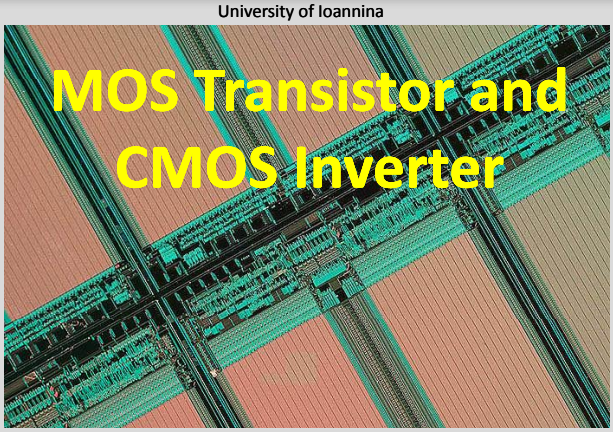
CMOS INTEGRATED CIRCUIT DESIGN TECHNIQUES

University of Ioannina



Dept. of Computer Science and Engineering

Survey on CMOS Digital Circuits

Y. Tsiatouhas

CMOS Integrated Circuit Design Techniques

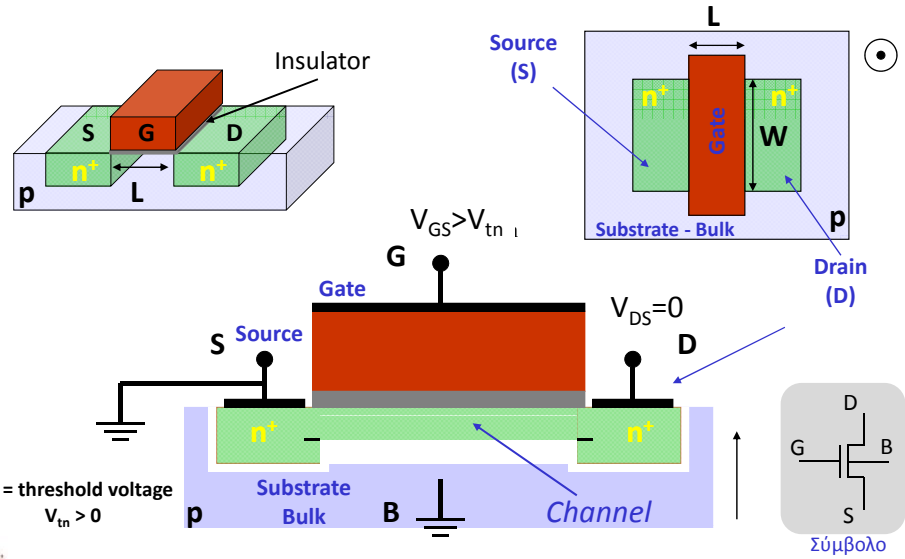
Overview

1. MOS transistor
2. Channel modulation
3. Threshold voltage
4. Velocity saturation
5. CMOS inverter
6. Noise margins
7. CMOS technology scaling

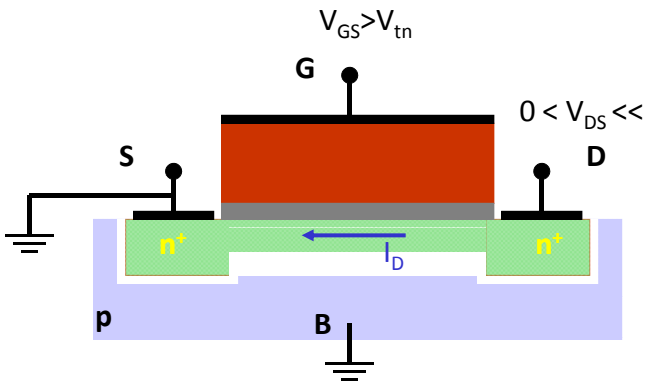
VLSI Systems
and Computer Architecture Lab

MOS Transistor

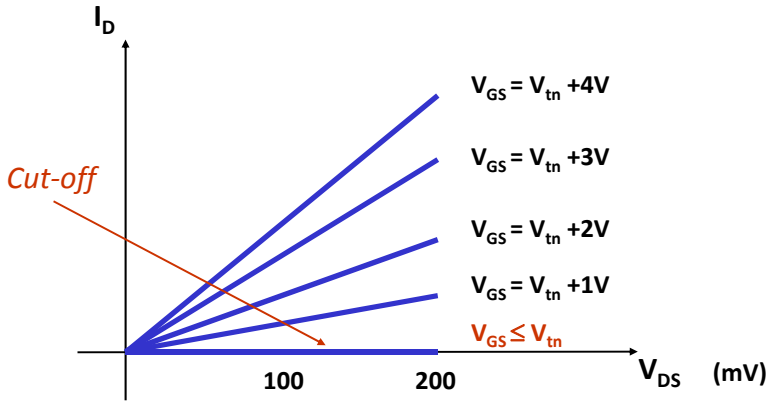
MOS-FET (nMOS Transistor)



nMOS Transistor – Operation (I)

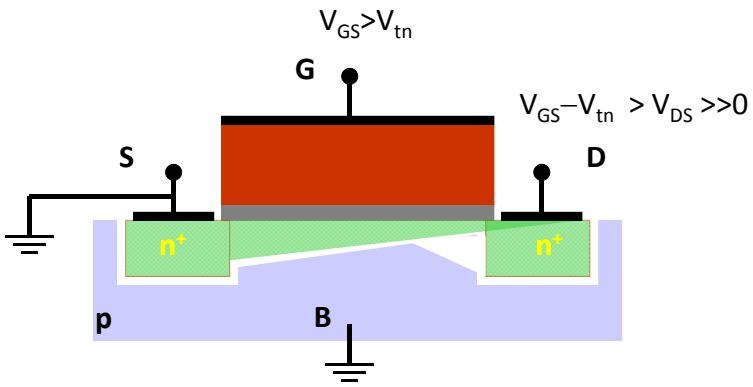


nMOS Transistor – Operation (II)



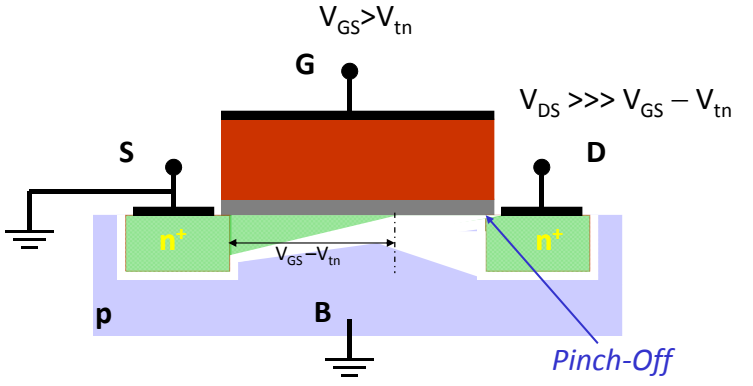
Linear Region

nMOS Transistor – Operation (III)



