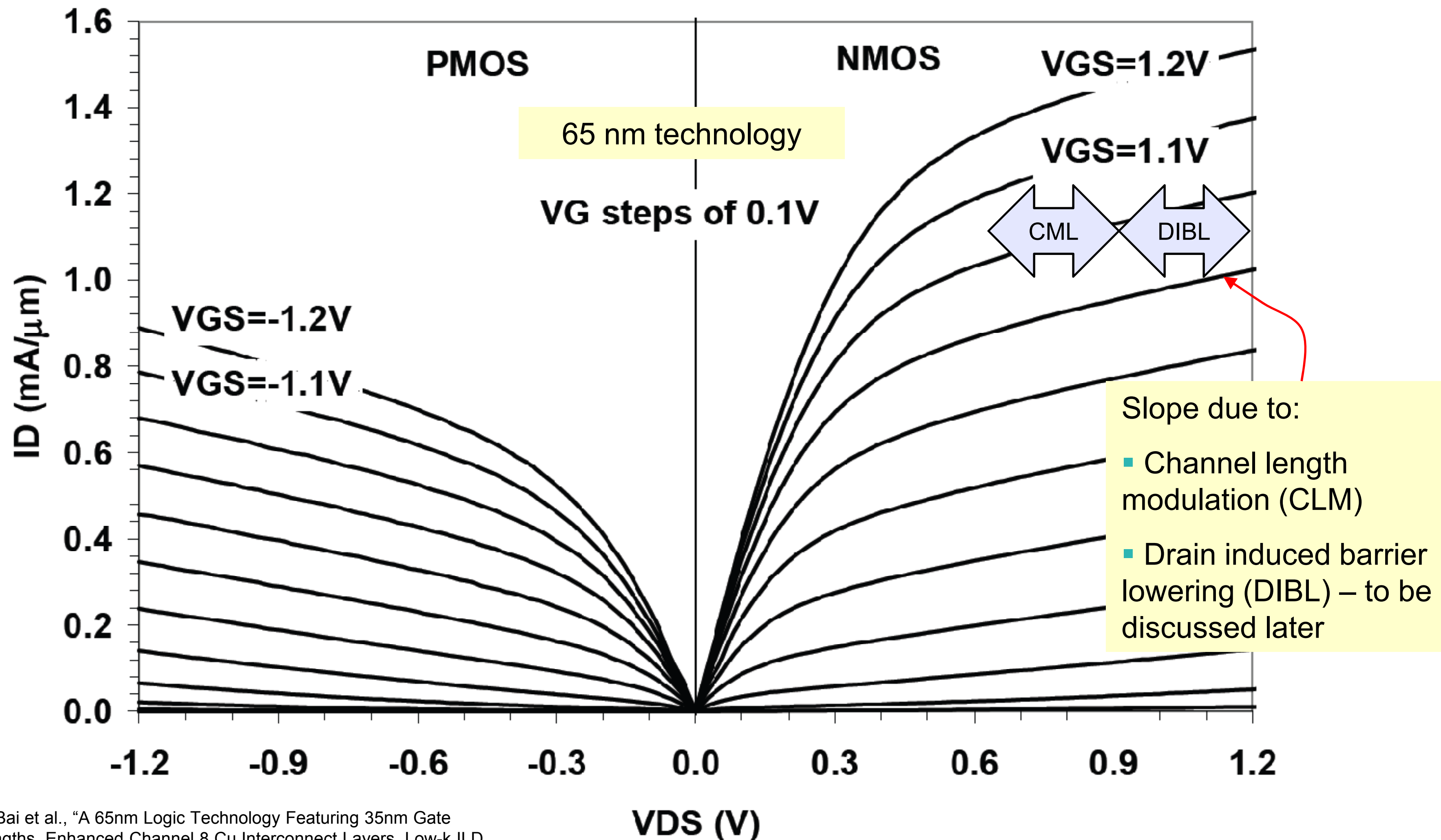
(b)



(c)

Channel length modulation

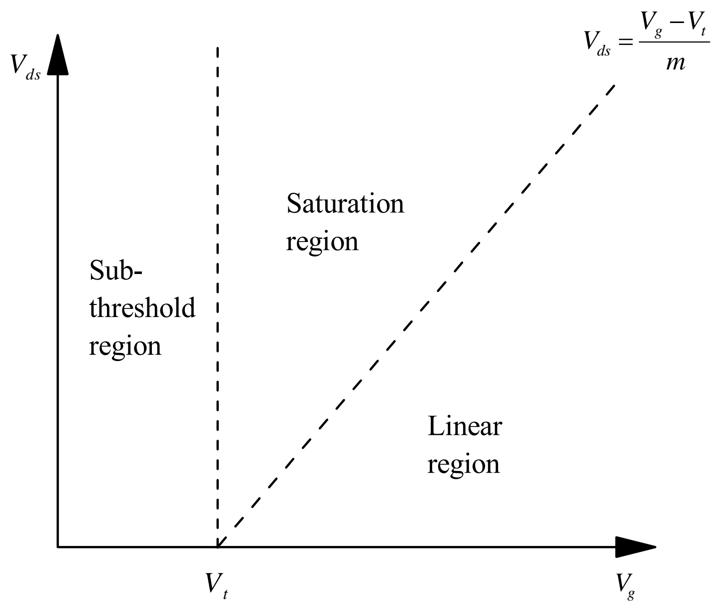
Saturation Characteristics – Experimental Example

