The CMOS Inverter

Slides adapted from:
N. Weste and D. Harris, ***CMOS VLSI Design***
Addison-Wesley, 3/e, 2004

Outline

- Robustness of CMOS Inverter – The Static Behavior
  - Switching threshold
  - Noise Margins
- Performance of CMOS Inverter – Dynamic Behavior
  - Propagation delay
- Power Dissipation
  - Static dissipation
  - Dynamic dissipation

Q&A

1. If the width of a transistor increases, the current will
   increase decrease not change
2. If the length of a transistor increases, the current will
   increase decrease not change
3. If the supply voltage of a chip increases, the maximum transistor current will
   increase decrease not change
4. If the width of a transistor increases, its gate capacitance will
   increase decrease not change
5. If the length of a transistor decreases, its gate capacitance will
   increase decrease not change
6. If the supply voltage of a chip increases, the gate capacitance of each transistor will
   increase decrease not change

Q&A

1. If the width of a transistor increases, the current will
   increase decrease not change
2. If the length of a transistor increases, the current will
   increase **decrease** not change
3. If the supply voltage of a chip increases, the maximum transistor current will
   increase decrease not change
4. If the width of a transistor increases, its gate capacitance will
   increase decrease not change
5. If the length of a transistor decreases, its gate capacitance will
   increase **decrease** not change
6. If the supply voltage of a chip increases, the gate capacitance of each transistor will
   increase decrease **not change**

CMOS Inverter Static Behavior: DC Analysis

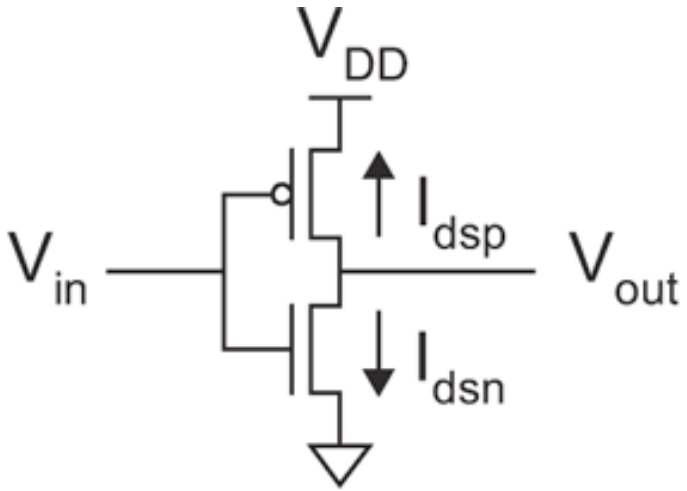


FIG 2.23 A CMOS inverter

CMOS Inverter: DC Analysis

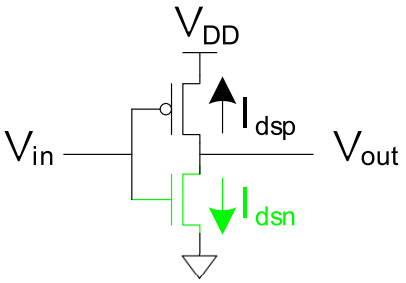
- DC Response: V_{out} vs. V_{in} for a gate
- Inverter
  - When $V_{in} = 0 \rightarrow V_{out} = V_{DD}$
  - When $V_{in} = V_{DD} \rightarrow V_{out} = 0$
  - In between, V_{out} depends on transistor current
  - By KCL, must settle such that
  - $I_{dsn} = |I_{dsp}|$
  - We can solve equations
  - Graphical solution gives very good insight

Transistors operation regions

- Current depends on transistor's operation region
- For what V_{in} and V_{out} are nMOS and pMOS in
  - Cutoff ?
  - Linear ?
  - Saturation ?

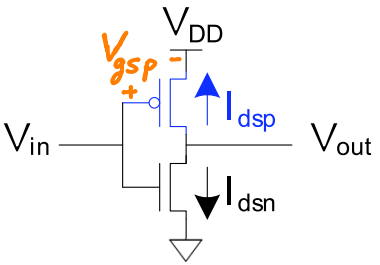
nMOS and pMOS operation

Table 2.2	Relationships between voltages for the three regions of operation of a CMOS inverter		
	Cutoff	**Linear**	**Saturated**
nMOS	$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
	$V_{in} < V_{tn}$	$V_{in} > V_{tn}$	$V_{in} > V_{tn}$
		$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$
		$V_{out} < V_{in} - V_{tn}$	$V_{out} > V_{in} - V_{tn}$
pMOS	$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
	$V_{in} > V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$
		$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$
		$V_{out} > V_{in} - V_{tp}$	$V_{out} < V_{in} - V_{tp}$



