Lecture 30 - The "Short" Metal-Oxide-Semiconductor Field-Effect Transistor

November 15, 2002

Contents:

1. Short-channel effects

Reading assignment:

P. K. Ko, "*Approaches to Scaling.*"

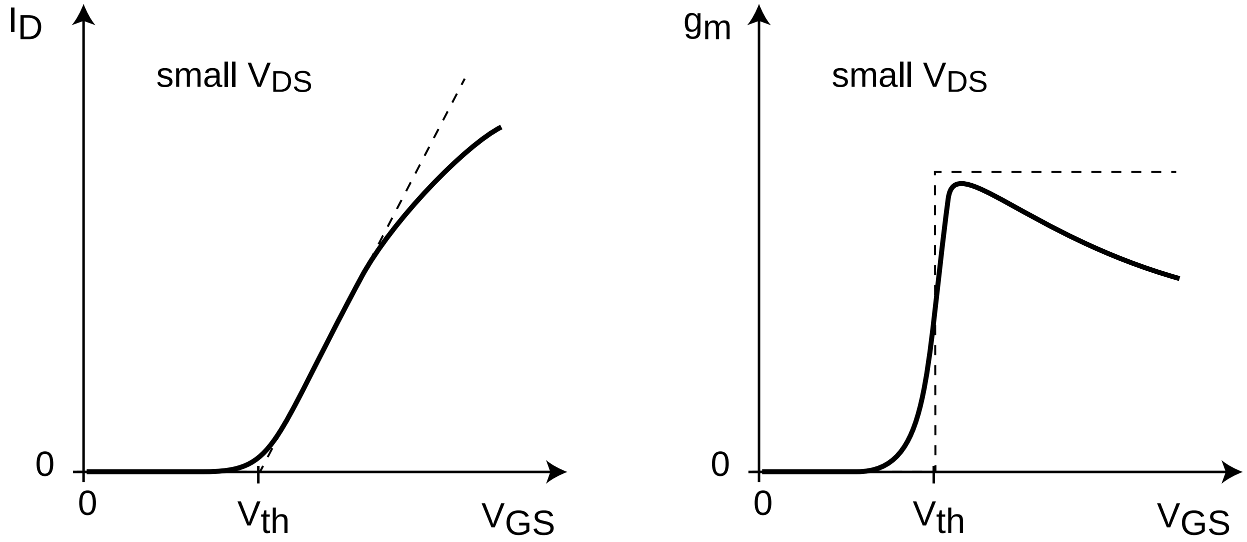
Key questions

- Why does it seem that in practice the drain current is significantly smaller than predicted by simple long MOSFET models?
- What is the impact of velocity saturation in the device characteristics?

1. Short-channel effects

□ **Mobility degradation**: mobility dependence on $\mathcal{E}_x$ (vertical field).

Experimental observation in linear regime:



Experimental observation:

$$\mu_{eff} = \frac{\mu_o}{1 + |\frac{\mathcal{E}_{av}}{\mathcal{E}_o}|^{\nu}}$$

where $\mathcal{E}_{av}$ is *average normal field in inversion layer*:

$$\mathcal{E}_{av} = \frac{Q_{dmax} + \frac{1}{2}Q_i}{\epsilon_s}$$

Due to *semiconductor-oxide interface roughness.*

μ_{eff} vs. $\mathcal{E}_{av}$: <u>universal relationship</u> for many MOSFET designs:

Fig. 4. Measured universal µeff vs. Eeff curves for electrons and holes in the inversion layer. (After Liang et al. [20].)

Parameters for the effective mobility models for electrons and holes:

	electrons	holes
μ_o $(cm^2/V \cdot s)$	670	160
$\mathcal{E}_o$ (MV/cm)	0.67	0.7
ν	1.6	1

Figure 6 graph from Inversion layer electron mobility data, graph from: Arora and Richardson, P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

Simplified expression of $\mathcal{E}_{av}$ for n$^+$-polySi gate:

1) Relationship between Q_{dmax} and V_{th}:

$$V_{th} = V_{FB} + \phi_{sth} - \frac{Q_{dmax}}{C_{ox}}$$

with:

$$V_{FB} = -\phi_{bi} = \frac{1}{q}(W_M - W_S) = \frac{1}{q}(W_M - \chi_s - \frac{E_g}{2} - q\phi_f) = -\frac{1}{q}\frac{E_g}{2} - \phi_f$$