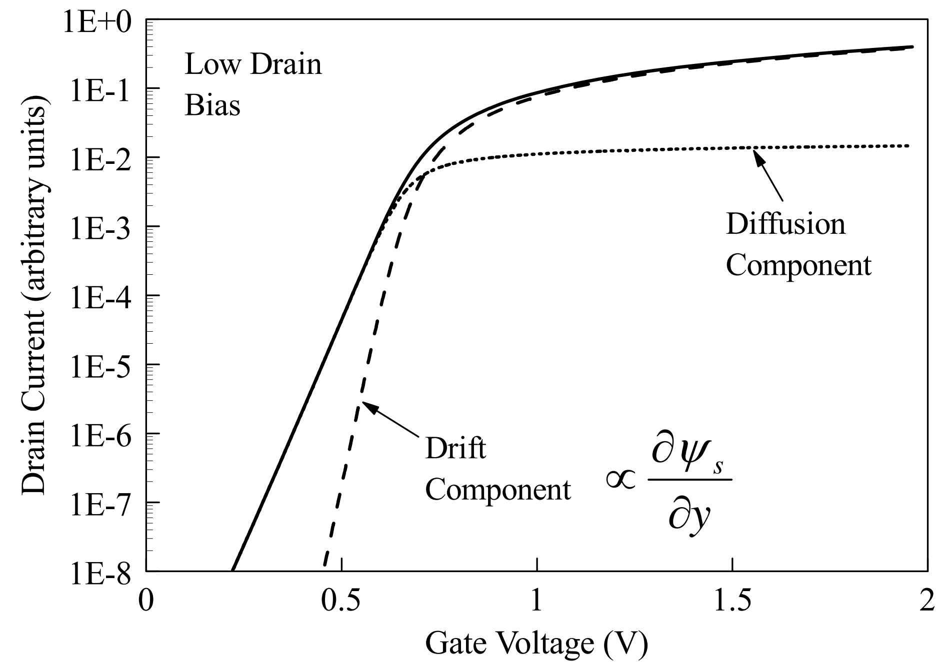


P. Bai et al., "A 65nm Logic Technology Featuring 35nm Gate Lengths, Enhanced Channel 8 Cu Interconnect Layers, Low-k ILD and 0.57 μm^2 SRAM Cell," *IEDM,* p. 657 (2004).

Subthreshold Region

$$J_n = q\mu_n nE + qD_n \frac{dn}{dy}$$





Subthreshold Currents

$$-Q_s = \varepsilon_{si} E_s = \sqrt{2\varepsilon_{si} kTN_a}\left[\frac{q\psi_s}{kT} + \frac{n_i^2}{N_a^2} e^{q(\psi_s - V)/kT}\right]^{1/2}$$

Power series expansion: 1st term Q_d, 2nd term Q_i,

$$-Q_i = \sqrt{\frac{\varepsilon_{si} q N_a}{2\psi_s}}\left(\frac{kT}{q}\right)\left(\frac{n_i}{N_a}\right)^2 e^{q(\psi_s - V)/kT}$$

$$I_{ds} = \mu_{eff}\frac{W}{L}\int_0^{V_{ds}} (-Q_i(V))dV$$

$$\Rightarrow I_{ds} = \mu_{eff}\frac{W}{L}\sqrt{\frac{\varepsilon_{si} q N_a}{2\psi_s}}\left(\frac{kT}{q}\right)^2\left(\frac{n_i}{N_a}\right)^2 e^{q\psi_s/kT}\left(1 - e^{-qV_{ds}/kT}\right)$$

Solving for ψ_s using

$$V_g = V_{fb} + \psi_s + \frac{\sqrt{2\varepsilon_{si} kTN_a}}{C_{ox}}\left[\frac{q\psi_s}{kT}\right]^{1/2}$$

or, $$I_{ds} = \mu_{eff} C_{ox}\frac{W}{L}(m-1)\left(\frac{kT}{q}\right)^2 e^{q(V_g - V_t)/mkT}\left(1 - e^{-qV_{ds}/kT}\right)$$