$V_{gsn} = V_{in}$

$V_{dsn} = V_{out}$

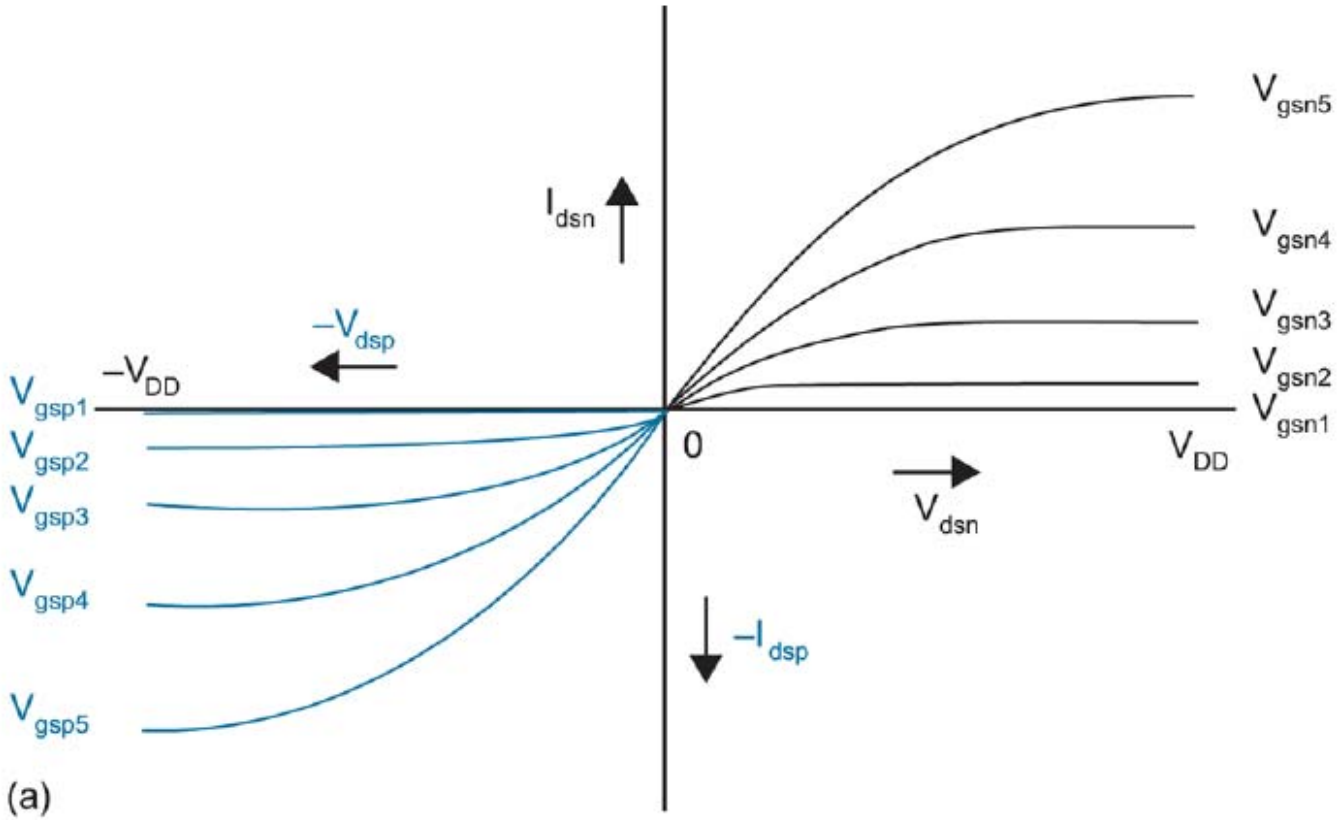


$V_{gsp} = V_{in} - V_{DD}$

$V_{dsp} = V_{out} - V_{DD}$

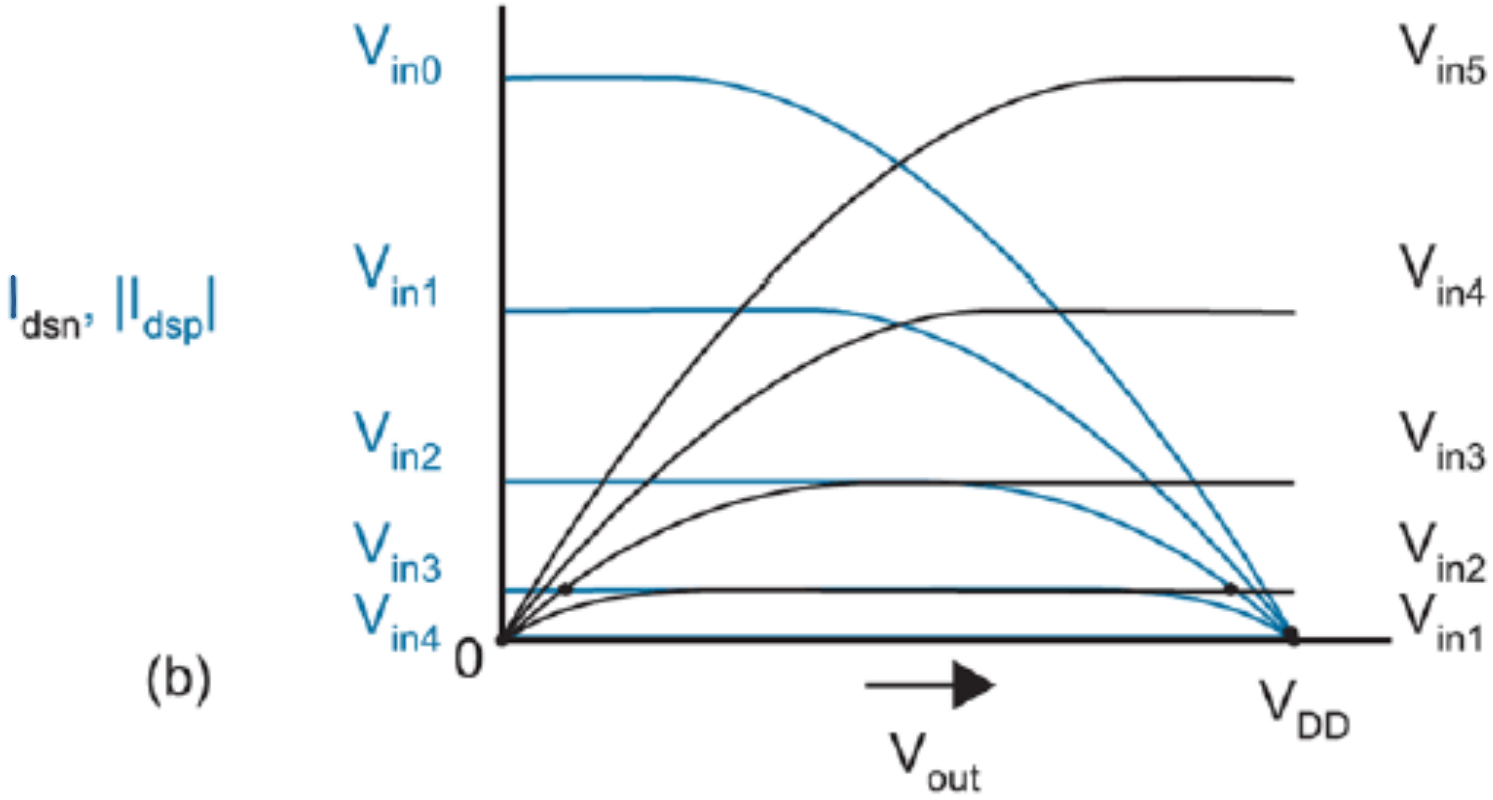
Graphical derivation of the inverter DC response: I-V Characteristics

- Make pMOS wider than nMOS such that $\beta_n = \beta_p$
- For simplicity let's assume $V_{tn}=-V_{tp}$



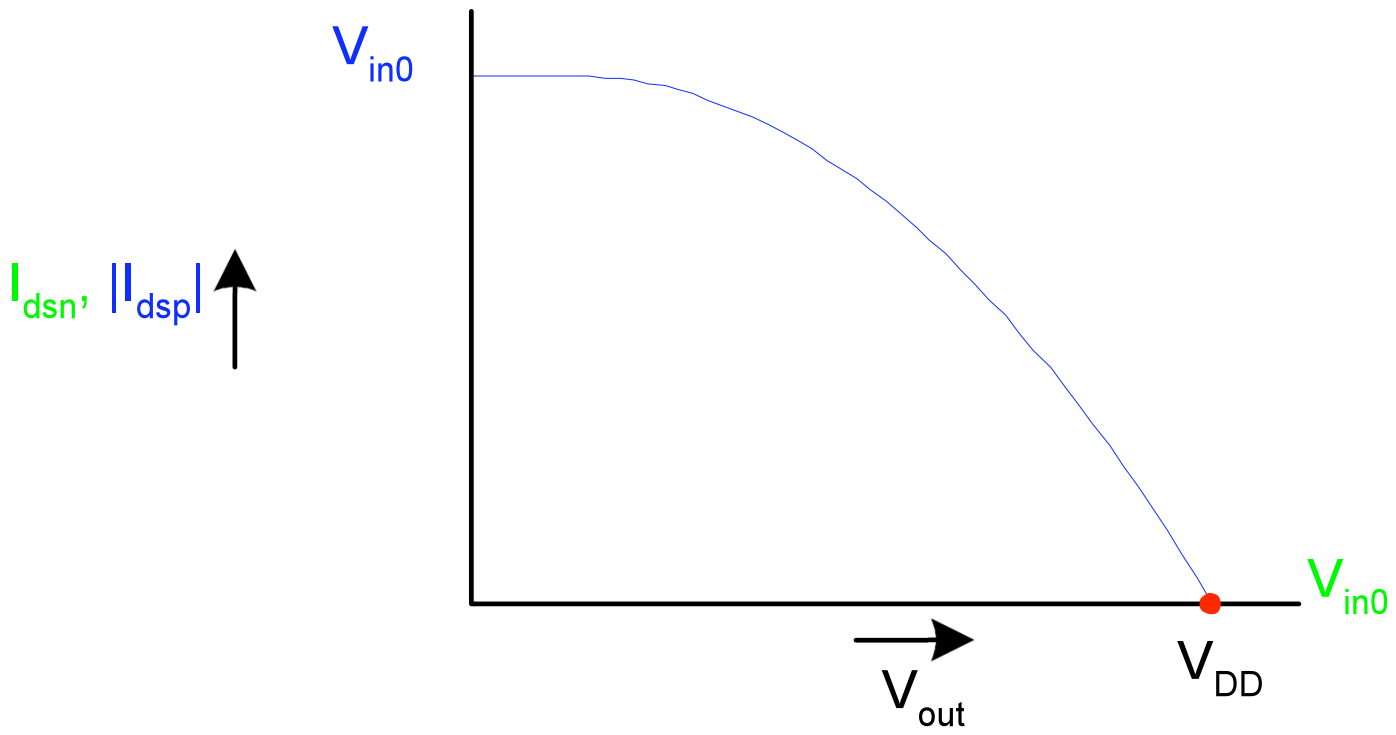
Graphical derivation of the inverter DC response: current vs. V_{out}, V_{in}

- Load Line Analysis:
  For a given V_{in}:
  - Plot I_{dsn}, I_{dsp} vs. V_{out}
  - V_{out} must be where |currents| are equal



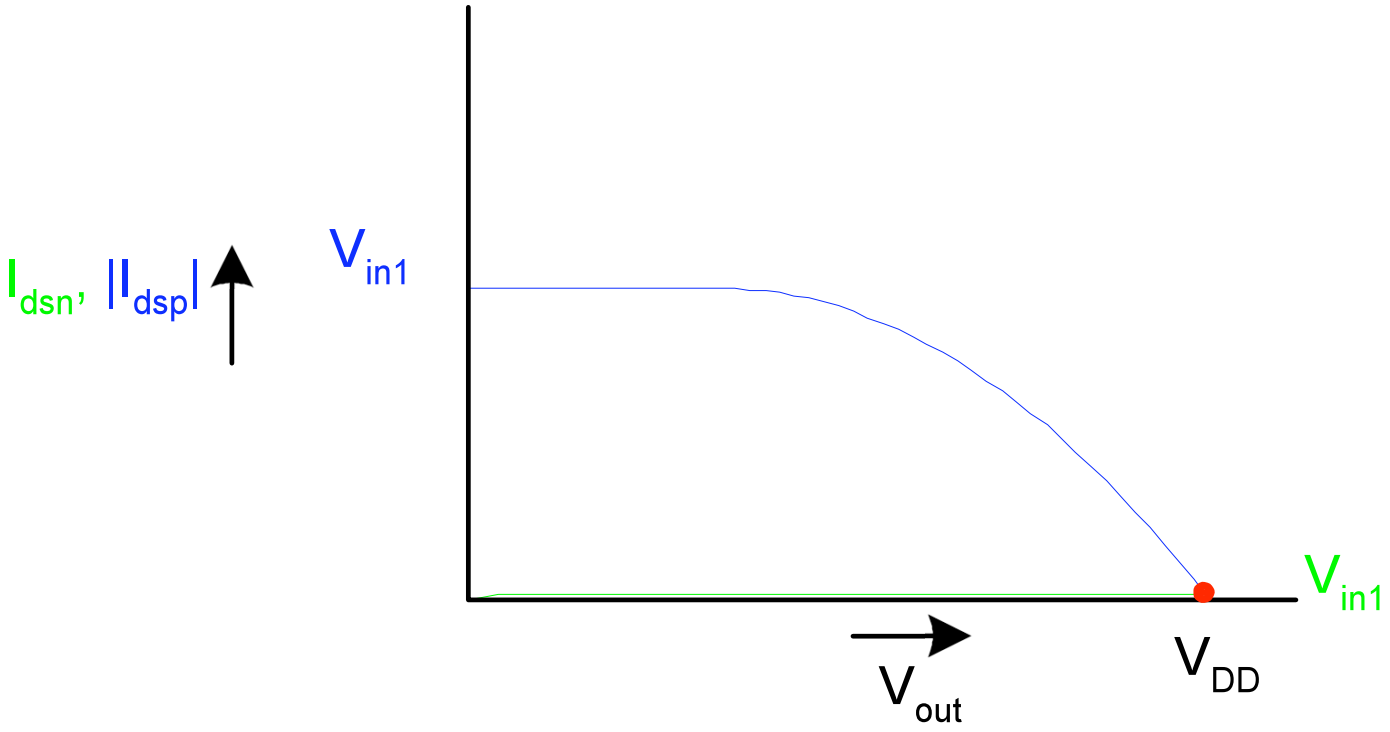
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = 0$