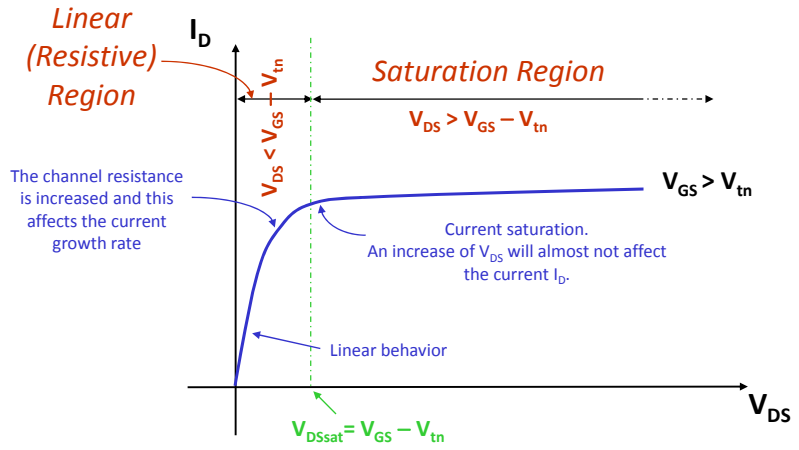
Linear region: $V_{DS} \leq V_{GS} - V_{tn}$

nMOS Transistor – Operation (IV)



Saturation: $V_{DS} > V_{GS} - V_{tn}$

nMOS I_D-V_{DS} Characteristic



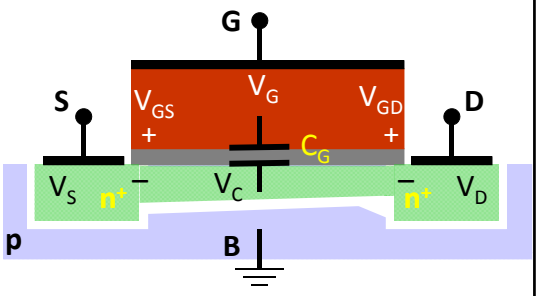
Transistor Current Estimation I

The channel charge is:

$$Q_C = C_G(V_{GC} - V_{tn})$$

The **mean** potential value at the channel is :

$$V_c = \frac{(V_S + V_D)}{2} = V_S + \frac{V_{DS}}{2}$$



Long-channel transistor, L >>

The mean gate-channel potential difference is:

$$V_{GS} = V_G - V_c = V_{GS} - \frac{V_{DS}}{2}$$

The gate-channel capacitance is expressed as:

$$C_G = \varepsilon_{ox}\varepsilon_0 \frac{WL}{t_{ox}} = C_{ox} WL$$

ε_{ox} = relative oxide permittivity
ε_0 = vacuum permittivity
t_{ox} = gate oxide thickness

$$C_{ox} = \frac{\varepsilon_{ox}\varepsilon_0}{t_{ox}}$$

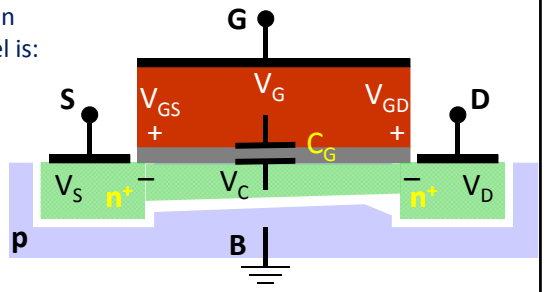
Transistor Current Estimation II

Under lateral electric fields, the mean velocity of the carriers at the channel is:

$$\upsilon = \mu_n E$$

μ_n = electron mobility
E = electric field tension

It stands: $E = \frac{V_{DS}}{L}$



The drain current can be expressed as:

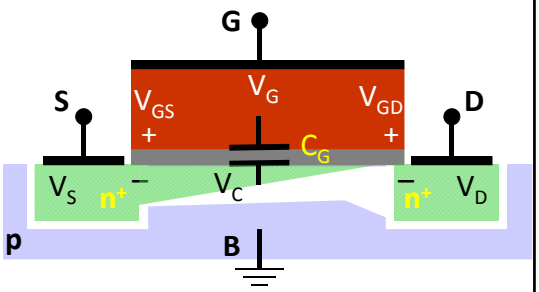
$$I_D = \frac{Q_C}{T} = \frac{Q_C}{L/\upsilon} = \frac{Q_C}{L}\mu_n E = \frac{Q_C}{L^2}\mu_n V_{DS} = \mu_n C_{ox}\frac{W}{L}\left(V_{GS} - V_{tn} - \frac{V_{DS}}{2}\right)V_{DS}$$

$$I_D = \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{tn})V_{DS} - \frac{V_{DS}^2}{2} = k_n(V_{GS} - V_{tn})V_{DS} - \frac{V_{DS}^2}{2}$$

$$k_n = \mu_n C_{ox}\frac{W}{L}$$

Transistor Current Estimation III

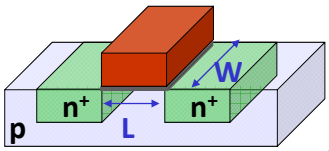
For $V_{DS} > V_{DSsat} = V_{GS} - V_{tn}$ (saturation) the channel is pinched off and the drain current does not depend on the V_{DS} .



Consequently, in the drain current equation the V_{DS} is replaced by the V_{DSsat} so that:

$$I_D = k_n(V_{GS} - V_{tn})(V_{GS} - V_{tn}) - \frac{(V_{GS} - V_{tn})^2}{2} = \frac{k_n}{2}(V_{GS} - V_{tn})^2$$

nMOS Transistor Current Equations



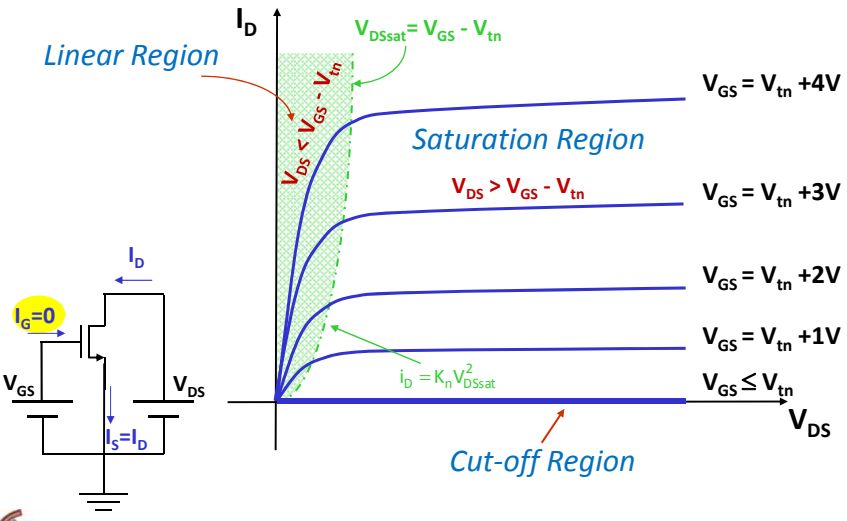
$$k_n = \mu_n C_{ox} \frac{W}{L} = \frac{\mu_n \varepsilon}{t_{ox}} \frac{W}{L}$$

