Recitation 11: Small Signal Model of MOSFET/MOSFET in Digital Circuits

Small Signal Models

On Tuesday we talked about *small signal models of MOSFETs*. Why do we need small signal modeling?

To linearize circuits. Linear circuits are much easier to work with: we can use Thevenin/Norton equivalent circuits, superposition, etc.

How to obtain a linearized circuit?

If we limit our signals to a relatively small amplitude, (this is in fact most often the case), the non-linear IV curves (as seen in 1 for example) can be considered piece-wise linear $\Longrightarrow$ small signal model.

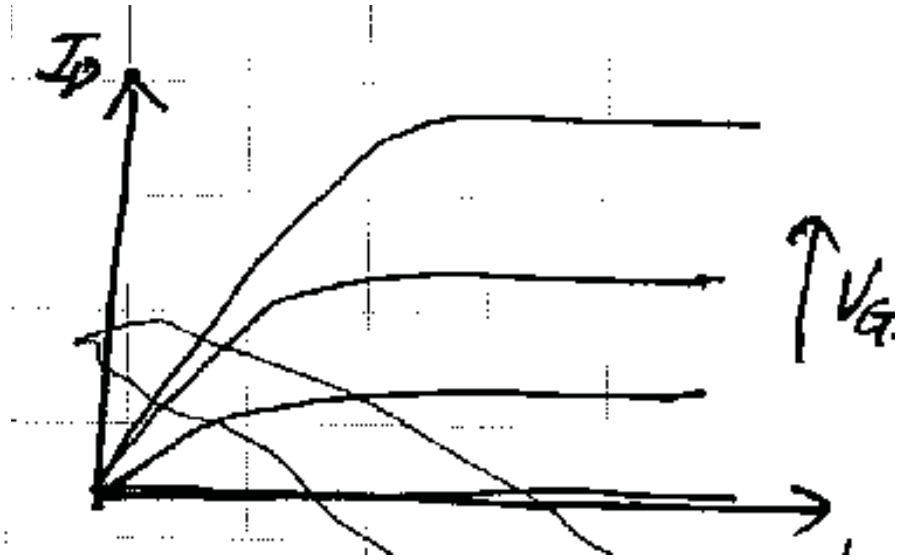


Figure 1: Non Linear IV Curves

A MOSFET is a 4-terminal device (NMOS as an example)

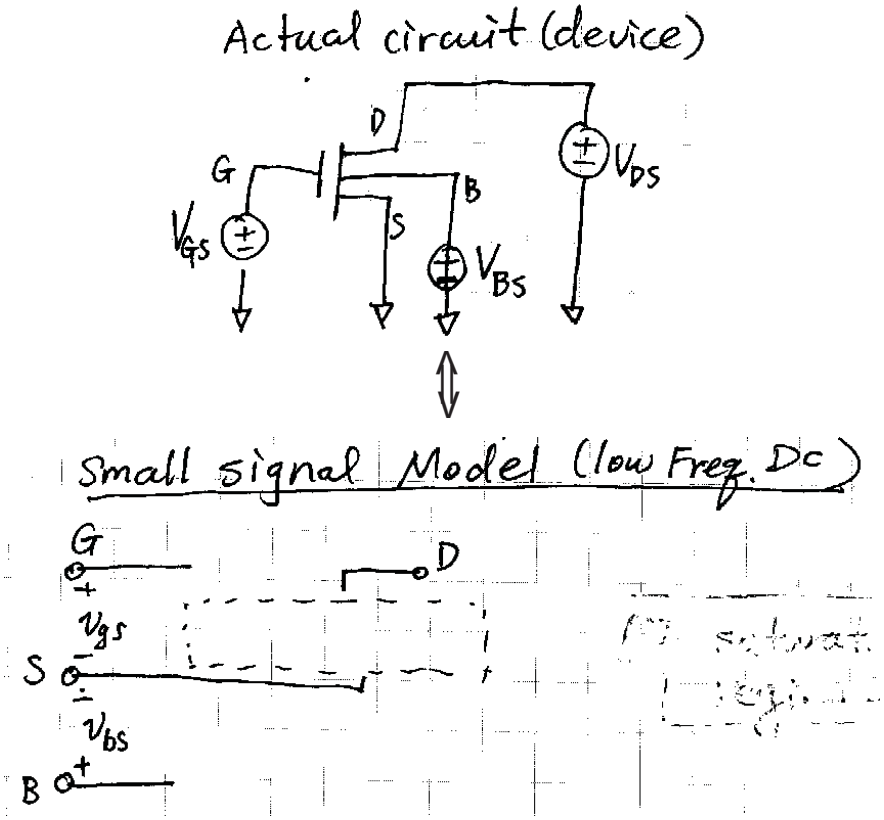


Figure 2: NMOS

- cut-off $i_{\mathrm{d}} = 0$
- triode/linear

$$i_{\mathrm{d}} = \frac{w}{L}\mu_{\mathrm{n}}C_{\mathrm{ox}}(V_{\mathrm{GS}} - \frac{V_{\mathrm{DS}}}{2} - V_{\mathrm{T}}) \cdot V_{\mathrm{DS}}$$