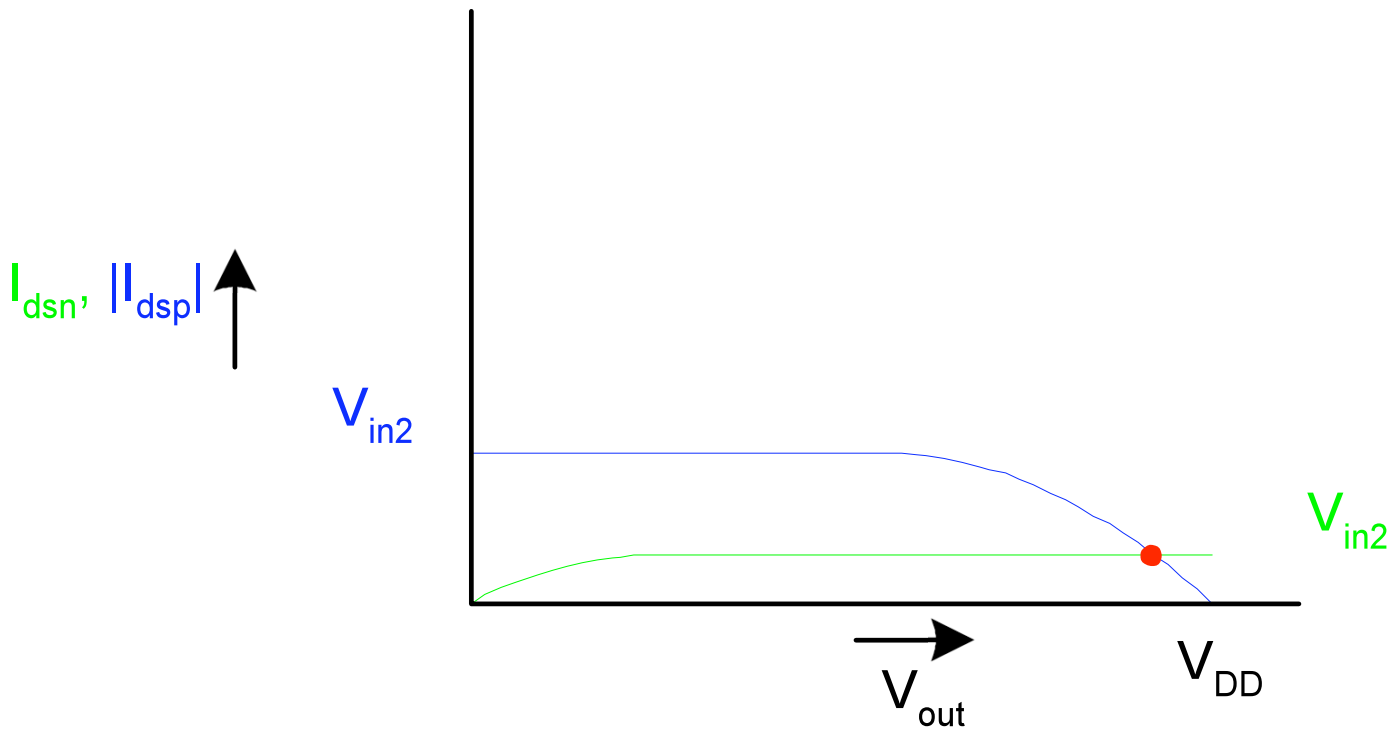
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = 0.2\ V_{DD}$



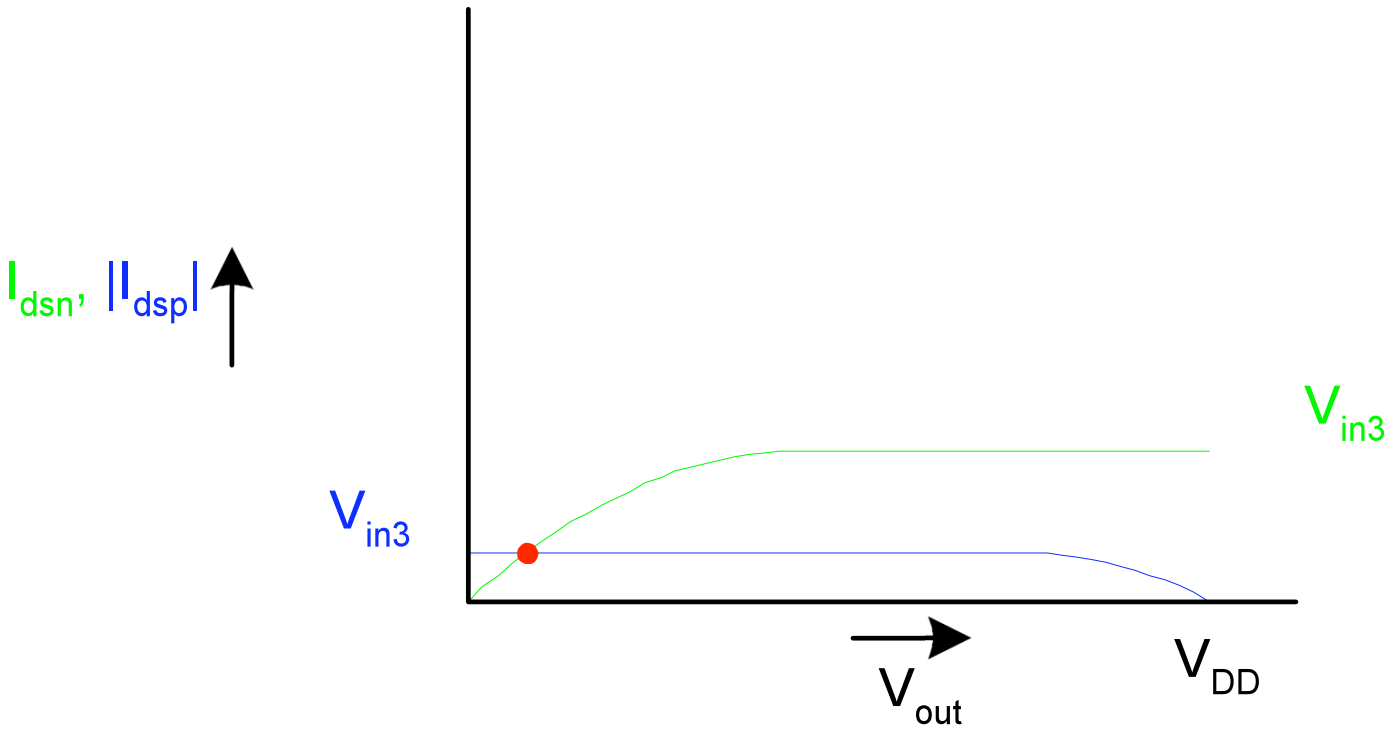
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = 0.4\ V_{DD}$