Plug into V_{th} and solve for Q_{dmax}:

$$Q_{dmax} = -C_{ox}(V_{th} + \frac{E_g}{2q} + \phi_f - 2\phi_f) = -C_{ox}(V_{th} + \frac{E_g}{2q} - \phi_f) \simeq -C_{ox}V_{th}$$

2) Relationship between Q_i and $V_{GS} - V_{th}$:

$$Q_i = -C_{ox}(V_{GS} - V_{th})$$

3) Plug Q_{dmax} and Q_i into $\mathcal{E}_{av}$:

$$|\mathcal{E}_{av}| \simeq |\frac{Q_{dmax} + \frac{1}{2}Q_i}{\epsilon_s}| = \frac{C_{ox}V_{th} + \frac{1}{2}C_{ox}(V_{GS} - V_{th})}{\epsilon_s} = \frac{\epsilon_{ox}}{\epsilon_s}\frac{V_{GS} + V_{th}}{2x_{ox}}$$

$$|\mathcal{E}_{av}| \simeq \frac{\epsilon_{ox}}{\epsilon_s}\frac{V_{GS}+V_{th}}{2x_{ox}}$$

Key dependences:

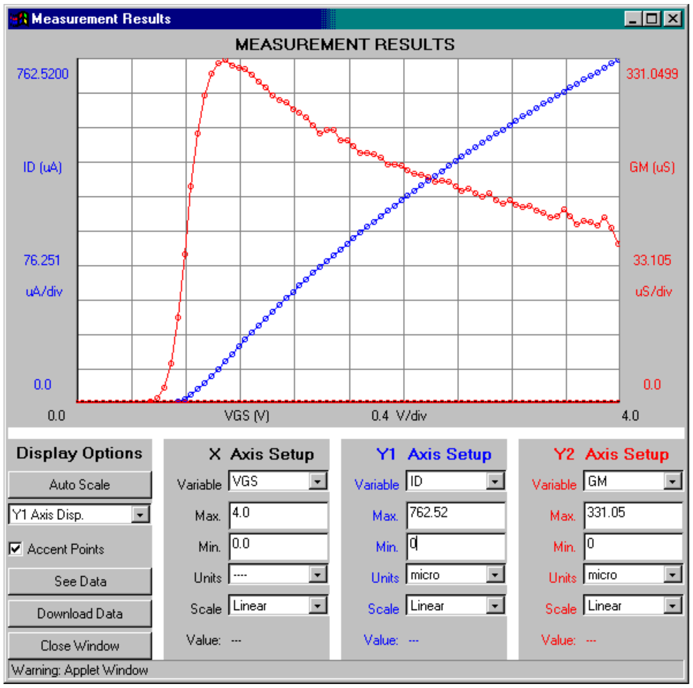
- $V_{GS} \uparrow \Rightarrow |\mathcal{E}_{av}| \uparrow \Rightarrow \mu_{eff} \downarrow$
- $V_{th} \uparrow \Rightarrow |\mathcal{E}_{av}| \uparrow \Rightarrow \mu_{eff} \downarrow$
- $x_{ox} \downarrow \Rightarrow |\mathcal{E}_{av}| \uparrow \Rightarrow \mu_{eff} \downarrow$

Fig. 5. Calculated μeff of current carrier, graph. From: VLSI Electronics P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

Several comments:

- Since $I_D \sim \mu_e$, mobility degradation more severe as V_{GS} increases $\Rightarrow I_D$ won't rise as fast with V_{GS}.
- Since μ_e depends on $|\mathcal{E}_{av}| \Rightarrow \mu_{eff}$ depends on y. Disregard to first order $\Rightarrow$ use same μ_{eff} everywhere.
- Mobility degradation considered "short-channel effect" because as $L \downarrow \Rightarrow x_{ox} \downarrow$ and μ degradation becomes important.

g_m in linear regime ($V_{DS} = 0.1\ V$) for $L_g = 1.5\ \mu m$ MOSFET:

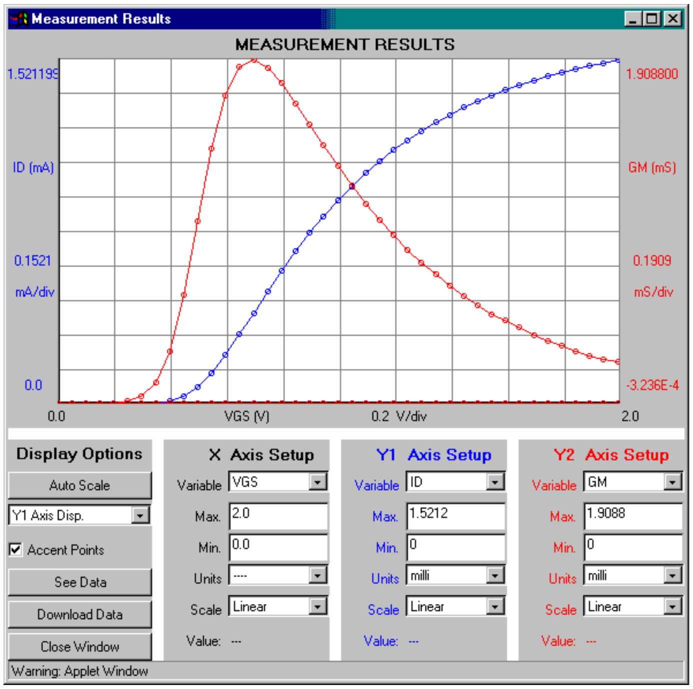


X Axis Setup
Variable: VGS
Max.: 4.0
Min.: 0.0
Units: ----
Scale: Linear
Value: ---

Y1 Axis Setup
Variable: ID
Max.: 762.52
Min.: 0
Units: micro
Scale: Linear
Value: ---

Y2 Axis Setup
Variable: GM
Max.: 331.05
Min.: 0
Units: micro
Scale: Linear
Value: ---

g_m in linear regime ($V_{DS} = 0.1\ V$) for $L_g = 0.18\ \mu m$ MOSFET:



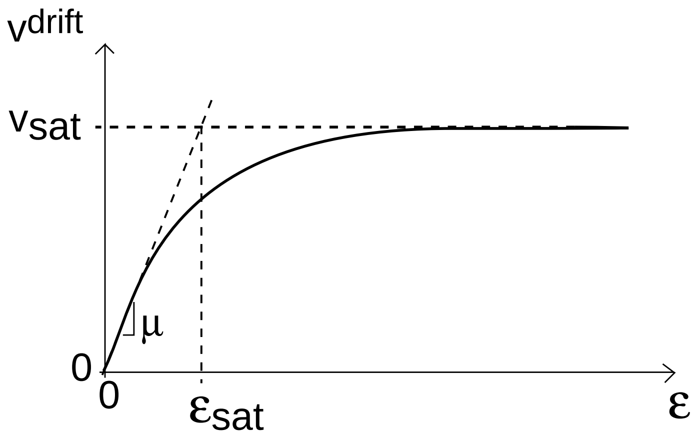
X Axis Setup
Variable: VGS
Max.: 2.0
Min.: 0.0
Units: ----
Scale: Linear
Value: ---

Y1 Axis Setup
Variable: ID
Max.: 1.5212
Min.: 0
Units: milli
Scale: Linear
Value: ---

Y2 Axis Setup
Variable: GM
Max.: 1.9088
Min.: 0
Units: milli
Scale: Linear
Value: ---

□ Velocity saturation

At high longitudinal fields, v_e cannot exceed v_{sat}:



Best fit to experiments:

$$v_e = \frac{\mu_e \mathcal{E}}{(1 + |\frac{\mu_e \mathcal{E}}{v_{sat}}|^n)^{1/n}}$$