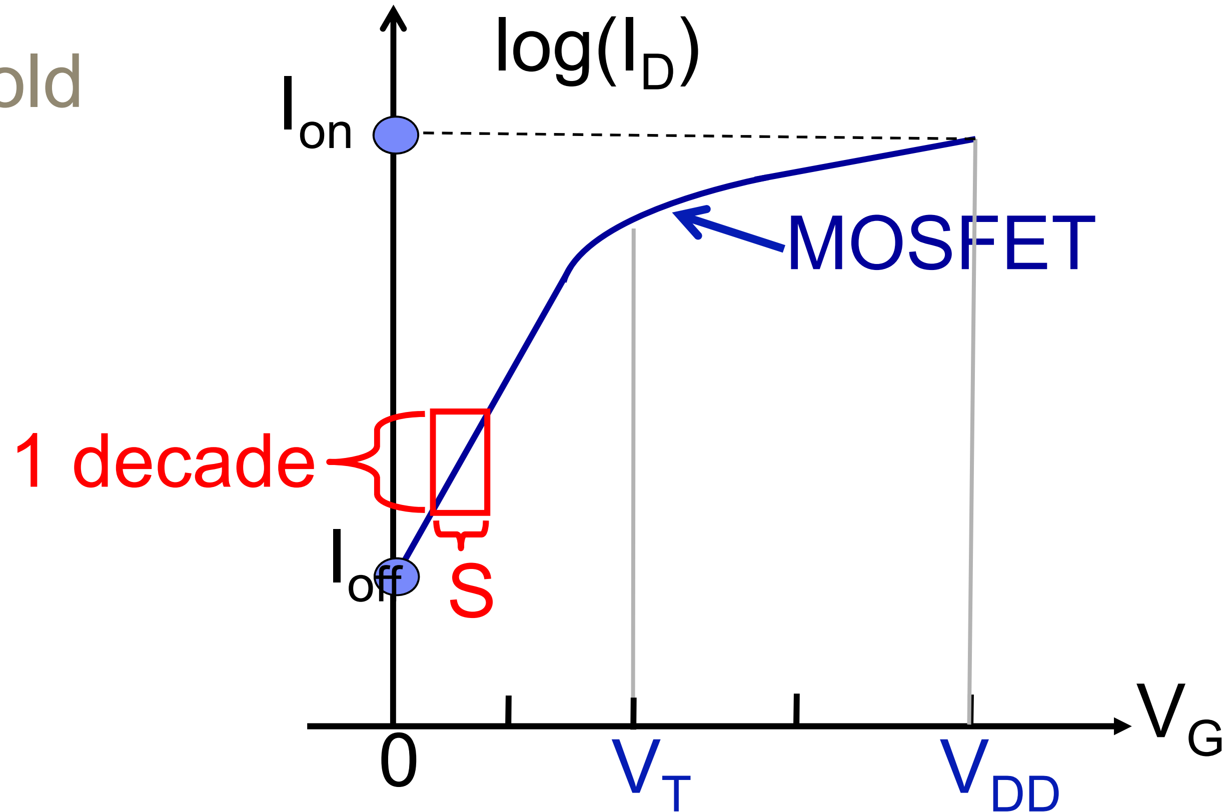
Inverse subthreshold slope:

$$S = \left(\frac{d(\log I_{ds})}{dV_g}\right)^{-1} = 2.3\frac{mkT}{q} = 2.3\frac{kT}{q}\left(1 + \frac{C_{dm}}{C_{ox}}\right)$$

$\ln(10)$

$(\ln(10)kT/q)^{-1} \approx 60\ mV/dec$ at 300K

Subthreshold Slope (S)