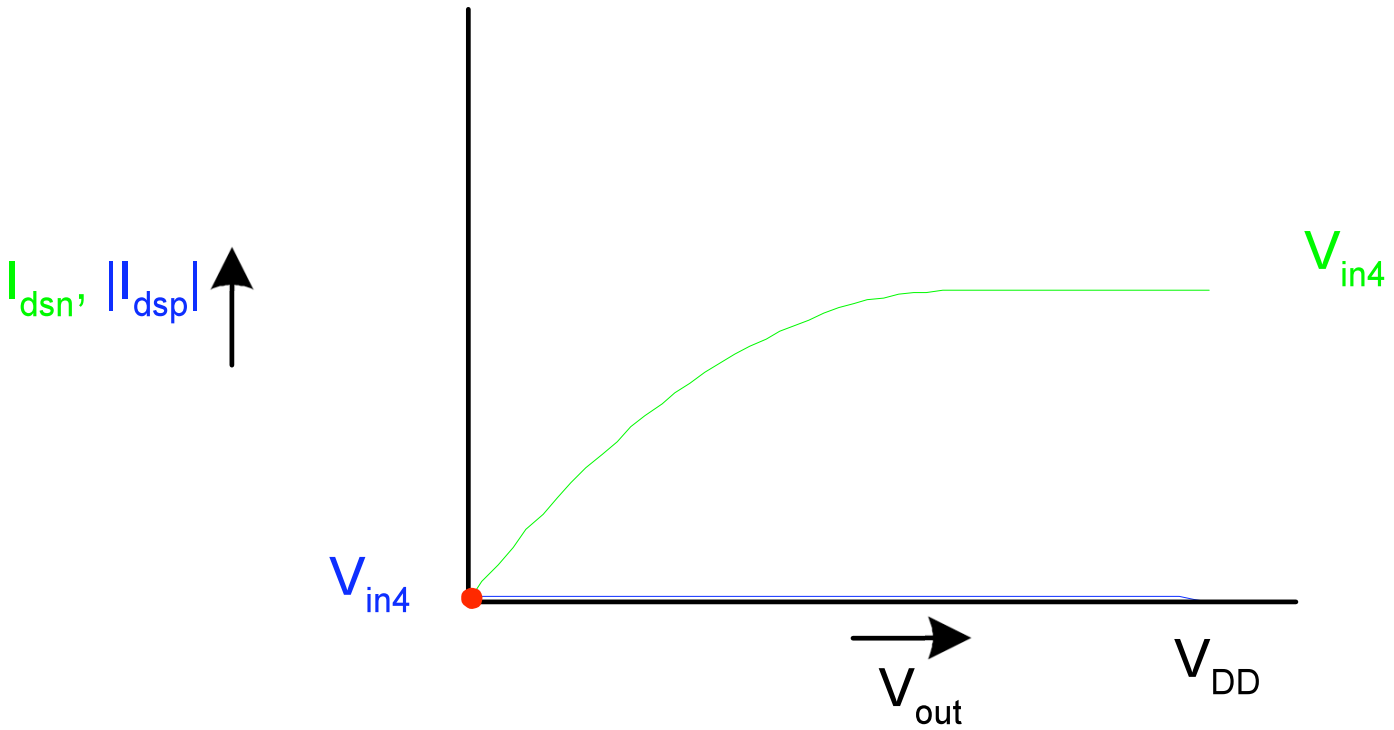
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = 0.6\ V_{DD}$



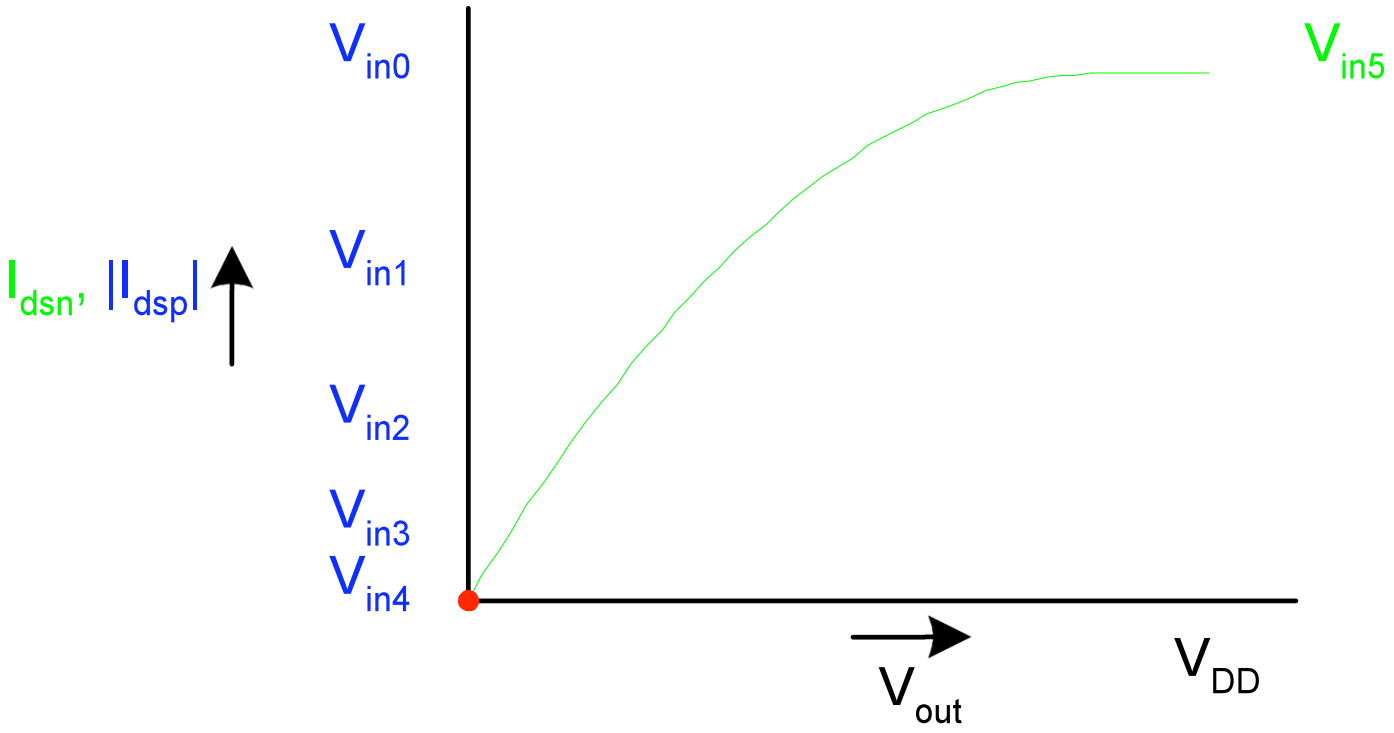
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = 0.8\ V_{DD}$



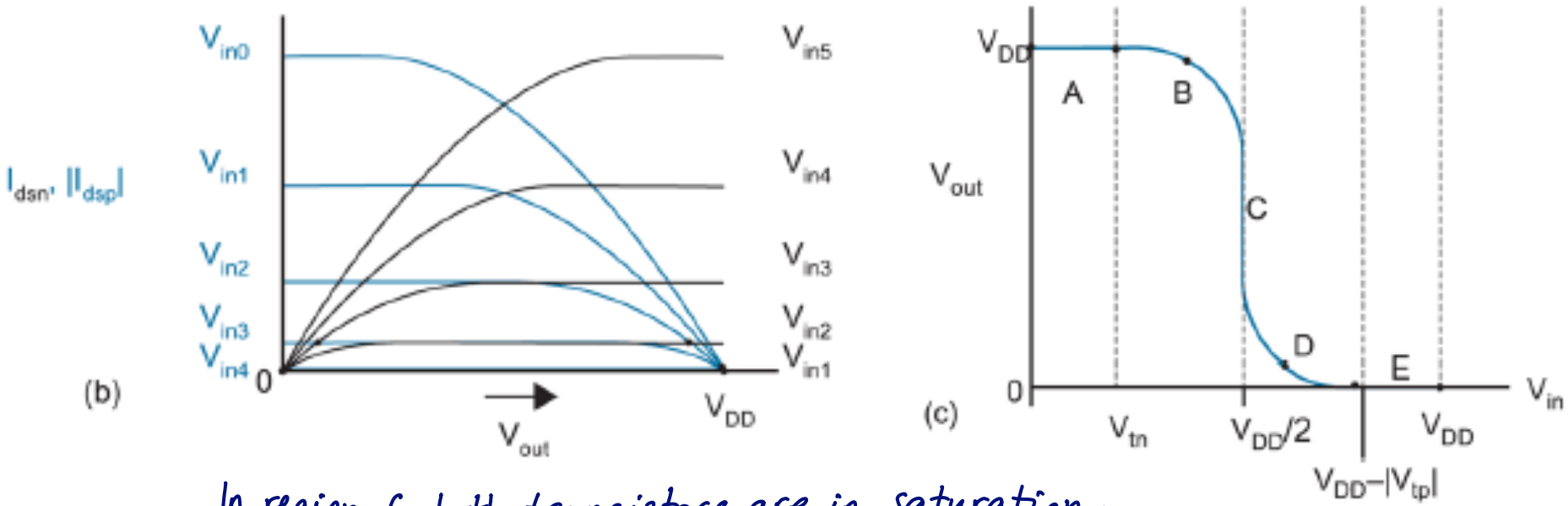
Graphical derivation of the inverter DC response: Load Line Analysis