Current gain factor

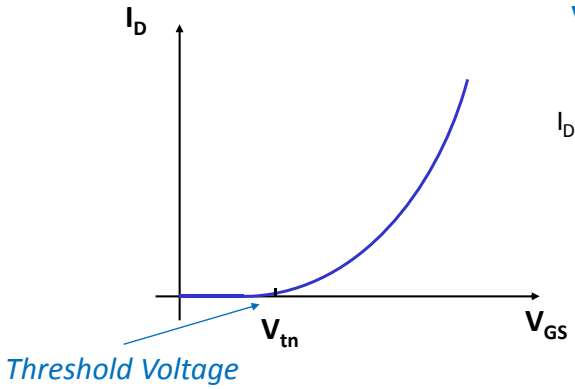
$\varepsilon = \varepsilon_{ox} \cdot \varepsilon_0$ = oxide permittivity

$$I_D = \begin{cases} 0 & V_{GS} - V_{tn} < 0 \text{ Cut-off region} \\ k_n \left[(V_{GS} - V_{tn}) V_{DS} - \frac{V_{DS}^2}{2} \right] & 0 < V_{DS} < V_{GS} - V_{tn} \text{ Linear region} \\ \frac{k_n}{2} (V_{GS} - V_{tn})^2 & 0 < V_{GS} - V_{tn} < V_{DS} \text{ Saturation region} \end{cases}$$

nMOS I_D-V_{DS} Characteristic



nMOS I_D-V_{GS} Saturation Characteristic



Saturation

$V_{DS} \geq V_{GS} - V_{tn}$

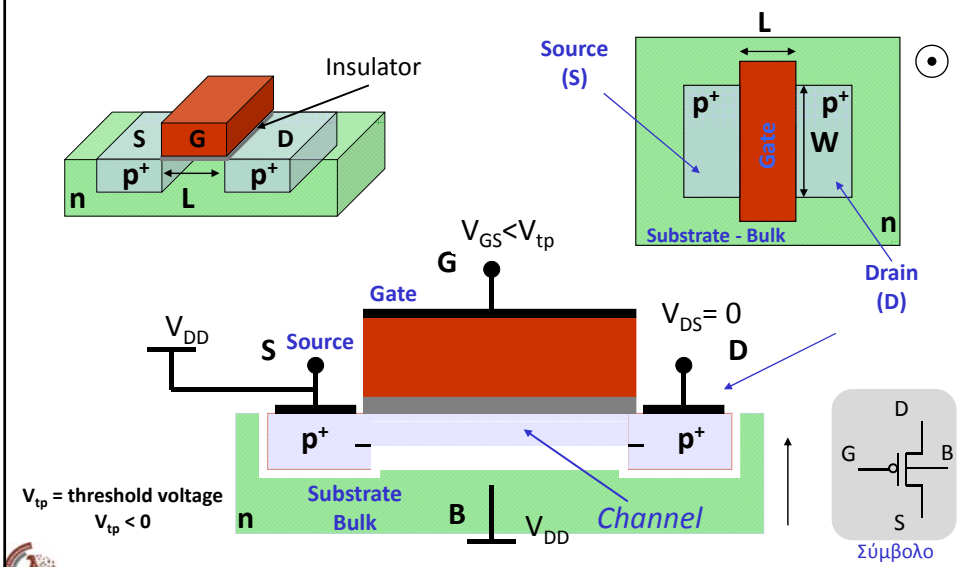
$$I_D = \frac{k_n}{2}(V_{GS} - V_{tn})^2$$

I_D Dependencies

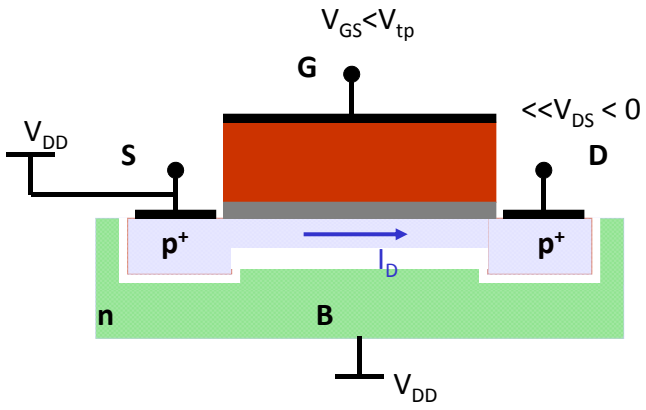
The drain current (on current) depends on:

- The channel length (L)
- The channel width (W)
- The threshold voltage (V_t)
- The gate oxide thickness (t_{ox})
- The gate oxide permittivity (ε)
- The electron/hole mobility (μ)

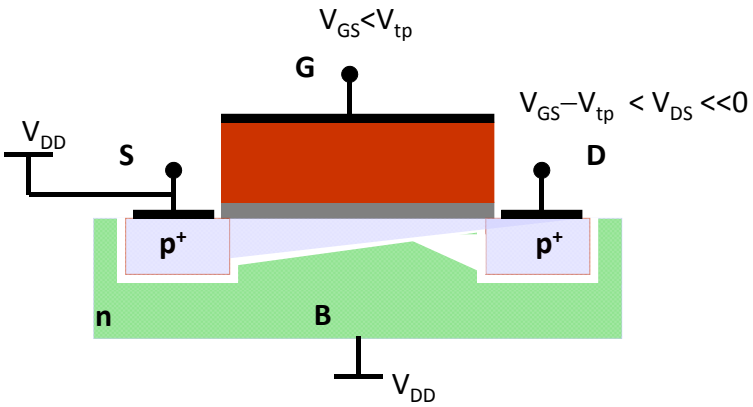
pMOS Transistor



pMOS Transistor – Operation (I)

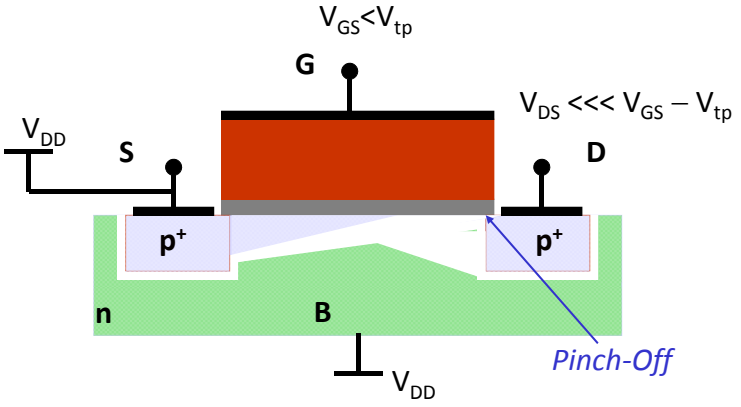


pMOS Transistor – Operation (II)



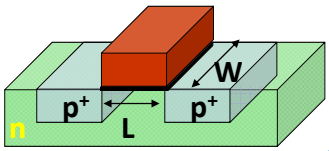
Linear region: $0 > V_{DS} > V_{GS} - V_{tp}$

pMOS Transistor – Operation (III)



Saturation: $V_{DS} < V_{GS} - V_{tp} < 0$

pMOS Transistor Current Equations



$$k_p = \mu_p C_{ox} \frac{W}{L} = \frac{\mu_p \varepsilon}{t_{ox}} \frac{W}{L}$$

Current gain factor

$\varepsilon = \varepsilon_{ox} \cdot \varepsilon_0$ = oxide permittivity

It stands that: $\mu_p < \mu_n$!