- Saturation

$$i_{\mathrm{d}} = \frac{w}{2L}\mu_{\mathrm{n}}C_{\mathrm{ox}}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2 \cdot (1 + \lambda V_{\mathrm{DS}})$$

In small signal modeling it is very important to differentiate:

$$\begin{aligned} i_{\mathrm{D}} &\leftarrow \text{total signal } (i_{\mathrm{D}} = I_{\mathrm{D}} + i_{\mathrm{d}}) \\ I_{\mathrm{D}} &\leftarrow \text{DC signal} \\ i_{\mathrm{d}} &\leftarrow \text{small signal} \end{aligned}$$

What Happens at Low Frequency?

To obtain a small signal equivalent circuit, we need to find an operation point first: $Q(V_{\rm DS}, V_{\rm GS}, V_{\rm BS})$. ($Q$ is a specific point on Fig. 1) thus $I_{\rm D}$ is also known.
Then, what is in between D & S in Figure **??** depends on:

$$\begin{aligned} i_{\rm d} &= \frac{\delta i_{\rm d}}{\delta V_{\rm GS}}\Big|_{\rm Q} \cdot V_{\rm GS} + \frac{\delta i_{\rm d}}{\delta V_{\rm DS}}\Big|_{\rm Q} \cdot V_{\rm DS} + \frac{\delta i_{\rm d}}{\delta V_{\rm BS}}\Big|_{\rm Q} \cdot V_{\rm BS} \\ &= g_{\rm m} \cdot V_{\rm GS} + g_{\rm o} \cdot V_{\rm DS} + g_{\rm MB} \cdot V_{\rm BS} \quad \text{three conductances in parallel } (\because \text{they add up}) \end{aligned}$$

- $g_{\rm m}$: trans-conductance (unit S)
- $g_{\rm o}$: output conductance (unit S)
- $g_{\rm mb}$: backgate transconductance (unit S)

On Tuesday, we derived the expression for $g_{\rm m}, g_{\rm o}, g_{\rm mb}$ *in saturation regime*

$$\begin{aligned} g_{\rm m} &= \frac{\delta i_{\rm d}}{\delta V_{\rm GS}}\Big|_{\rm Q} = \frac{\delta[\frac{w}{2L}\mu_{\rm n}C_{\rm ox}(V_{\rm GS}-V_{\rm T})^2(1+\lambda V_{\rm DS})]}{\delta V_{\rm GS}}\Big|_{Q(V_{\rm GS},V_{\rm DS},V_{\rm BS})} \\ &= \frac{w}{2L}\mu_{\rm n}C_{\rm ox}(1+\lambda V_{\rm DS})(V_{\rm GS}-V_{\rm T})\cdot 2 \\ &\simeq \frac{w}{L}\mu_{\rm n}C_{\rm ox}(V_{\rm GS}-V_{\rm T}) = \sqrt{\frac{2w}{L}\mu_{\rm n}C_{\rm ox}I_{\rm D}} \to g_{\rm m} \propto \sqrt{\frac{w}{L}I_{\rm D}} \\ g_{\rm o} &= \frac{\delta i_{\rm D}}{\delta V_{\rm DS}}\Big|_{\rm Q} = \frac{w}{2L}\mu_{\rm n}C_{\rm ox}(V_{\rm GS}-V_{\rm T})^2\cdot\lambda \simeq \lambda I_{\rm D} \\ \gamma_{\rm o} &= \frac{1}{g_{\rm o}} = \frac{1}{\lambda I_{\rm D}} \propto \frac{L}{I_{\rm D}} \\ g_{\rm mb} &= \frac{\delta i_{\rm D}}{\delta V_{\rm BS}}\Big|_{\rm Q} = \frac{w}{2L}\mu_{\rm n}C_{\rm ox}(1+\lambda V_{\rm DS})\cdot(-2)(V_{\rm GS}-V_{\rm T})\cdot\frac{\delta V_{\rm T}}{\delta V_{\rm BS}}\Big|_{\rm Q} \\ &\simeq \underbrace{-\frac{w}{L}\mu_{\rm n}C_{\rm ox}(V_{\rm GS}-V_{\rm T})}_{g_{\rm m}}\frac{\delta V_{\rm T}}{\delta V_{\rm BS}}\Big|_{\rm Q} \\ &= +g_{\rm m}\cdot\gamma\cdot\frac{1}{2\sqrt{-2\phi_{\rm p}-V_{\rm BS}}} \\ V_{\rm T} &= V_{\rm T_o} + \gamma(\sqrt{-2\phi_{\rm p}-V_{\rm BS}} - \sqrt{-2\phi_{\rm p}}) \end{aligned}$$

Now what does the small signal circuit look like?