- $V_{in} = V_{DD}$



DC Transfer Curve

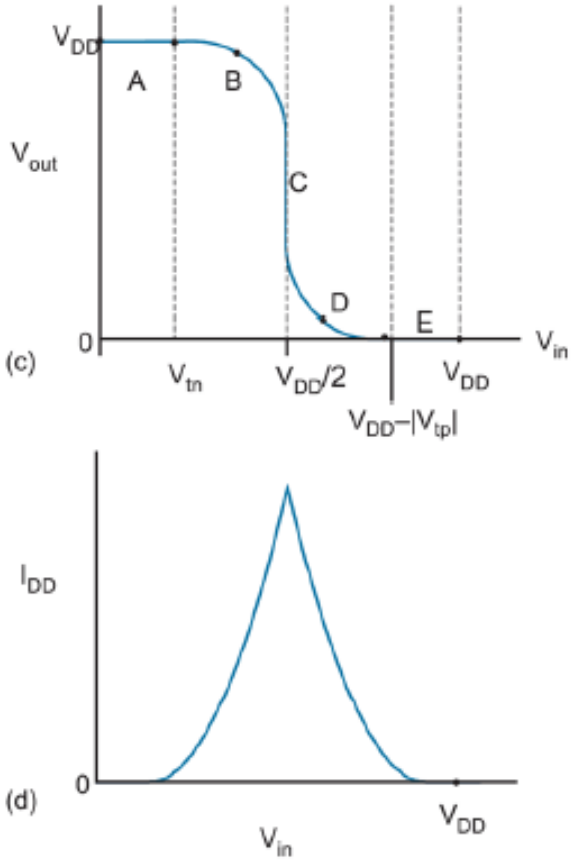
- Transcribe points onto V_{in} vs. V_{out} plot



In region C both transistors are in saturation.
Ideal transistors are only in region C for $V_{in} = \frac{V_{DD}}{2}$
and the DC curve slope in C is $-\infty$.

The crossover point where $V_{in} = V_{out}$ is called input threshold.

DC transfer curve: operating regions



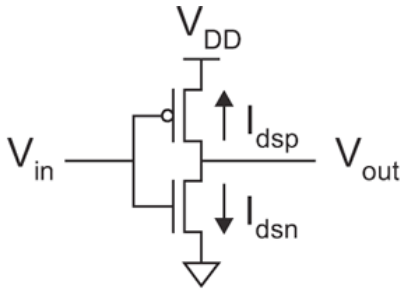


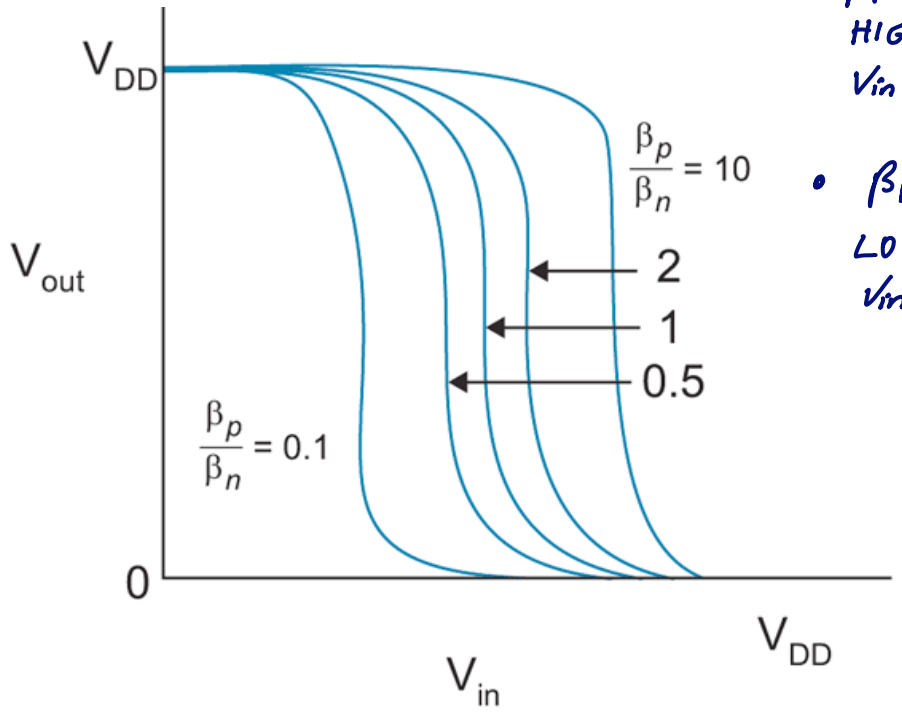
FIG 2.23 A CMOS inverter

Table 2.3 Summary of CMOS inverter operation

Region	Condition	p-device	n-device	Output
A	$0 \le V_{in} < V_{tn}$	linear	cutoff	$V_{out} = V_{DD}$
B	$V_{tn} \le V_{in} < V_{DD}/2$	linear	saturated	$V_{out} > V_{DD}/2$
C	$V_{in} = V_{DD}/2$	saturated	saturated	V_{out} drops sharply
D	$V_{DD}/2 < V_{in} \le V_{DD} - \lvert V_{tp}\rvert$	saturated	linear	$V_{out} < V_{DD}/2$
E	$V_{in} > V_{DD} - \lvert V_{tp}\rvert$	cutoff	linear	$V_{out} = 0$

Beta Ratio

- If $\beta_p / \beta_n \neq 1$, switching point will move from $V_{DD}/2$
- Called *skewed* gate



- $\beta_p > \beta_n$
  HIGH SKEWED inverter
  $V_{in} = \frac{V_{DD}}{2} \rightarrow V_{out} > \frac{V_{DD}}{2}$
- $\beta_p < \beta_n$
  LOW SKEWED inverter
  $V_{in} = \frac{V_{DD}}{2} \rightarrow V_{out} < \frac{V_{DD}}{2}$

FIG 2.26 Transfer characteristics of skewed inverters

Noise Margins

- How much noise can a gate input see before it does not recognize the input ?

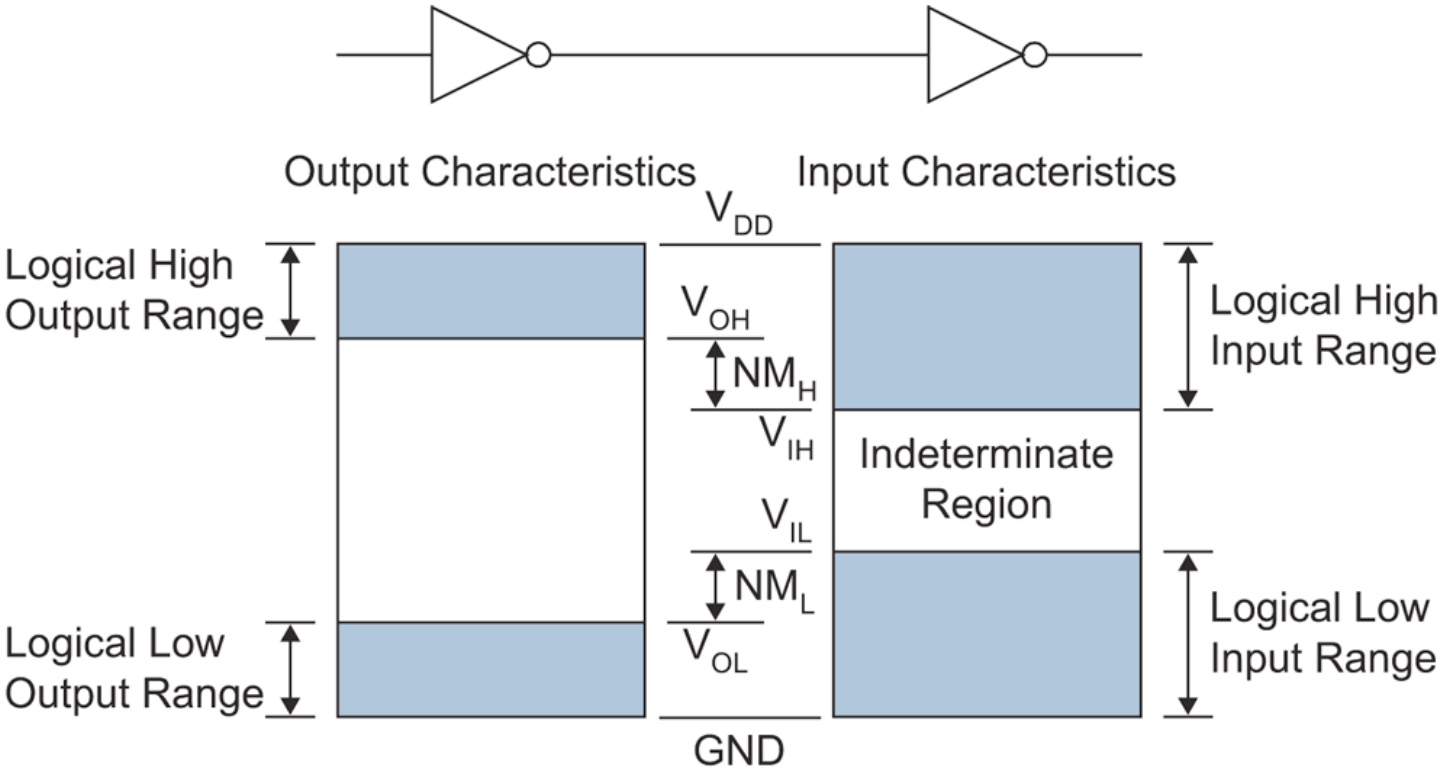


FIG 2.27 Noise margin definitions

Noise Margins

- To maximize noise margins, select logic levels at unity gain point of DC transfer characteristic