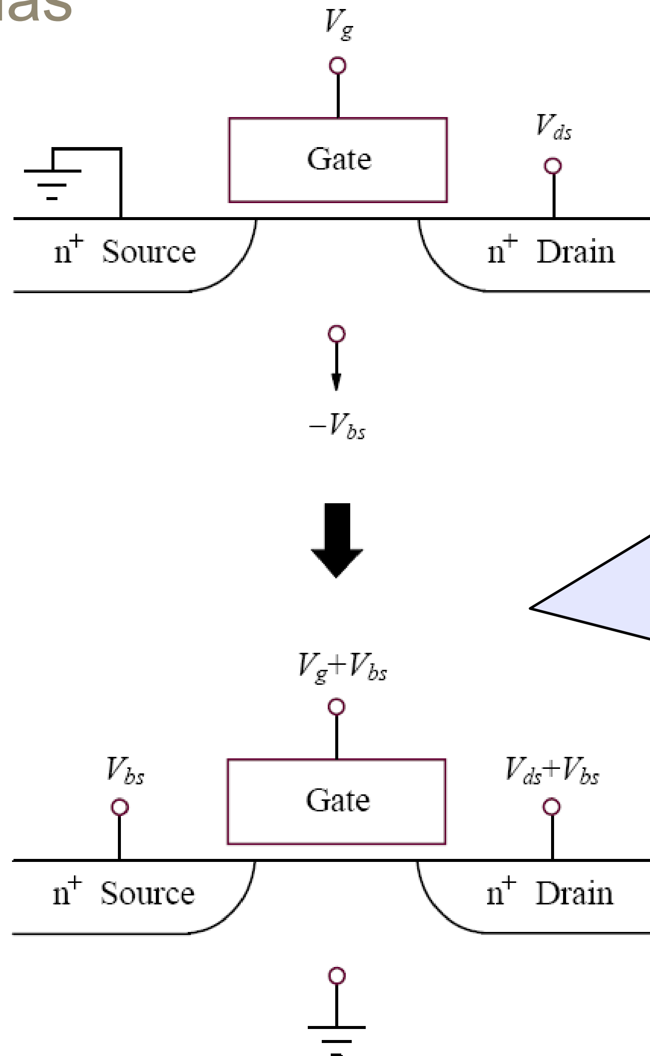
$$S = \left(\frac{d\log I_D}{dV_G}\right)^{-1} = \frac{\partial V_G}{\partial \Psi_S}\frac{\partial \Psi_S}{\partial \log I_D} = \left(1+\frac{C_{dm}}{C_{ox}}\right)\frac{kT}{q}\ln(10)$$

Gate to channel potential coupling: >1 in MOSFETs

60 mV/dec in MOSFETs due to Fermi-Dirac distribution

Questions?

Body Effect: Dependence of Threshold Voltage on Substrate Bias



You can either

1. Start with the Poisson's equation solution for Q_s, Q_d, Q_i with the quasi-Fermi levels of the holes and electrons separated by the substrate bias V_{bs}, or
2. Keep the substrate at zero (as the reference) and shift the source , drain, and gate biases by V_{bs}

Body Effect: Dependence of Threshold Voltage on Substrate Bias

You can either

1. Start with the Poisson's equation solution for Q_s, Q_d, Q_i with the quasi-Fermi levels of the holes and electrons separated by the substrate bias V_{bs}, or
2. Keep the substrate at zero (as the reference) and shift the source , drain, and gate biases by V_{bs}

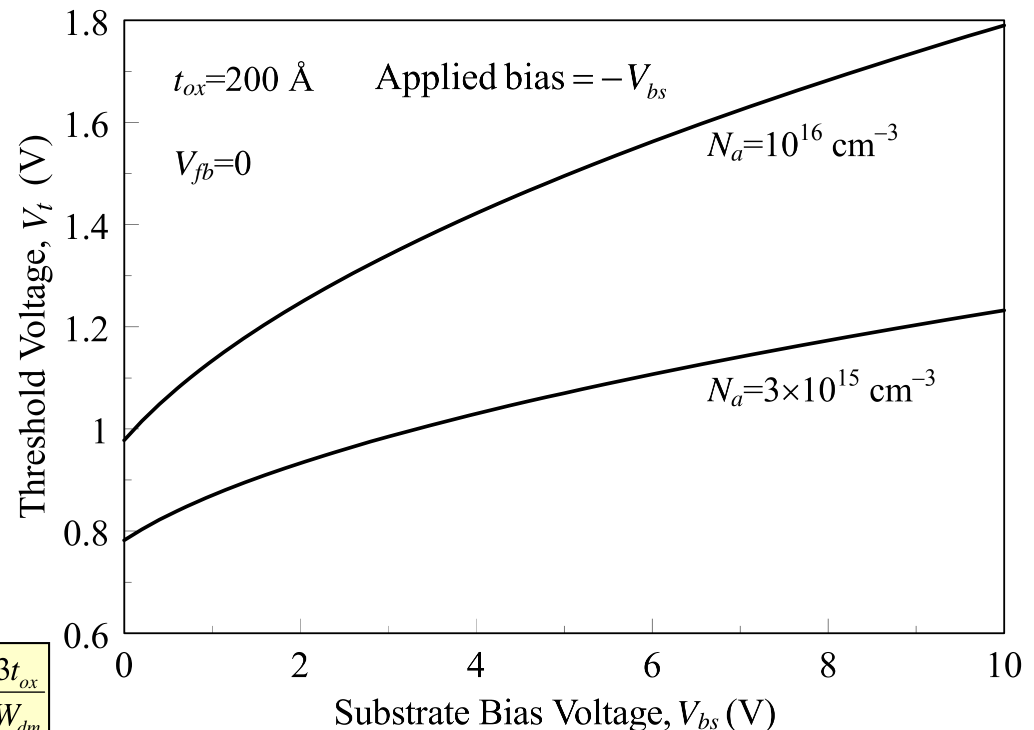
If $V_{bs} \neq 0$,

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L} \left\{ \left(V_g - V_{fb} - 2\psi_B - \frac{V_{ds}}{2} \right) V_{ds} - \frac{2\sqrt{2\varepsilon_{si} q N_a}}{3C_{ox}} \left[(2\psi_B + V_{bs} + V_{ds})^{3/2} - (2\psi_B + V_{bs})^{3/2} \right] \right\}$$