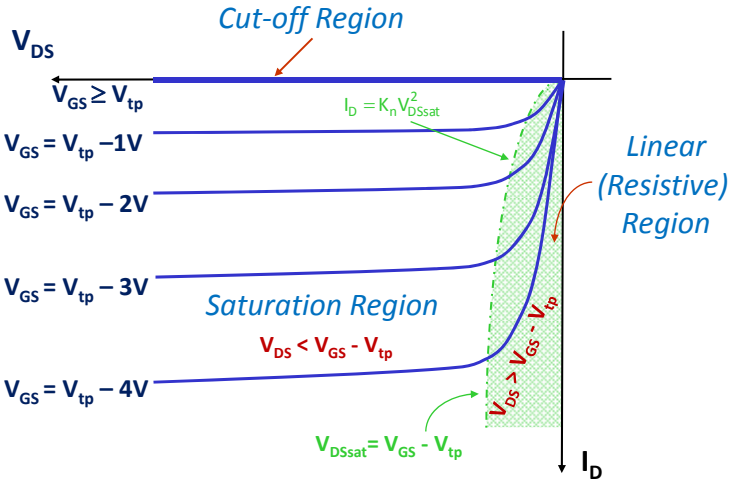
$$I_D = \begin{cases} 0 \\ k_p\left[(V_{GS} - V_{tp})V_{DS} - \frac{V_{DS}^2}{2}\right] \\ \frac{k_p}{2}(V_{GS} - V_{tp})^2 \end{cases}$$

$V_{GS} - V_{tp} > 0$
Cut-off region

$0 > V_{DS} > V_{GS} - V_{tp}$
Linear region

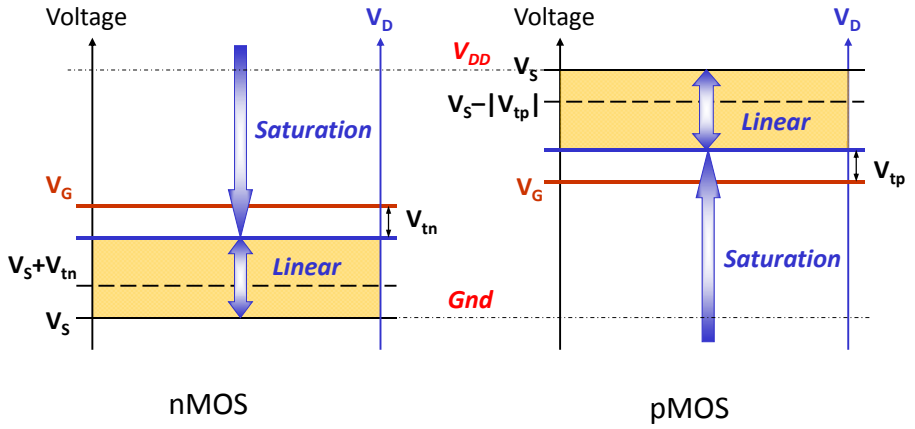
$0 > V_{GS} - V_{tp} > V_{DS}$
Saturation region

pMOS I_D-V_{DS} Characteristic

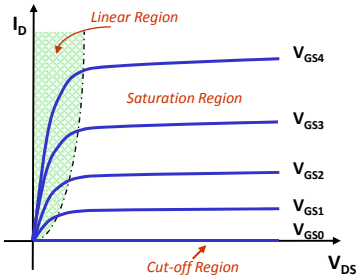


$V_{tp}<0$

Linear and Saturation Regions



Transconductance



Linear Region

The channel resistance (transistor ourput resistance) is given by:

$$\left.\frac{dI_D}{dV_{DS}}\right|_{V_{DS}\to 0} \approx k(V_{GS}-V_t) \Rightarrow R_C = \frac{1}{k(V_{GS}-V_t)}$$

The transconductance (g_m) represents the relation between I_D and V_{GS} and it is defined as follows:

$$g_m = \left.\frac{dI_D}{dV_{GS}}\right|_{V_{DS}=\sigma\tau\alpha\theta.} = k\cdot V_{DS}$$

Saturation

The MOS transistor operates like a current source since the current I_D is almost independent of the voltage V_{DS}. The transconductance is provided by:

$$g_m = k(V_{GS}-V_t)$$

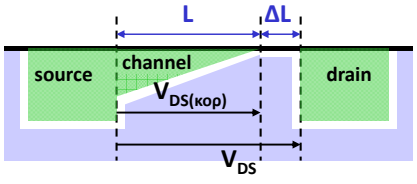
Additional MOS Transistor Dependencies

The equations and curves presented earlier are approximations of the MOS transistor operation. A more detailed analysis must consider the following phenomena:

- the channel-length modulation (*διαμόρφωση μήκους καναλιού*)
- the *body effect* (*φαινόμενο σώματος*) and other V_t dependencies
- the *velocity saturation* (*κορεσμός ταχύτητας*)

Channel Length Modulation

Saturation:



Increasing the V_{DS} by ΔV_{DS} above $V_{DS(κορ)}$ [$V_{DS}=V_{DS(κορ)}+\Delta V_{DS}$] the channel length L is decreased by ΔL. Since I_D is conversely proportional of the channel length, I_D is increased. Consequently, in saturation stands that:

$$I_D = \frac{k}{2}(V_{GS} - V_t)^2(1 + \lambda V_{DS})$$

The Body Effect

The threshold voltage is the minimum gate-to-source voltage in a transistor in order to establish an effective current I_D. In general the threshold voltage depends on:

- the gate material
- the gate insulator (oxide) material
- the oxide thickness
- the channel doping concentration and
- the voltage difference between the source and the bulk (V_{SB})

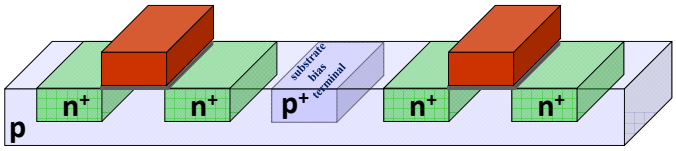
For a given gate material the threshold voltage is provided by the following equation :

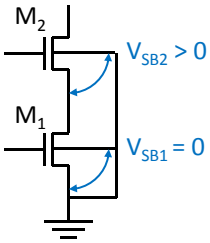
$$V_t = V_{t0} \pm \gamma\left(\sqrt{|(-2)\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right) \quad \text{where} \quad \gamma = \frac{t_{ox}}{\varepsilon_{ox}}\sqrt{2q\varepsilon_{Si}N}$$

Body Effect Coefficient

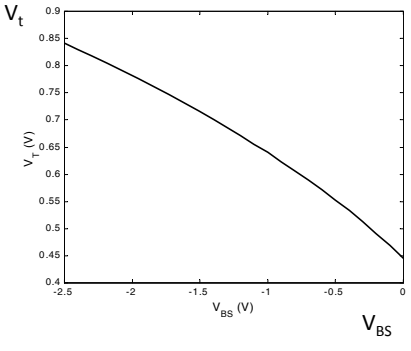
where V_{t0} is the threshold voltage for V_{SB}=0, ϕ_F is the Fermi potential (≈ -0.3V). t_{ox} is the oxide thickness, ε_{ox} is the relative dielectric constant of the oxide, q is the electron charge, ε_{Si} the relative dielectric constant of the silicon and N is the channel doping concentration. The sign +/- refers to an nMOS / pMOS transistor respectively.