For inversion layer:

	electrons	holes
v_{sat} (cm/s)	8×10^6 cm/s	6×10^6 cm/s
n	2	1

To develop analytical model, use $n = 1$:

$$v_e = \frac{\mu_e \mathcal{E}}{1 + |\frac{\mu_e \mathcal{E}}{v_{sat}}|}$$

New current model:

$$J_e = \mu_e Q_i \frac{dV}{dy} = \frac{\mu_e}{1+|\frac{\mathcal{E}}{\mathcal{E}_{sat}}|} Q_i \frac{dV}{dy}$$

with

$$\mathcal{E}_{sat} = \frac{v_{sat}}{\mu_e}$$

Rewrite current equation:

$$J_e = [\mu_e Q_i - \frac{J_e}{\mathcal{E}_{sat}}] \frac{dV}{dy}$$

Plug in fundamental charge relationship:

$$Q_i = -C_{ox}(V_{GS} - V - V_{th})$$

and integrate along channel:

$$J_e L = -\mu_e C_{ox} \int_0^{V_{DS}} (V_{GS} - V - V_{th}) dV - \frac{J_e}{\mathcal{E}_{sat}} V_{DS}$$

Solve for J_e:

$$J_e = -\frac{\mu_e C_{ox}}{L + \frac{V_{DS}}{\mathcal{E}_{sat}}} (V_{GS} - V_{th} - \frac{1}{2} V_{DS}) V_{DS}$$

Terminal drain current in linear regime:

$$I_D = \frac{W}{L} \frac{\mu_e C_{ox}}{1 + \frac{V_{DS}}{\mathcal{E}_{sat} L}} (V_{GS} - V_{th} - \frac{1}{2} V_{DS}) V_{DS}$$

Effectively, impact of velocity saturation:

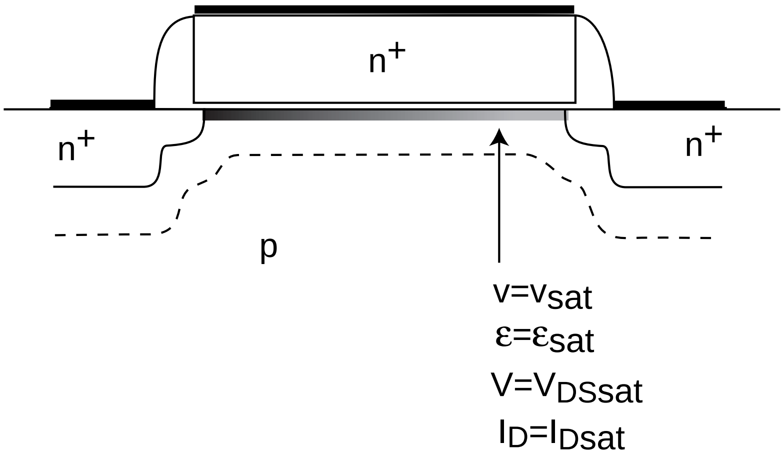
$$\mu_e \Rightarrow \frac{\mu_e}{1 + \frac{V_{DS}}{\mathcal{E}_{sat} L}}$$

with $\frac{V_{DS}}{L} \equiv$ *average longitudinal field.*

- For $\frac{V_{DS}}{L} \ll \mathcal{E}_{sat} \Rightarrow$ velocity saturation irrelevant (*mobility regime*).
- For $\frac{V_{DS}}{L} \gg \mathcal{E}_{sat} \Rightarrow$ velocity saturation prominent (*velocity saturation regime*).

Since $\mathcal{E}_{sat} \simeq 10^5\ V/cm$ and V_{DS} order $1 - 10\ V$, velocity saturation important if $L \sim 0.1 - 1\ \mu m$.

Current saturation occurs when v_{sat} reached anywhere in the channel $\Rightarrow I_D$ won't increase anymore with V_{DS}:



Bottleneck is current flowing through v_{sat} region:

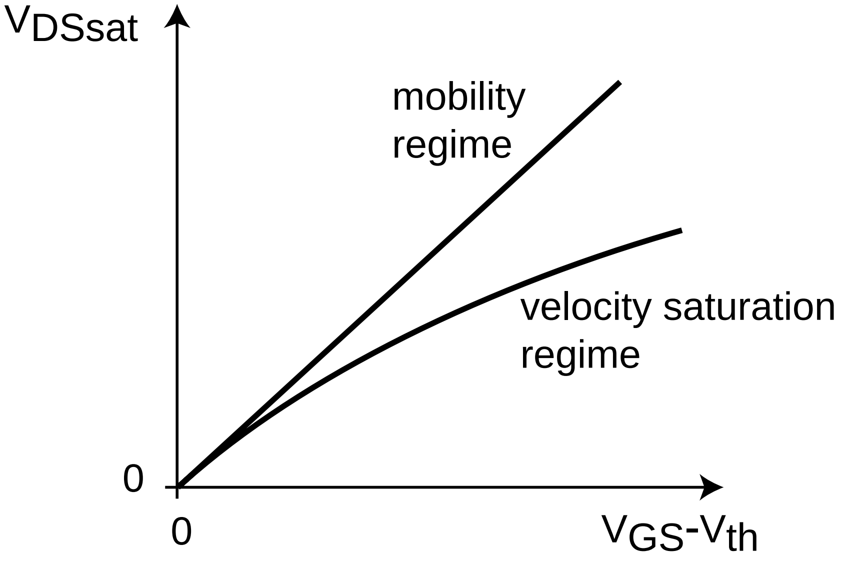
$$\begin{aligned} I_{Dsat} &= -Wv_eQ_i \\ &= Wv_{sat}C_{ox}(V_{GS}-V_{th}-V_{DSsat}) \\ &= \frac{W}{L}\frac{\mu_e C_{ox}}{1+\frac{V_{DSsat}}{\mathcal{E}_{sat}L}}(V_{GS}-V_{th}-\frac{1}{2}V_{DSsat})V_{DSsat} \end{aligned}$$

Solve for V_{DSsat}:

$$V_{DSsat} = \mathcal{E}_{sat}L(\sqrt{1+2\frac{V_{GS}-V_{th}}{\mathcal{E}_{sat}L}}-1)$$

$$V_{DSsat} = \mathcal{E}_{sat}L(\sqrt{1 + 2\frac{V_{GS} - V_{th}}{\mathcal{E}_{sat}L}} - 1)$$

- For long L, $V_{DSsat} \simeq V_{GS} - V_{th}$
- For short L, $V_{DSsat} \simeq \sqrt{2\mathcal{E}_{sat}L(V_{GS} - V_{th})} < V_{GS} - V_{th}$



Velocity saturation results in *premature current saturation* and less current:

$$I_{Dsat} = Wv_{sat}C_{ox}(V_{GS} - V_{th} - V_{DSsat})$$

Experiments:

Fig. 7. The saturation voltage for several channel lengths, graph. From: VLSI P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

Current-voltage characteristics:

Fig. 15 Comparison of drain characteristics for constant mobility case and P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

Impact of velocity saturation on transconductance:

$$g_m = \frac{\partial I_{Dsat}}{\partial V_{GS}} = W v_{sat} C_{ox} (1 - \frac{\partial V_{DSsat}}{\partial V_{GS}})$$

with

$$\frac{\partial V_{DSsat}}{\partial V_{GS}} = \frac{1}{\sqrt{1 + 2\frac{V_{GS}-V_{th}}{\mathcal{E}_{sat} L}}}$$

Then:

$$g_m = W v_{sat} C_{ox} (1 - \frac{1}{\sqrt{1 + 2\frac{V_{GS}-V_{th}}{\mathcal{E}_{sat} L}}})$$

In the limit of short L:

$$g_m = W v_{sat} C_{ox}$$

In the limit of short L, g_m determined only by x_{ox}.

Fig. 10: Deep submicon NMOS transistors, graph. From: VLSI Electronics P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

theoretical
limit for L=0

Fig. 9: Measured and calculated gm, graph. From: VLSI Electronics Microstructure P. Ko, "Approaches to Scaling," in VLSI Electronics: Microstructure Science, vol. 18, chapter 1, pp. 1--37, Academic Press, 1989.