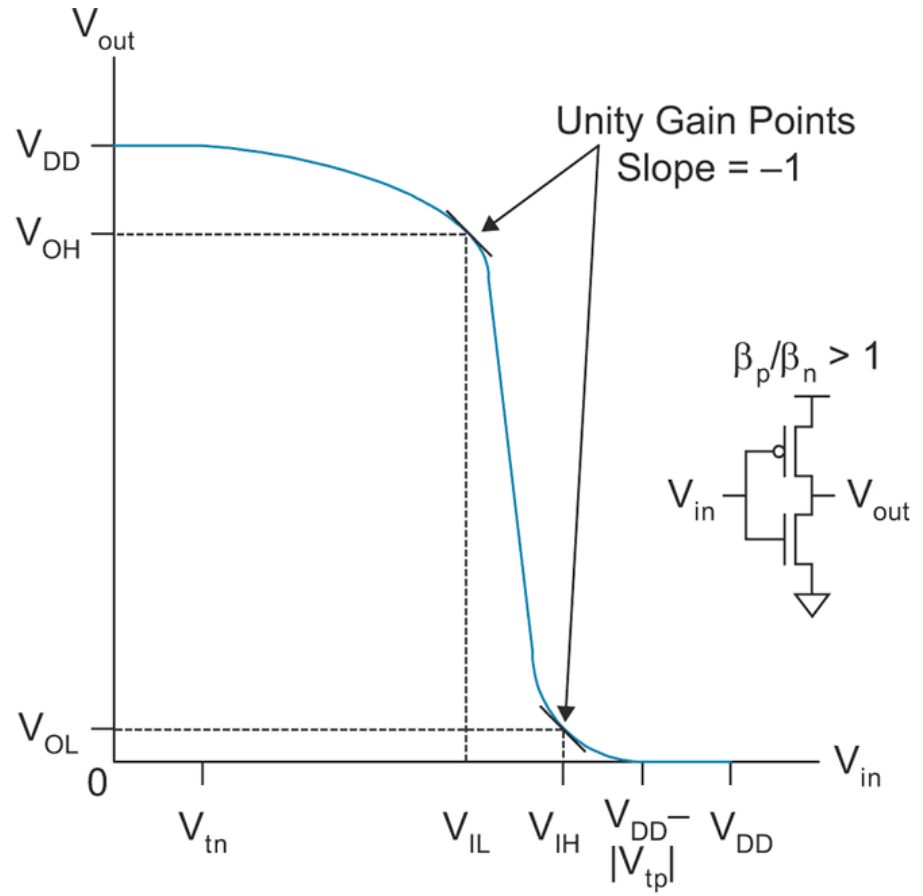


FIG 2.28 CMOS inverter noise margins

DC parameters

- Input switching threshold: V_{TH}
- Minimum high output voltage: V_{OH}
- Maximum low output voltage: V_{OL}
- Minimum HIGH input voltage: V_{IH}
- Maximum LOW input voltage: V_{IL}

symmetry of the logic gate behavior

quality of the logic values provided by the gate

ability of the logic gate to tolerate noise

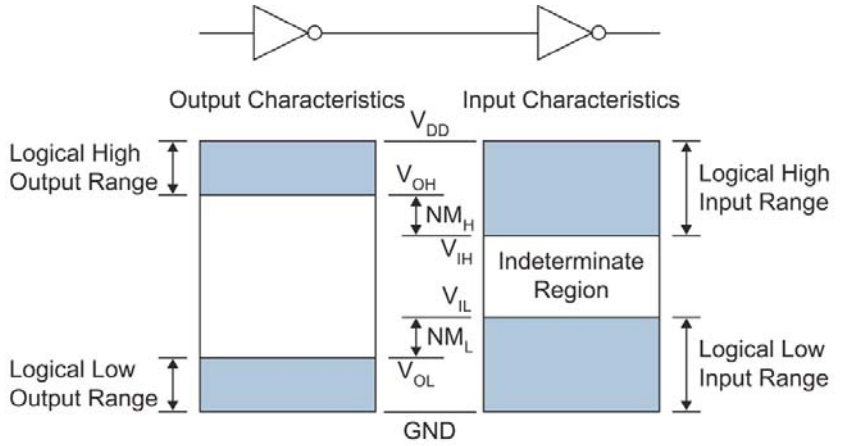


FIG 2.27 Noise margin definitions

CMOS Inverter Dynamic Behavior: AC Analysis

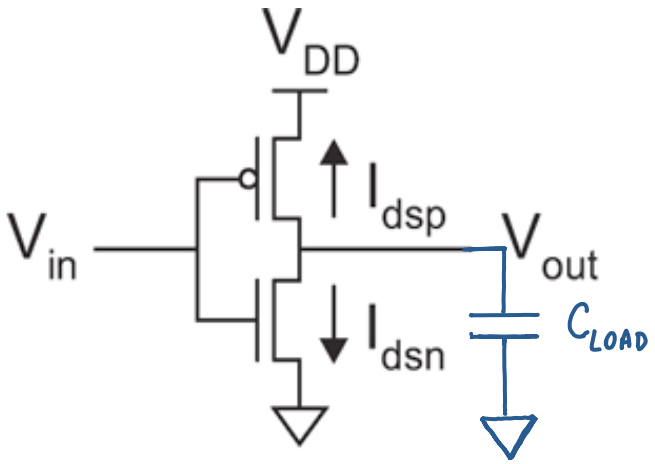


FIG 2.23 A CMOS inverter

- DC analysis tells Vout if Vin is constant
- AC analysis tells Vout(t) if Vin(t) changes
  - Requires solving differential equations
- Input is usually considered to be a step or ramp from 0 to VDD or vice versa

AC analysis = transient analysis = switching analysis = dynamic analysis

CMOS Inverter Dynamic Behavior: AC Analysis

- The switching characteristic (Vout(t) given Vin(t)) of a logic gate tells the speed at which the gate can operate
- The switching speed of a logic gate can be measured in terms of the time required to charge and discharge a capacitive load
- Critical paths
- Timing Analyzers automatically finds the slowest paths in a logic design
- Critical paths can be affected at various levels:
  - Architecture/ Micro-architecture Level
  - Logic Level
  - Circuit Level
  - Layout level

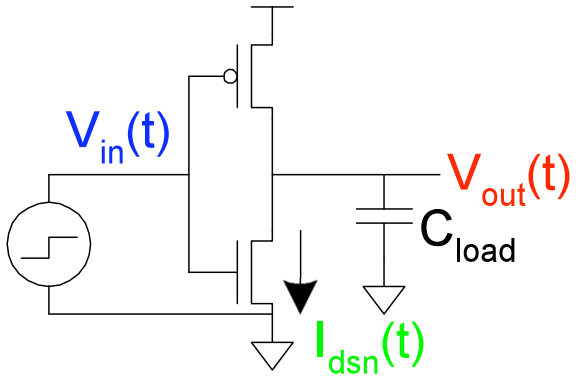
Inverter Step Response

- Find step response of inverter driving load cap

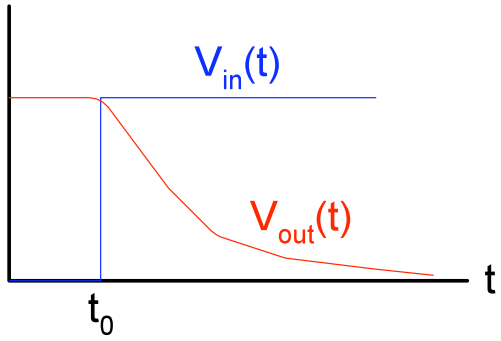
$$V_{in}(t) = u(t - t_0)V_{DD}$$

$$V_{out}(t < t_0) = V_{DD}$$

$$\frac{dV_{out}(t)}{dt} = -\frac{I_{dsn}(t)}{C_{load}}$$



$$I_{dsn}(t) = \begin{cases} 0 & t \le t_0 \\ \frac{\beta}{2}(V_{DD} - V_t)^2 & V_{out} > V_{DD} - V_t \\ \beta\left(V_{DD} - V_t - V_{out}(t)/2\right)V_{out}(t) & V_{out} < V_{DD} - V_t \end{cases}$$



Delay Parameters

- **t_r: rise time**
  - From output crossing 0.2 V_{DD} to 0.8 V_{DD}
- **t_f: fall time**
  - From output crossing 0.8 V_{DD} to 0.2 V_{DD}
- **t_{pdr}: rising propagation delay**
  - From input crossing $V_{DD}/2$ to rising output crossing $V_{DD}/2$
- **t_{pdf}: falling propagation delay**
  - From input crossing $V_{DD}/2$ to falling output crossing $V_{DD}/2$
- **t_{pd}: average propagation delay**
  - $t_{pd} = (t_{pdr} + t_{pdf})/2$

Delay Parameters cont.