The Body Effect



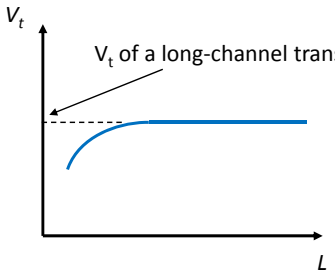


$V_{t2} > V_{t1}$

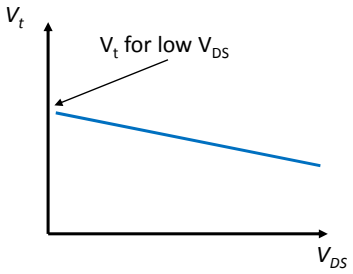


The increment of V_{SB} results in the increment of the channel depletion region which changes the transistor geometry so that the substrate turns to act as a second gate and thus the threshold voltage is increased.

Threshold Voltage Dependencies



Threshold voltage as a function of the channel length for low V_{DS}



Drain-induced barrier lowering (DIBL) for low L

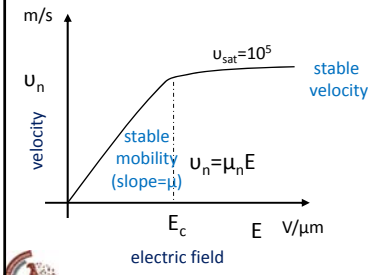
Mobility Degradation (μ)

Under the presence of high vertical electric fields (V_{GS}/t_{ox}) the carriers scatter off the oxide interface more often so that the mobility (μ) is degraded ($\mu_{eff} < \mu$) !
Thus, for a certain V_{GS} a lower current I_D is present.

Velocity Saturation

All I-V equations above apply under the assumption that the carriers' velocity υ in a transistor is proportional to the electric field E, which means that the carriers mobility μ is stable. However, the velocity $\upsilon_{p/n}$ is saturated when the electric field reaches a critical level E_c, due to the scattering of the carriers in the channel.

In short channel transistors (L<<) and under the velocity saturation influence, the current is saturated earlier and at a lower value than the expected, according to the following equation:

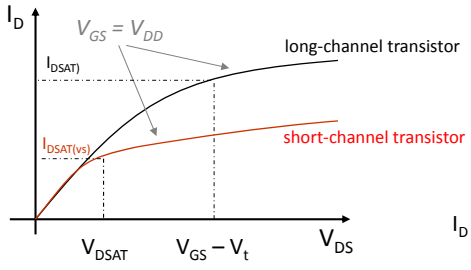


Linear region current

$$I_D = k_n \cdot \left[(V_{GS} - V_{tn}) V_{DS} - \frac{V_{DS}^2}{2} \right] \cdot \frac{1}{1 + (V_{DS}/E_c \cdot L)}$$

Velocity saturation impact

Velocity Saturation Impact (I)

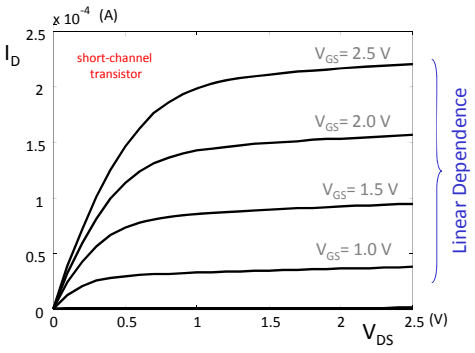


The W/L ratio and the V_{GS} of the two transistors are the same in both cases

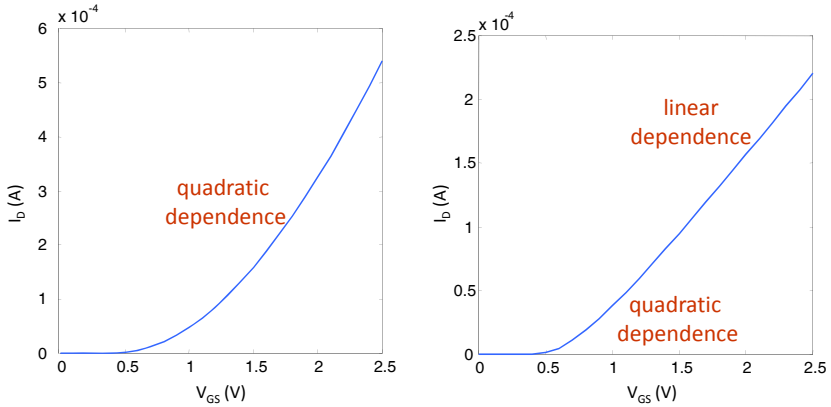
$V_{DSAT} < V_{GS} - V_t$

$I_{DSAT(vs)} < I_{DSAT}$

$$V_{DSAT} \cong E_c \cdot L$$



Velocity Saturation Impact (II)



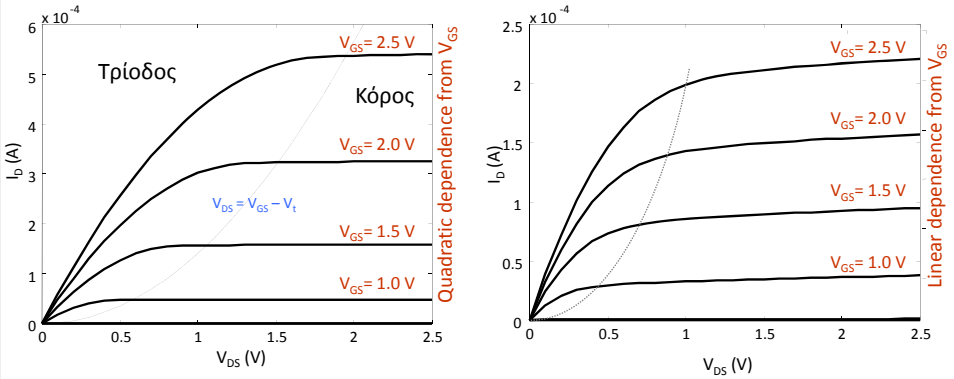
Long-channel transistor

Short-channel transistor

Impact of velocity saturation on short-channel transistors: Increasing the V_{DS} voltage the drain current I_D reaches the saturation earlier than the expected level (V_{GS}–V_t) and I_D has a linear dependence on the V_{GS} voltage.