L=1 μm, x_{ox}=20 nm

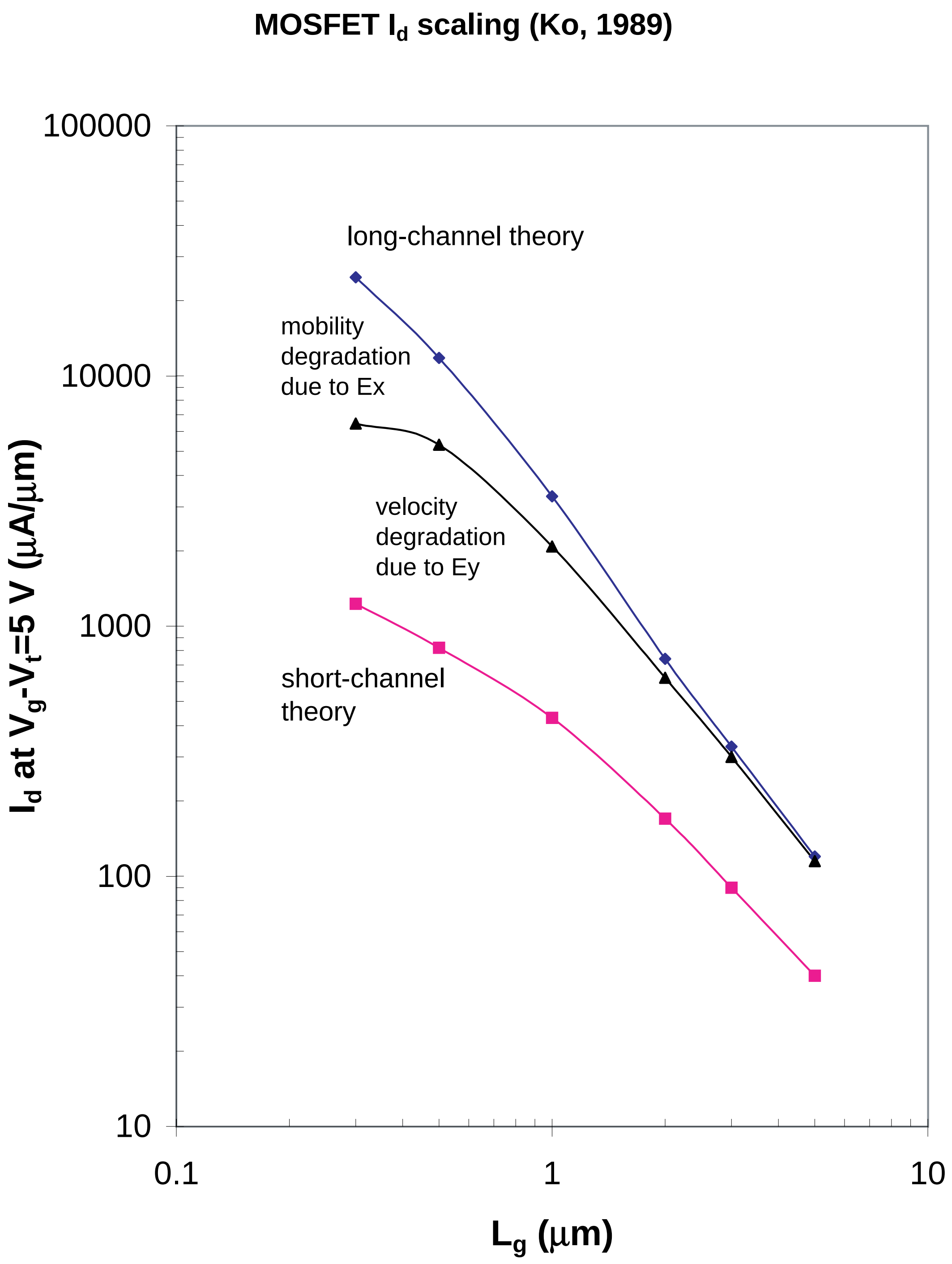
L=2 μm, x_{ox}=30 nm

L=3 μm, x_{ox}=50 nm

L=1.3 μm, x_{ox}=20 nm

L=2.3 μm, x_{ox}=30 nm

L=3.3 μm, x_{ox}=50 nm


MOSFET I_d scaling (Ko, 1989)
I_d at V_g-V_t=5 V (µA/µm)
100000
10000
1000
100
10
0.1
1
10
L_g (µm)
long-channel theory
mobility degradation due to Ex
velocity degradation due to Ey
short-channel theory

Key conclusions

- Inversion layer mobility degraded by transveral field due to roughness of semiconductor-oxide interface ⇒ I_D lower than predicted by simple models.
- There is a *universal relationship between mobility and* average *transversal field in inversion layer.*
- For short L, velocity saturation in inversion layer important ⇒ I_D lower than predicted by simple models.
- Velocity saturation ⇒ premature MOSFET saturation ⇒ V_{DSsat} lower than predicted by simple models.
- Velocity saturation ⇒ g_m saturation in limit of short L:

$$g_m = W v_{sat} C_{ox}$$