$$V_t = V_{fb} + 2\psi_B + \frac{\sqrt{2\varepsilon_{si} q N_a (2\psi_B + V_{bs})}}{C_{ox}}$$

$$\frac{dV_t}{dV_{bs}} = \frac{\sqrt{\varepsilon_{si} q N_a / 2(2\psi_B + V_{bs})}}{C_{ox}}$$

$$m = 1 + \frac{\sqrt{\varepsilon_{si} q N_a / 4\psi_B}}{C_{ox}} = 1 + \frac{C_{dm}}{C_{ox}} = 1 + \frac{3t_{ox}}{W_{dm}}$$



Body Biasing for Low Power

IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 40, NO. 7, JULY 2005 1549

Efficiency of Body Biasing in 90-nm CMOS for Low-Power Digital Circuits

Klaus von Arnim, Eduardo Borinski, Peter Seegebrecht, *Member, IEEE*, Horst Fiedler, *Member, IEEE*, Ralf Brederlow, Roland Thewes, *Member, IEEE*, Jörg Berthold, and Christian Pacha, *Member, IEEE*

***Abstract*—The efficiency of body biasing for leakage reduction and performance improvement in a 90-nm CMOS low-power technology with triple-well option is evaluated. Static measurements of single devices and dynamic measurements of ring oscillators and 32-b parallel prefix adders are presented. Whereas forward biasing still provides a significant performance improvement of up to 37% for low-leakage devices with 2.2-nm gate oxide thickness, the application of reverse biasing to reduce subthreshold leakage currents is inefficient due to additional leakage currents such as gate leakage and gate-induced drain leakage. Experimental results confirm that, in 90-nm CMOS circuits, the efficiency of body biasing strongly depends on the device type and operating temperature. Moreover, the impact of the zero-temperature coefficient point on static device and dynamic circuit performance is investigated.**

***Index Terms*—Body biasing, CMOS digital integrated circuits, zero-temperature coefficient point.**

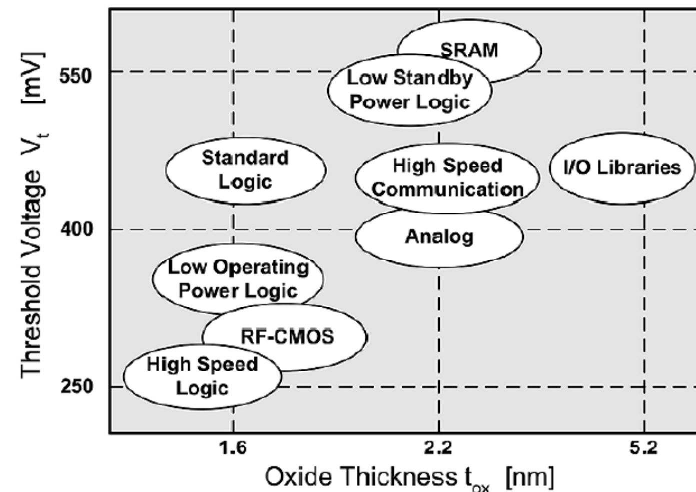


Fig. 1. Classifications of different circuit applications in the 90-nm CMOS system-on-chip technology.

I. Introduction

Application of Body Bias for Controlling Variations

1396 IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 37, NO. 11, NOVEMBER 2002

Adaptive Body Bias for Reducing Impacts of Die-to-Die and Within-Die Parameter Variations on Microprocessor Frequency and Leakage

James W. Tschanz, *Member, IEEE*, James T. Kao, *Member, IEEE*, Siva G. Narendra, *Member, IEEE*, Raj Nair, *Member, IEEE*, Dimitri A. Antoniadis, *Fellow, IEEE*, Anantha P. Chandrakasan, *Senior Member, IEEE*, and Vivek De, *Member, IEEE*

***Abstract*—Bidirectional adaptive body bias (ABB) is used to compensate for die-to-die parameter variations by applying an optimum pMOS and nMOS body bias voltage to each die which maximizes the die frequency subject to a power constraint. Measurements on a 150-nm CMOS testchip which incorporates on-chip ABB, show that ABB reduces variation in die frequency by a factor of seven, while improving the die acceptance rate. An enhancement of this technique, that compensates for within-die parameter variations as well, increases the number of dies accepted in the highest frequency bin. ABB is therefore shown to provide bin split improvement in the presence of increasing process parameter variations.**

***Index Terms*—Body bias, CMOS digital integrated circuits, forward bias, low-power circuits, microprocessors, parameter variations, substrate bias, within-die variation.**