Velocity Saturation Impact (III)

The W/L ratio of the two transistors is the same in both cases.



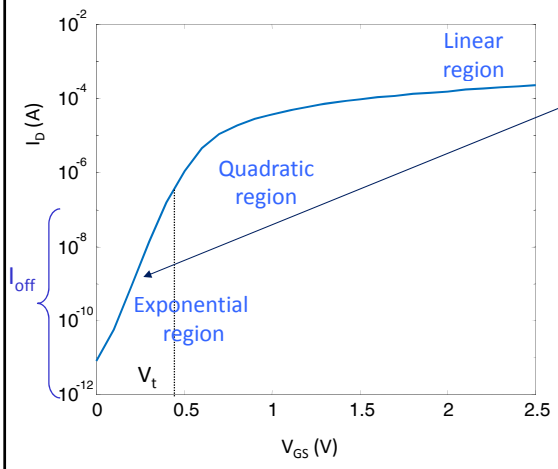
Long-channel transistor

Short-channel transistor

$$I_D = \frac{k}{2}(V_{GS} - V_{tn})^2 \qquad \text{saturation} \qquad I_D = \frac{k}{2}(V_{GS} - V_{tn})^{\alpha}$$

$1 < \alpha < 2$

Weak Inversion Current



$$I_{off} = I_S e^{\frac{V_{GS}-V_t}{nkT/q}} \left(1 - e^{-\frac{V_{DS}}{kT/q}}\right)(1+\lambda V_{DS})$$

I_S and n are empirical parameters, with $n \geq 1$

⌘

Slope factor:

$$S = n\left(\frac{kT}{q}\right)\ln(10)$$

expresses the reduction rate (that is the required reduction of V_{GS} in order to reduce I_D by 10).

For n=1.5 the slop factor is : S=90mV/decade

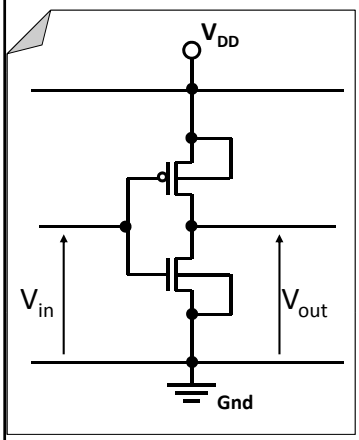
Additional Phenomena

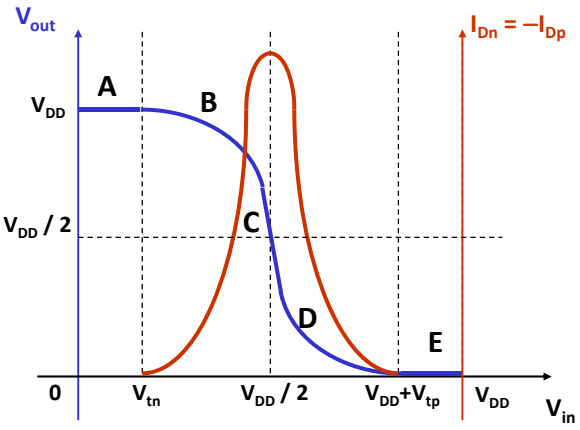
- **Mobility Variation**
  The mobility (μ) of the carriers (electrons/holes) is decreased by increasing the doping concentration and/or the temperature.
- **Fowler-Nordheim and Direct Tunnelling**
  For thin gate oxides in modern nanotechnologies a current is present through this insulator. This current is due to quantum-mechanics tunneling phenomena and is proportional to the gate surface area.
- **Hot Electrons**
  As the channel length (L) decreases the drain electric field increases so that the electrons gain enough energy (hot electrons) as they move in the channel to deflect towards the gate under the influence of the gate potential.
- **Substrate Breakdown**
  For short channels (L) and high drain voltages, the depletion area of the drain expands towards the source so that a drain-to-source current is established which is independent of the gate voltage. This phenomenon does not result in a permanent transistor failure.
- **Gate Oxide Dielectric Breakdown**
  For high gate-to-source voltages the gate oxide may breakdown. This will lead to a permanent transistor failure.

CMOS Inverter

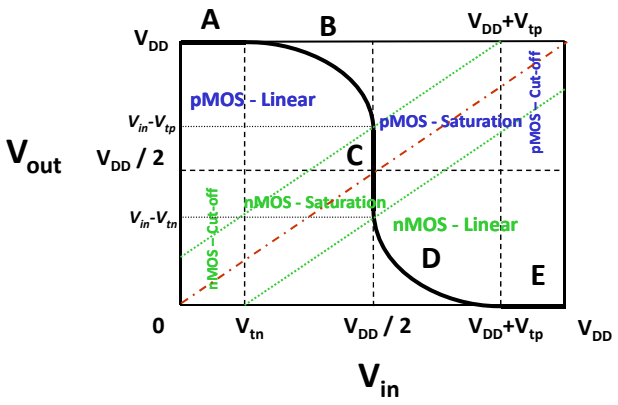
Input-Output Characteristic Curve (I)

Input-Output (Static) Characteristic Curve [$V_{out}=f(V_{in})$]





Input-Output Characteristic Curve (II)



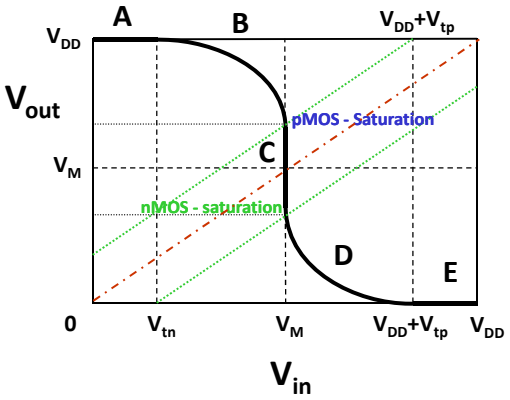
$V_{tp}<0$

At the C region both transistors are in saturation and behave as current sources.

There is an input voltage level for which it stands $V_{in}=V_{out}$ and it is called *transition threshold* (*κατώφλι μετάβασης*) V_M of the logic gate.
At this voltage the whole system is not stable.

(In the figure, the transition threshold is equal to $V_{DD}/2$ given that $k_p=k_n$).