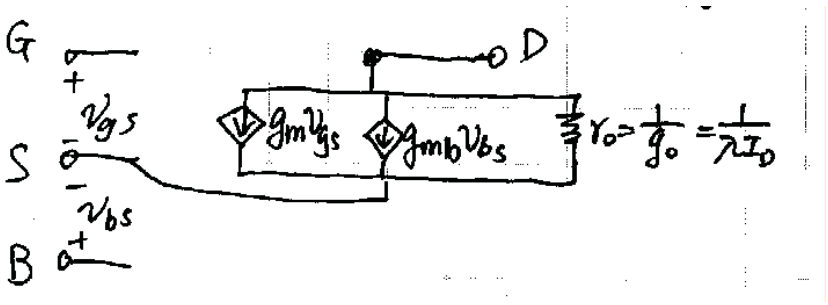


Figure 3: Low Frequency Small Signal Circuit. Two of them are voltage controlled current sources, one is a resistor. Why?

What Happens at High Frequency?

There are intrinsic or parasitic capacitances related to the MOSFET structure, as we know $Z_c = \frac{1}{jwC}$. At low frequency, Z_c is very large, can be approximated to open circuit, however at high frequency, Z_c is small enough we have to consider.
We have 4 terminals. Considering the possible combinations between them, we have:

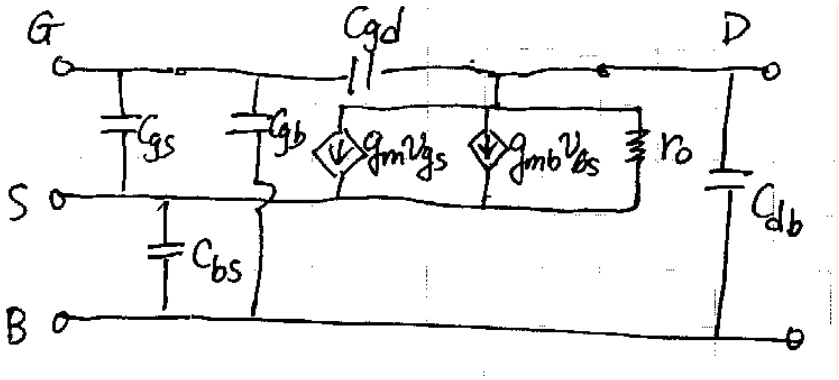


Figure 4: Between D & S, we have conduction channel, no capacitance between D & S

What are these capacitances (under saturation)?

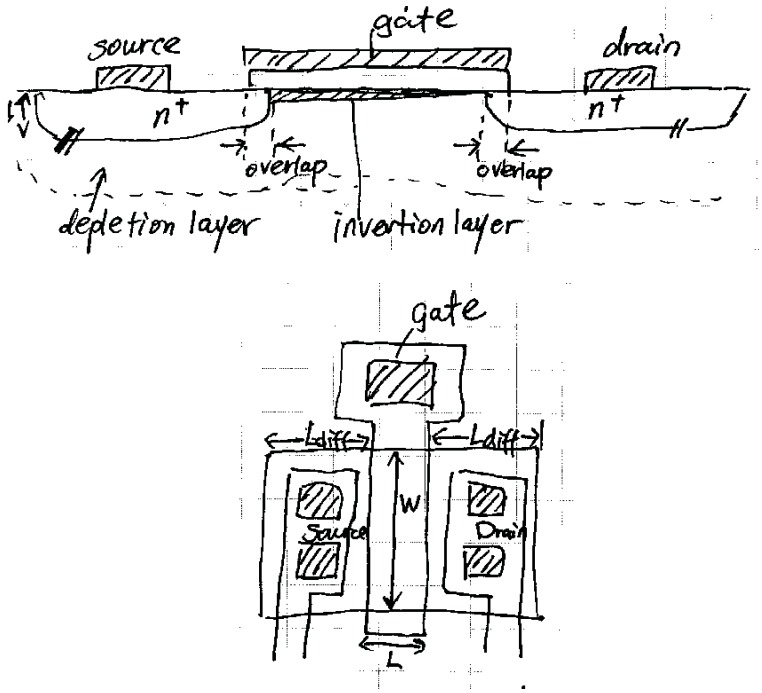


Figure 5: Capacitances (under saturation)