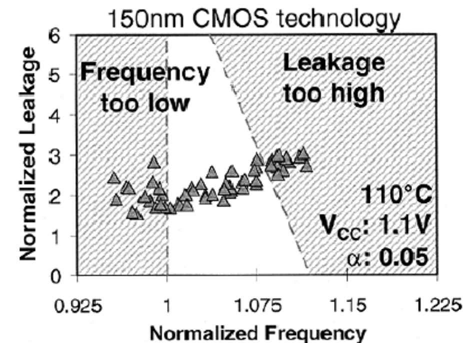


Fig. 1. Measured leakage power and frequency for 62 dies.

cannot be accepted because of either low operating frequency

Dependence of Threshold Voltage on Temperature

$$V_t = V_{fb} + 2\psi_B + \frac{\sqrt{4\varepsilon_{si} q N_a \psi_B}}{C_{ox}}$$

For n+ poly gated nMOSFET, $V_{fb} = -(E_g/2q) - \psi_B$

$$\Rightarrow \quad V_t = -\frac{E_g}{2q} + \psi_B + \frac{\sqrt{4\varepsilon_{si} q N_a \psi_B}}{C_{ox}}$$

$$\frac{dV_t}{dT} = -\frac{1}{2q}\frac{dE_g}{dT} + \left(1 + \frac{\sqrt{\varepsilon_{si} q N_a / \psi_B}}{C_{ox}}\right)\frac{d\psi_B}{dT} = -\frac{1}{2q}\frac{dE_g}{dT} + (2m-1)\frac{d\psi_B}{dT}$$

$$\Rightarrow \quad \frac{dV_t}{dT} = -(2m-1)\frac{k}{q}\left[\ln\left(\frac{\sqrt{N_c N_v}}{N_a}\right) + \frac{3}{2}\right] + \frac{m-1}{q}\frac{dE_g}{dT}$$

See Taur & Ning p. 167-168 for derivation steps

From Table 2.1, $dE_g/dT \approx -2.7\times10^{-4}$ eV/K and $(N_c N_v)^{1/2} \approx 2.4\times10^{19}$ cm^{-3}.

For $N_a \sim 10^{16}$ cm^{-3} and $m \sim 1.1$,
dV_t/dT is typically –1 mV/K.

Note: Operating temperature is specified at 85 °C for microprocessors and 150 °C for automotive applications

Carrier Transport and Gate Capacitance

Linear Region:

$$I_{ds} = \mu_{eff} C_{ox} \frac{W}{L} (V_g - V_t) V_{ds}$$

Saturation Region:

$$I_{ds} = I_{dsat} = \mu_{eff} C_{ox} \frac{W}{L} \frac{(V_g - V_t)^2}{2m}$$

Will come back to a more elaborate discussion later in the course about carrier transport.

Let's first digress briefly about the gate capacitance C_{ox} and the effective mobility μ_{eff} right now (we will return to them later again)

MOSFET Channel Mobility

$$\mu_{eff} = \frac{\int_0^{x_i} \mu_n n(x)dx}{\int_0^{x_i} n(x)dx}$$

Weighted average with inversion carrier density

It was empirically found that when μ_{eff} is plotted against an effective normal field E_{eff}, there exists a "universal relationship" independent of the substrate bias, doping concentration, and gate oxide thickness (Sabnis and Clemens, IEDM 1979).

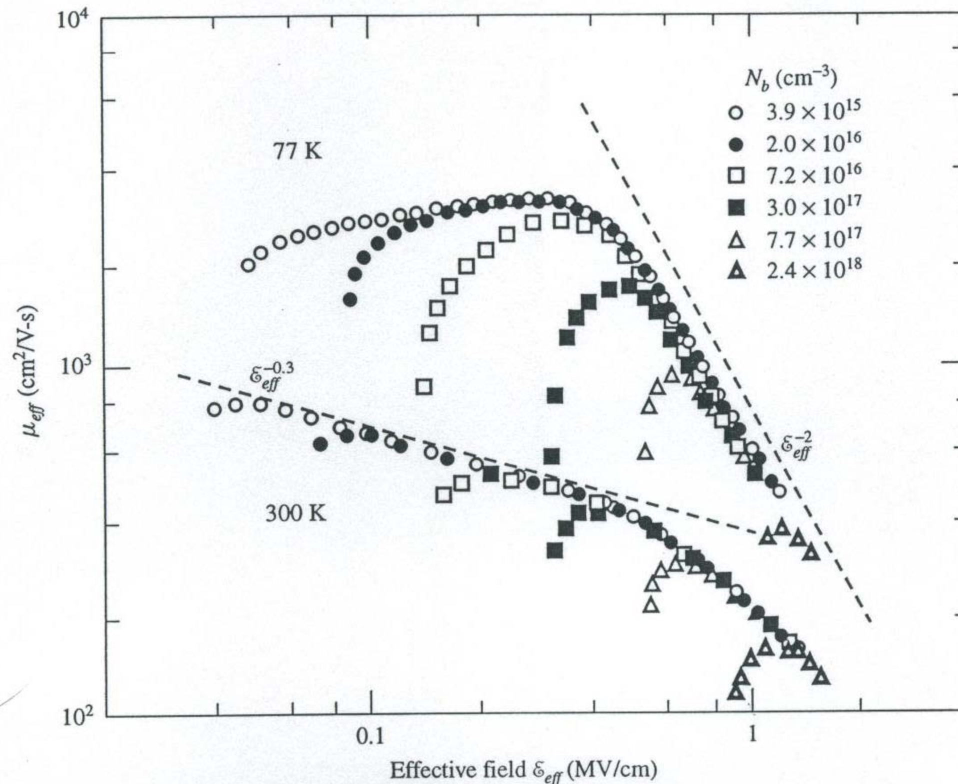
Here
$$E_{eff} = \frac{1}{\varepsilon_{si}}\left(|Q_d| + \frac{1}{2}|Q_i|\right)$$