The Transition Threshold



At the C region both transistors are in saturation. In addition, both transistor currents are equal and for $V_{in}=V_M$ it stands:

$$\frac{k_n}{2}(V_M - V_{tn})^2 = \frac{k_p}{2}(V_M - V_{DD} - V_{tp})^2$$

$$\Leftrightarrow$$

$$\frac{k_n}{k_p} = \frac{(V_M - V_{DD} - V_{tp})^2}{(V_M - V_{tn})^2}$$

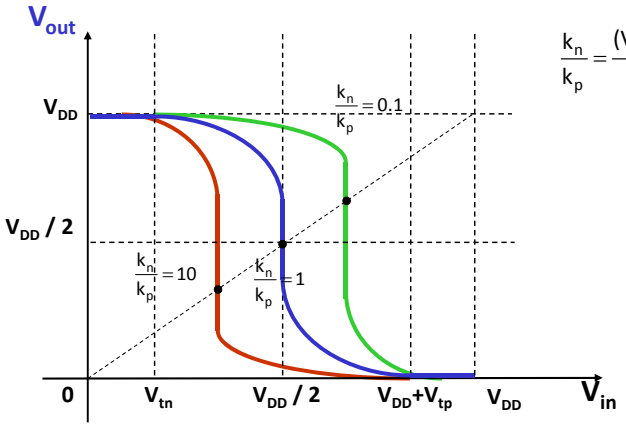
$$k_{p/n} = \frac{\mu_{p/n} \cdot \varepsilon}{t_{ox}} \cdot \frac{W}{L}$$

In case that a specific transition threshold V_M is required, then the appropriate transistor width ratio is given by the next equation:

$$\frac{W_n}{W_p} = \frac{\mu_p}{\mu_n} \frac{(V_M - V_{DD} - V_{tp})^2}{(V_M - V_{tn})^2}$$

$L_n = L_p$

k_n/k_p Effect on the Input-Output Curve

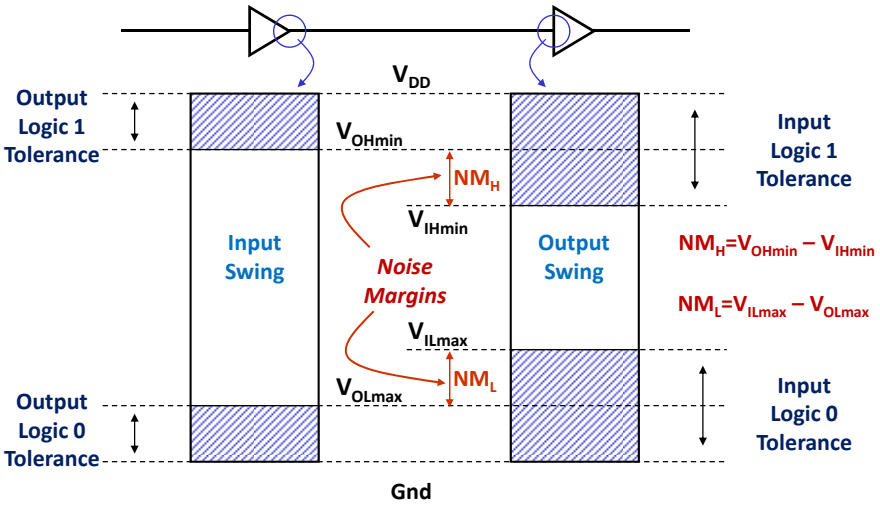


$$\frac{k_n}{k_p} = \frac{(V_M - V_{DD} - V_{tp})^2}{(V_M - V_{tn})^2}\Bigg|_{V_M = \frac{V_{DD}}{2}} = 1$$

$V_{tn} \approx V_{tp}$

$$k_{p/n} = \mu_{p/n}\varepsilon\frac{W}{t_{ox}}\frac{W}{L}$$

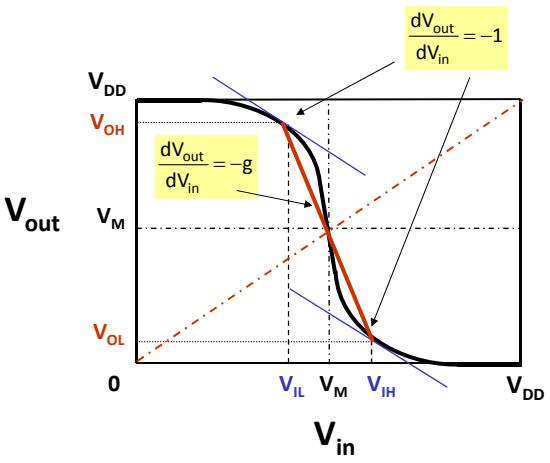
Noise Margins



$NM_H = V_{OHmin} - V_{IHmin}$

$NM_L = V_{ILmax} - V_{OLmax}$

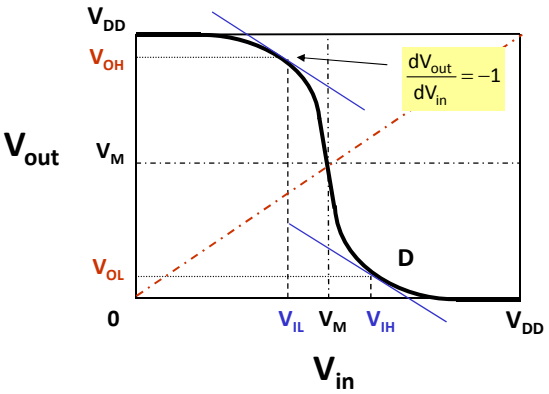
Noise Margins (II)



$$V_{IH} - V_{IL} = -\frac{(V_{OH} - V_{OL})}{g}$$

Noise Margins (III)

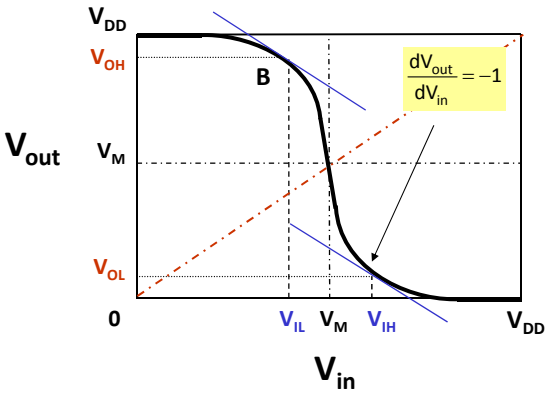
$$\frac{dV_{out}}{dV_{in}} = \left[(V_{in} - V_{tp}) + \left\{ (V_{in} - V_{tp})^2 - 2(V_{in} - V_{DD}/2 - V_{tp})V_{DD} - k_n/k_p (V_{in} - V_{tp})^2 \right\}^{1/2} \right] = -1 \Big|_{V_{in} = V_{IL}}$$



$$V_{IL} = \left. \frac{3V_{DD} - 3|V_{tp}| + 5V_{tn}}{8} \right|_{k_n/k_p = 1}$$

$$NM_L = V_{IL}\big|_{V_{OL} \cong 0}$$

Noise Margins (IV)