- t_r, t_f
  - Tells how steep can be the waveform that the logic gate is able to provide at its output
- t_{pdr}, t_{pdf}
  - Input-to-output delay of the logic gate (time needed for the output to respond to a change in the input)

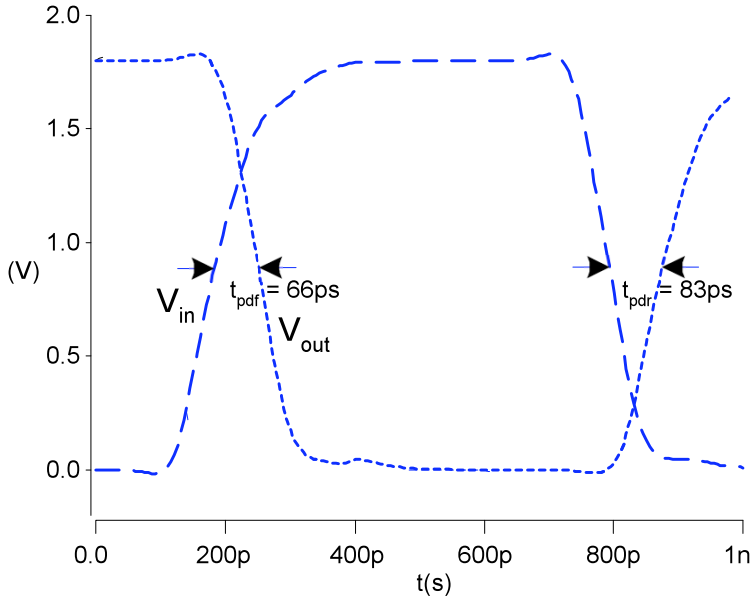
Factors affecting delay

- C_{LOAD} (= $C_{intrinsic}$ + $C_{extrinsic}$)
  - intrinsic capacitance
    (parasitic capacitance of the driving logic gate)
  - extrinsic capacitance
    (interconnect capacitance + capacitance of the stage driven)

- Slope of the input waveform
  - As the voltage on the gate terminal of a transistor change so does its capacitance

Simulated Inverter Delay

- Solving differential equations by hand is too hard
- SPICE simulator solves the equations numerically
  - Uses more accurate I-V models too!
- But simulations take time to write
- It is important to develop back of the envelope techniques to rapidly estimate delay, understand its origin, and figure out how it can be reduced



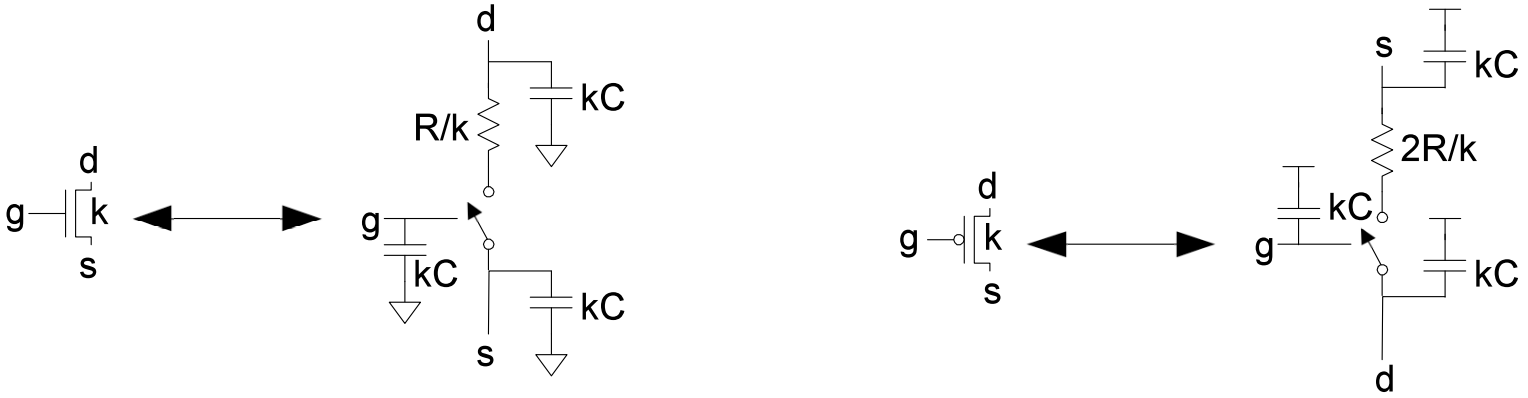
Delay Estimation (***Revisited***)

- We would like to be able to easily estimate delay
  - Not as accurate as simulation
  - But easier to ask "What if ?"
- The step response usually looks like a 1st order RC response with a decaying exponential.
- Use RC delay models to estimate delay
  - C = total capacitance on output node
  - Use *effective resistance* R
  - So that t_{pd} = RC
- Characterize transistors by finding their effective R
  - Depends on average current as gate switches

$$R = \ln 2 \cdot \Re$$

RC Delay Models

- Use equivalent circuits for MOS transistors
  - Ideal switch + capacitance and ON resistance
  - Unit nMOS has resistance R, capacitance C
  - Unit pMOS has resistance 2R, capacitance C
- Capacitance proportional to width
- Resistance inversely proportional to width



Power Dissipation

- Static CMOS gates are very power-efficient because they dissipate nearly zero power while idle

- Instantaneous power: $P = i_{DD}(t) \cdot V_{DD}$

- Energy consumed: $E = \int_0^T i_{DD} \cdot V_{DD} dt$

- Average power: $P_{avg} = \frac{1}{T} \int_0^T i_{DD} \cdot V_{DD} dt$

Power Dissipation

- Power dissipation in CMOS circuits comes from two components:
  - Static Dissipation
    - Subtreshold conduction
    - Tunneling current
    - Leakage through reverse biased diodes
  - Dynamic Dissipation
    - Charging and discharging (switching) of the load capacitance
    - “Short-Circuit” current while both pMOS and nMOS networks are partially ON

Static Dissipation

Assume $I_{leakage} \approx$ constant

$$P_{static} = V_{DD} \cdot I_{leakage}$$

- OFF transistors still conduct a certain amount of current :
  - Sub threshold current
  - Current through reverse biased diodes
  - gate tunneling current

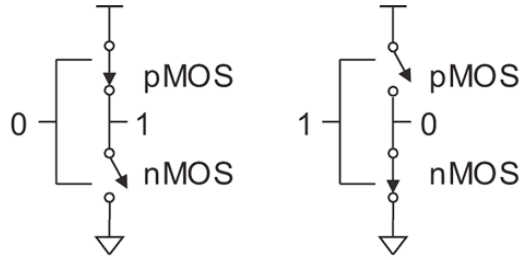


FIG 4.26 CMOS inverter model for static power dissipation evaluation

- In 130 nm processes and beyond leakage is becoming a major design issue and vendors now provide leakage data (often in the form of nA/μm of gate length)

Dynamic Dissipation

$$P_{dynamic} = P_{sw} + P_{sc} = \frac{1}{T}\int_0^T i_{DD}(t) \cdot V_{DD} dt = \frac{V_{DD}}{T}\int_0^T i_{DD}(t) dt$$

Assuming a logic gate goes through one complete charge/discharge cycle every clock cycle:

$$P_{sw} = C \cdot V_{DD}^2 \cdot f_{clock}$$

Because most gates do not switch every clock cycle, we introduce a corrective activity factor α:

$$P_{sw} = \alpha \cdot C \cdot V_{DD}^2 \cdot f_{clock}$$

A clock has α=1 because it rises and fall every cycle, but most data have a maximum activity factor α=0.5 because they transition only once every cycle

Dynamic Dissipation

- Because, input rise/fall time is greater than zero, both nMOS and pMOS will be ON for a short period of time (while the input is between V_{tn} and VDD-$|V_{tp}|$)
- This results in a “short-circuit” current pulse from VDD to GND
- Typically this increases power dissipation by about 10%

Low Power Design

- Power Dissipation is a major problem !!!