1. C_{gs}, two contributions, first is the MOS capacitor capacitance under inversion was C_{ox} before, but since under saturation regime, we have large V_{DS}, the inversion layer is non-uniform in charge density, need to do integration of q_{G} to consider this. We will skip the math here since it is derived in lecture already. The result is $wLC_{\text{ox}} \longrightarrow \frac{2}{3}C_{\text{ox}} \cdot wL$. The other contribution is from the overlap between the source and gate Cov

$$\therefore C_{\text{GS}} = \frac{2}{3}wL \cdot C_{\text{ox}} + w \cdot Cov \quad Cov \text{ unit is} \quad \text{F/cm}$$

2. $C_{\text{GD}} = w \cdot Cov$ L, Ldiff, w, see Fig. 5 above

3. C_{BS} or C_{SB} = pn junction capacitance underneath the source area and side wall, is:

$$\begin{aligned} &= w \cdot L_{\text{diff}} C_{\text{j}}^{(s)} + (2L_{\text{diff}} + w) \cdot C_{\text{jsw}}(s) \\ C_{\text{j}}^{(s)} &= \sqrt{\frac{q\epsilon_{\text{s}} N_{\text{a}}}{2(\phi_{\text{B}} - V_{\text{BS}})}} \end{aligned}$$

$$C_{\text{jsw}}(s) \quad \text{is} \quad \text{usually given, should be} \sqrt{\frac{q\epsilon_{\text{s}} N_{\text{a}}}{2(\phi_{\text{B}} - V_{\text{BS}})}} \cdot d$$

4. C_{bd} or C_{db} = pn junction capacitance underneath the drain area and side wall, is:

$$\begin{aligned} &= w \cdot L_{\text{diff}} \cdot C_{\text{j}}^{(D)} + (2L_{\text{diff}} + w) \cdot C_{\text{jsw}}^{(D)} \\ C_{\text{j}}^{(D)} &= \sqrt{\frac{q\epsilon_{\text{s}} N_{\text{a}}}{2(\phi_{\text{B}} - V_{\text{BD}})}} \end{aligned}$$

5. C_{gb} is due to the presence of inversion layer (screening) under inversion, the capacitance of C_{gb} can be ignored (it only present at cut off).

MOSFETS in Digital Circuits

Now we have both the low frequency and high frequency version of the small signal equivalent circuit. What has it to do with our MOSFET digital circuit discussion?
For digital circuits, there are two important aspects:

1. Noise Margin: larger noise margin $\longrightarrow$ higher noise immunity, better
2. Speed: the concept of "propagation delay". We want circuit response to be fast, $\Longrightarrow$ *low* propagation delay

Now the question is, when we design a circuit, what parameters affect the noise margin and what affects the delay?

Noise Margin

Our inverter is shown below:

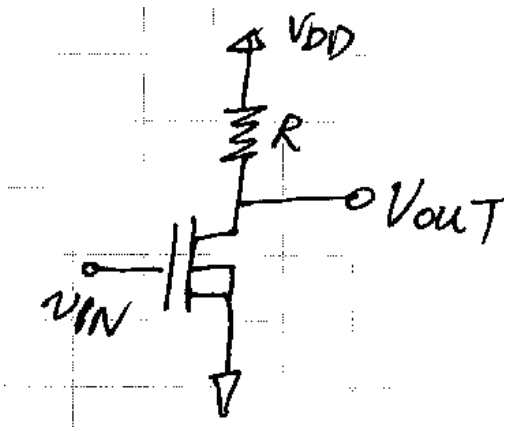
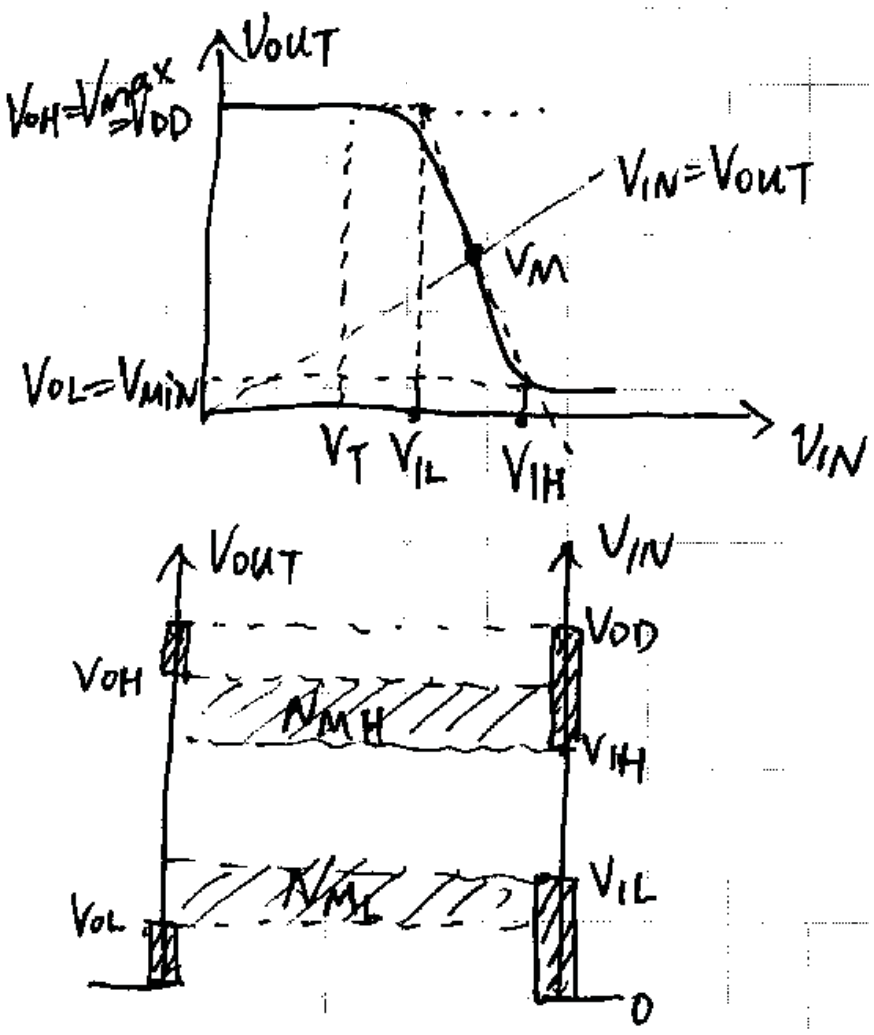


Figure 6: n-MOS inverter

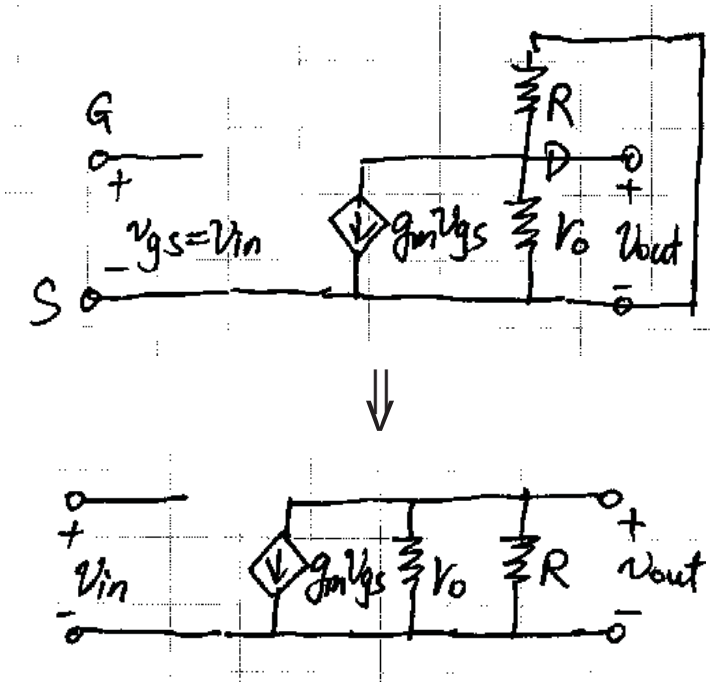
$$N_{\mathrm{MH}} = V_{\mathrm{OH}} - V_{\mathrm{IH}}$$
$$N_{\mathrm{ML}} = V_{\mathrm{IL}} - V_{\mathrm{OL}}$$

In order to have a high noise margin, we want high slope at $V_{\mathrm{m}} \implies$ high gain Av at V_{m}.



Gain

How to find gain at V_{m}? Use a small signal circuit:



$$A_v = \frac{V_{out}}{V_{in}} = -g_m(r_o||R) \simeq -g_m R$$

The larger the R the bigger the noise margin. However, as we will see later, the larger R the slower the circuit. There is a tradeoff since capacitances in the circuit add delay.

6.012 Microelectronic Devices and Circuits
Spring 2009