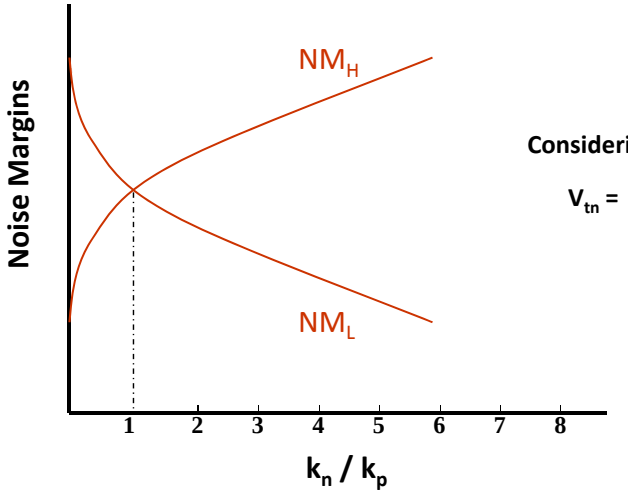


$$V_{IH} = \left.\frac{5V_{DD} - 5|V_{tp}| + 3V_{tn}}{8}\right|_{k_n/k_p=1}$$

$$NM_H = V_{DD} - V_{IH}\big|_{V_{OH}=V_{DD}} \Rightarrow$$

$$NM_H = \left.\frac{3V_{DD} + 5|V_{tp}| - 3V_{tn}}{8}\right|_{k_n/k_p=1}$$

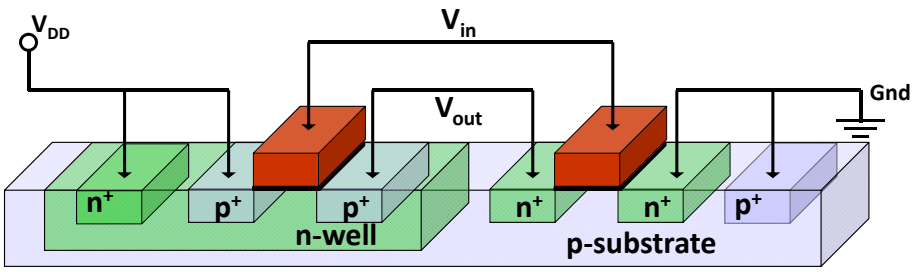
Noise Margins (V)



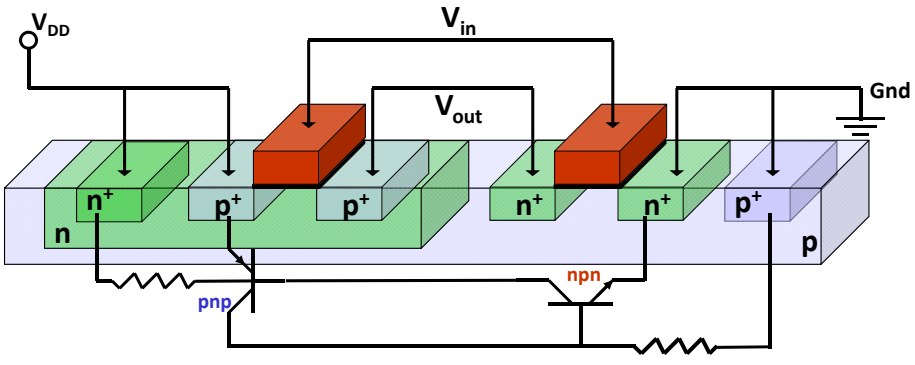
Considering that:

$V_{tn} = |V_{tp}|$

The CMOS Inverter



Latch-Up



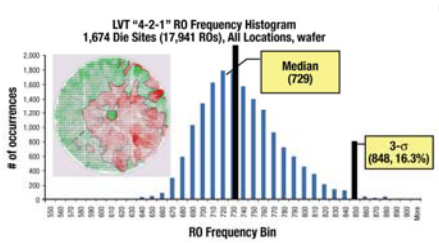
CMOS Technology Scaling

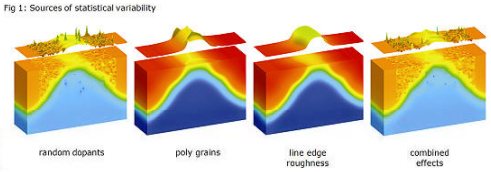
Process Variations

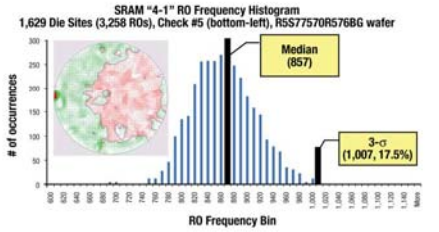
Transistor manufacturing is characterized by process variations either at the wafer level or the die level. Process variations are random and uncorrelated and lead to variations of transistor parameters and geometric characteristics like the W/L ratio, the oxide thickness, the doping concentration, the diffusion depth, the metal widths e.t.c. which result to variations on the transistor transconductance, threshold voltage, parasitic capacitance and resistance …

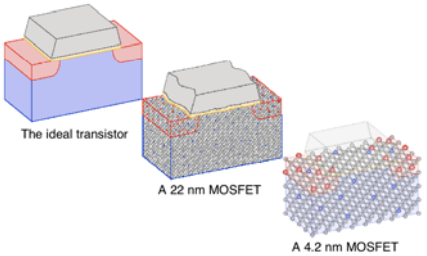
Consequently, process variations affect the expected circuit performance (speed, power consumption, reliability). Aiming to alleviate the situation, the fabs provide "fast" and "slow" models as well as statistical models of the devices for corner and Monte-Carlo analysis respectively.

Process Variations

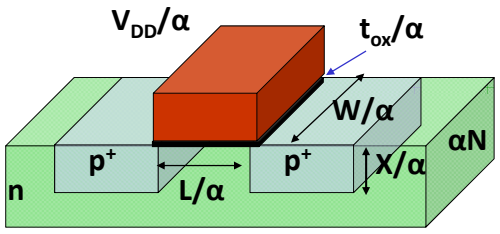








MOS Transistor Size Scaling



Parameter	Scaling Factor	Parameter	Scaling Factor
Electric Field:	*1*	*Power-Speed Product:*	*$1/\alpha^3$*
Depletion Layer:	*$1/\alpha$*	*Gate Area:*	*$1/\alpha^2$*
Parasitic Capacitance:	*$1/\alpha$*	*Power Density:*	*1*
Gate Delay:	*$1/\alpha$*	*Current Density:*	*α*
DC Power Dissipation:	*$1/\alpha^2$*	*Transconductance:*	*1*
Dynamic Power Dissipation:	*$1/\alpha^2$*		

Yield

$$Y = \frac{\#of\_good\_chips\_on\_wafer}{totall\_number\_of\_chips}$$

Seed Model: $Y = e^{-\sqrt{A \cdot D}}$ *A >> & Y<30%*

Murphy Model: $Y = \left(\frac{1 - e^{-A \cdot D}}{A \cdot D} \right)^2$ *A << & Y>30%*

where: A = die area
D = defect density (defects/cm²)

Bibliography

- "Digital Integrated Circuits," J. Rabaey, A. Chandrakasan and B. Nikolic, Prentice Hall, 2003.
- "CMOS VLSI Design: A Circuits and Systems Perspective," N. Weste and D. Harris Addison Wesley, 2010.