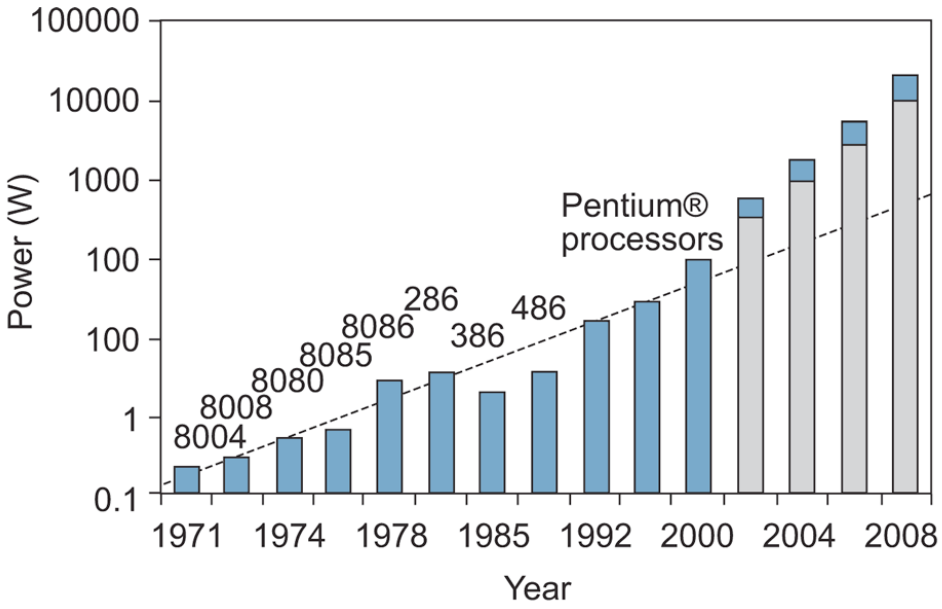


FIG 4.69 Intel processor power consumption. © IEEE 2001.

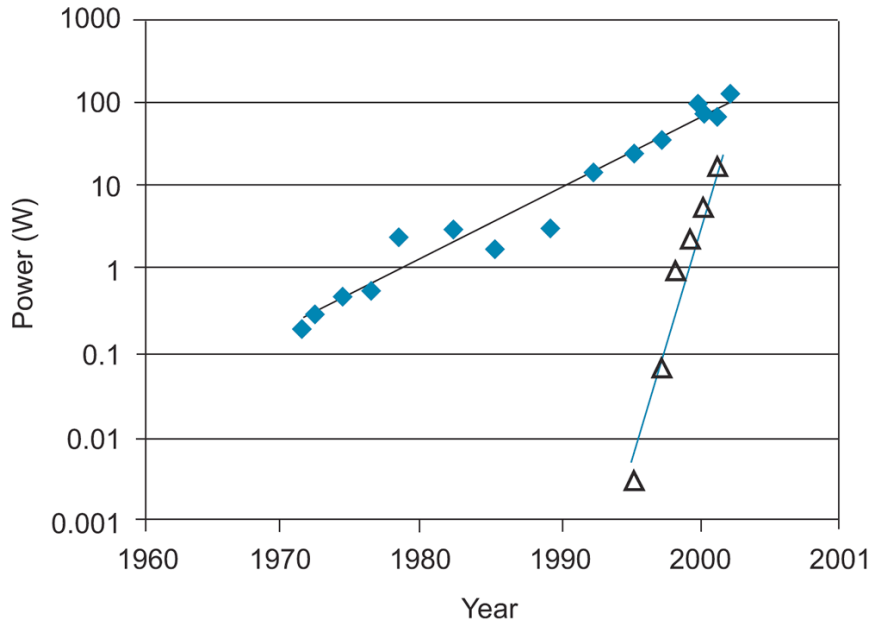


FIG 4.70 Dynamic and static power trends. © IEEE 2003.

Dynamic Power Reduction

- Decrease activity factor
  - Selective clock gating
  - Drawback: if the system transitions rapidly from an idle mode to a fully active mode a large di/dt spike will occur
- Decrease switching capacitance
  - Small transistors
  - Careful floor planning to reduce interconnect
- Decrease power supply
  - Adjust voltage depending on the operating mode
- Decrease operating frequency

Static Power Reduction

- Subthreshold current can be reduced by increasing V_t
  - Selective application of multiple threshold (low-V_t transistors on critical paths, high V_t transistors on other paths)
  - Control V_t through the body voltage

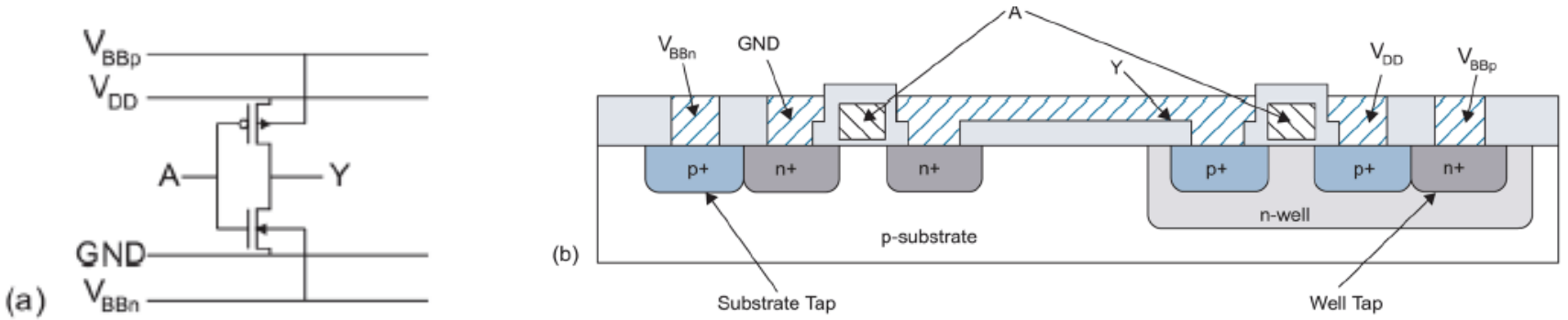


FIG 4.27 Body bias

Static Power Reduction cont.

- Turn off the power supply entirely. MTCMOS circuits use low V_t transistors for computation and high V_t transistor as a switch to disconnect the power supply during idle mode

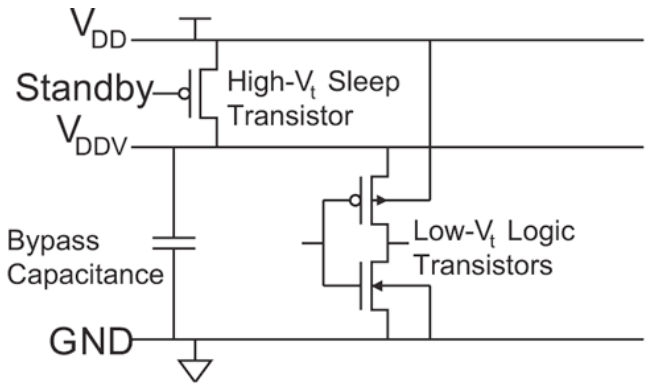


FIG 4.28 MTCMOS

- The leakage through two series OFF transistor is much lower (10-20x) than that of a single transistor (stack effect)

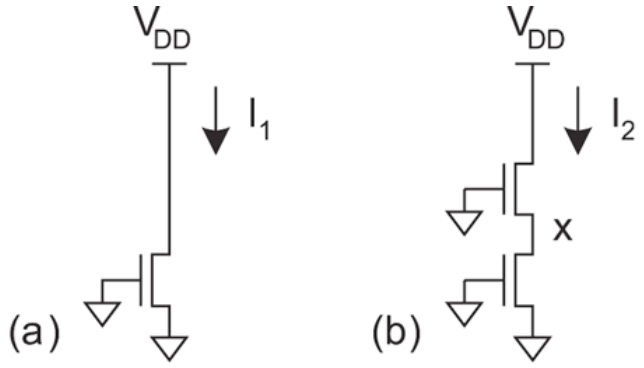


FIG 4.29 Leakage stack effect

Energy Delay Product (1/2)

- Delay is proportional to CV_{DD}/I_{dsat} ($\tau = Q/I$)

$$\tau = \kappa \frac{C \cdot V_{DD}}{(V_{DD} - V_T)^{\eta}}$$

κ reflect technology parameters

η *account for velocity saturation*

- Switching energy is proportional to CV_{DD}^2

$$E_{sw} = \alpha \cdot C \cdot V_{DD}^2$$

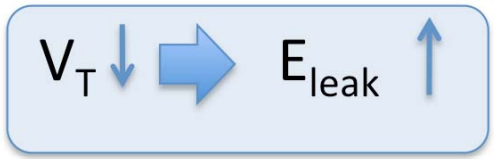
Energy Delay Product (2/2)

$$EDP = K \frac{C^2 \cdot V_{DD}^3}{(V_{DD} - V_T)^\eta}$$

- **BUT** … unfortunately there is also the leakage energy:

$$E_{leak} = I_{leak} \cdot \frac{V_{DD}}{f} = \underbrace{I_{off} \cdot e^{\frac{Vgs - V_T}{n v_{TH}}} (1 - e^{-Vds/v_{TH}})}_{} \cdot \frac{V_{DD}}{f}$$

Source: Kaushik Roy



I_{off} is the subthreshold current at Vgs=0 and Vds=V_{DD}
v_{TH} is the thermal voltage