Since $|Q_d| = \sqrt{4\varepsilon_{si} q N_a \psi_B} = C_{ox}(V_t - V_{fb} - 2\psi_B)$ and $|Q_i| \approx C_{ox}(V_g - V_t)$,

$$\Rightarrow \quad E_{eff} = \frac{V_t - V_{fb} - 2\psi_B}{3t_{ox}} + \frac{V_g - V_t}{6t_{ox}}$$

For n⁺ poly gated nMOSFET,
$$E_{eff} = \frac{V_t + 0.2}{3t_{ox}} + \frac{V_g - V_t}{6t_{ox}}$$

Electron Mobility



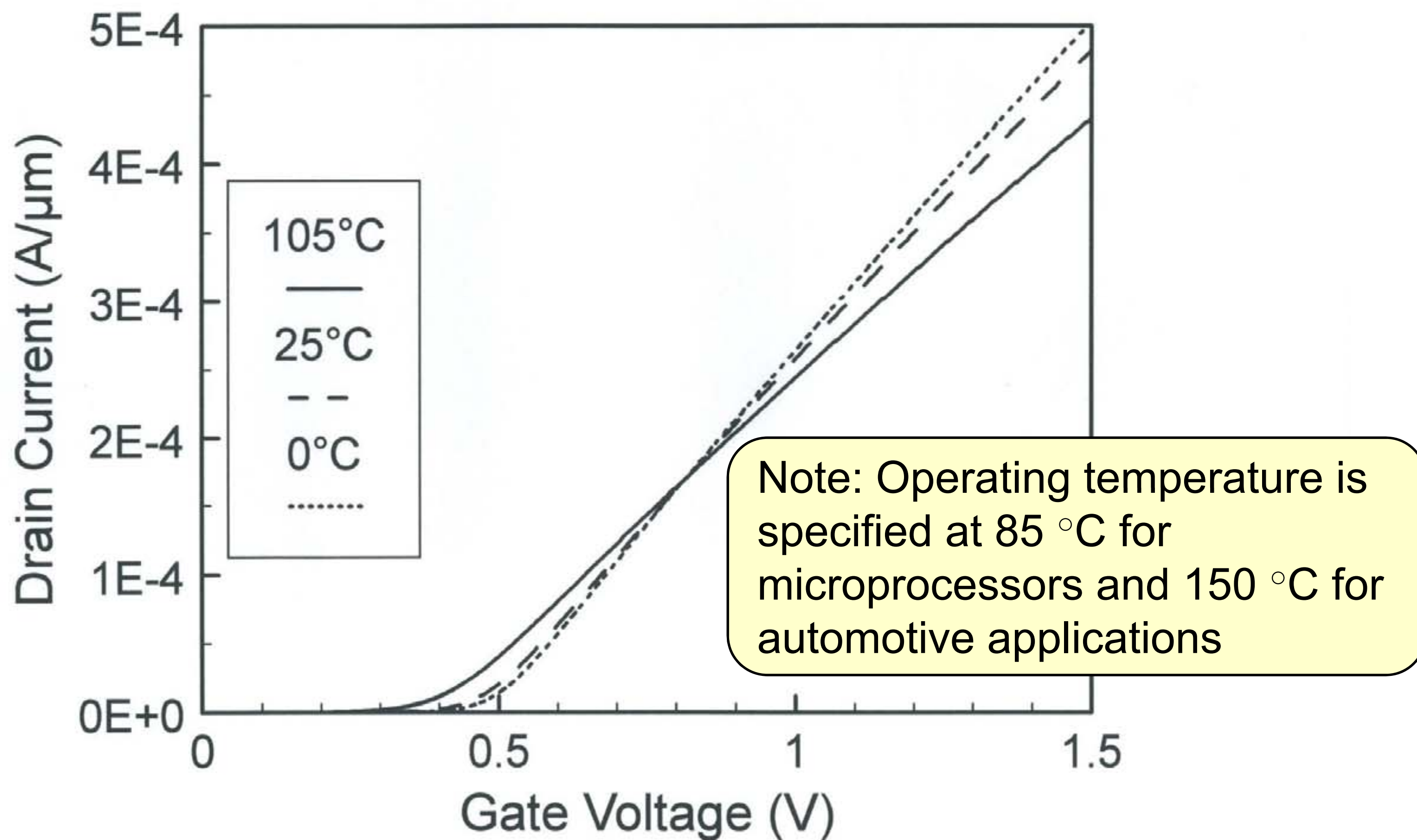
- Low field region (low electron density): Limited by impurity or Coulomb scattering (screened at high electron densities).

- Intermediate field region: Limited by phonon scattering,

$$\mu_{eff} \approx 32500 \times \mathrm{E}^{-1/3}$$

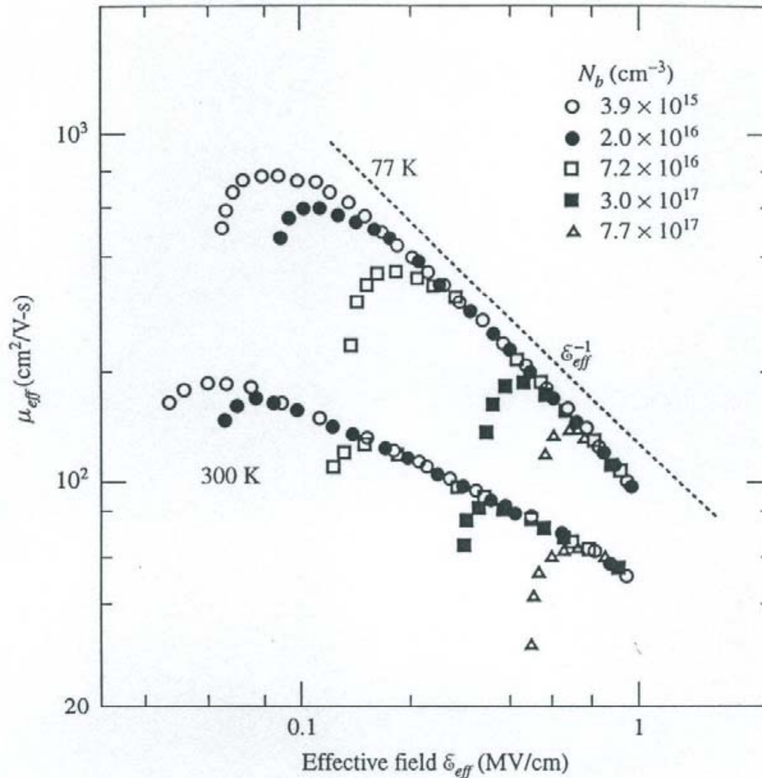
- High field region (> 1 MV/cm): Limited by surface roughness scattering (less temp. dependence).

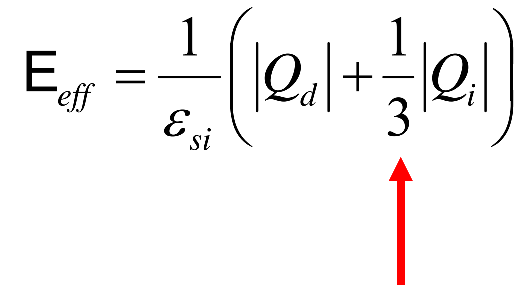
Temperature Dependence of MOSFET Current



Note: Operating temperature is specified at 85 °C for microprocessors and 150 °C for automotive applications

Hole Mobility





$$E_{eff} = \frac{1}{\varepsilon_{si}}\left(\left|Q_d\right| + \frac{1}{3}\left|Q_i\right|\right)$$

In general, pMOSFET mobility does not exhibit "universal" behavior as well as nMOSFET.

Intrinsic MOSFET Capacitance

- Subthreshold region: $C_g = WL\left(\frac{1}{C_{ox}} + \frac{1}{C_d}\right)^{-1} \approx WLC_d$
- Linear region: $C_g = WLC_{ox}$
- Saturation region:

$$Q_i(y) = -C_{ox}(V_g - V_t)\sqrt{1 - \frac{y}{L}}$$

$$\Rightarrow \quad C_g = \frac{2}{3}WLC_{ox}$$

Gate to body capacitance

Gate to channel capacitance

What is the gate to source and gate to drain capacitance in subthreshold?

What is the gate to body capacitance in the linear region?

See Taur & Ning p. 131 for derivation steps