

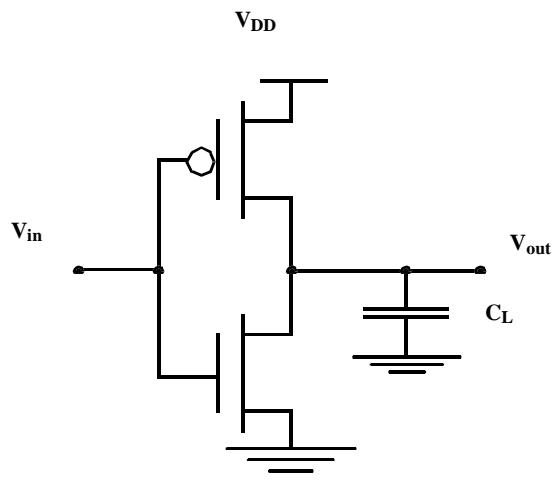
# CMOS INVERTER

## The CMOS Inverter – Static Model

### Outline

- First Glance
- Digital Gate Characterization
- Static Behavior (Robustness)
  - VTC
  - Switching Threshold
  - Noise Margins

# The CMOS Inverter: A First Glance

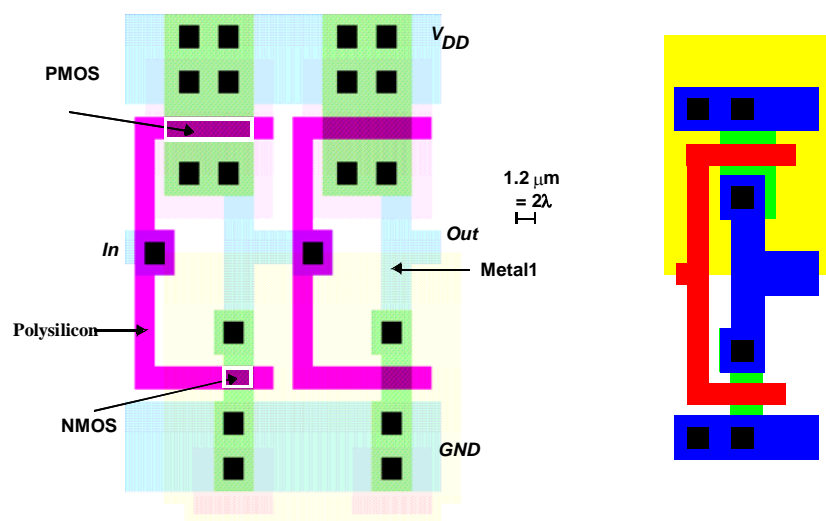


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

3

# CMOS Inverters (1)

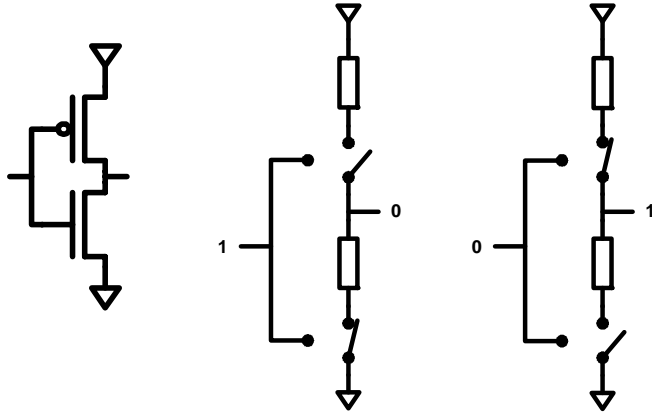


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

4

## CMOS Inverter Operation Principle



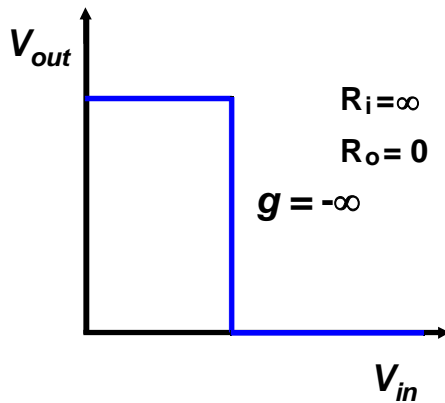
$$V_{OH} = V_{DD} \quad V_{OL} = 0$$



## Digital Gate Fundamental Parameters

- **Functionality**
- **Reliability, Robustness**
- **Area**
- **Performance**
  - **Speed (delay)**
  - **Power Consumption**
  - **Energy**

## The Ideal Inverter

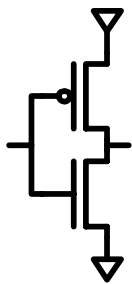


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

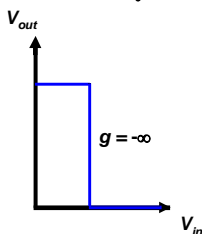
7

## Static CMOS Properties



Basic inverter belongs to class of **static circuits**: output always connected to either  $V_{DD}$  or  $V_{SS}$ . **Not ideal but:**

- Rail to rail voltage swing
- Ratio less design
- Low output impedance
- Extremely high input impedance
- No static power dissipation
- Good noise properties/margins



**Exercise:** prioritize the list above

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

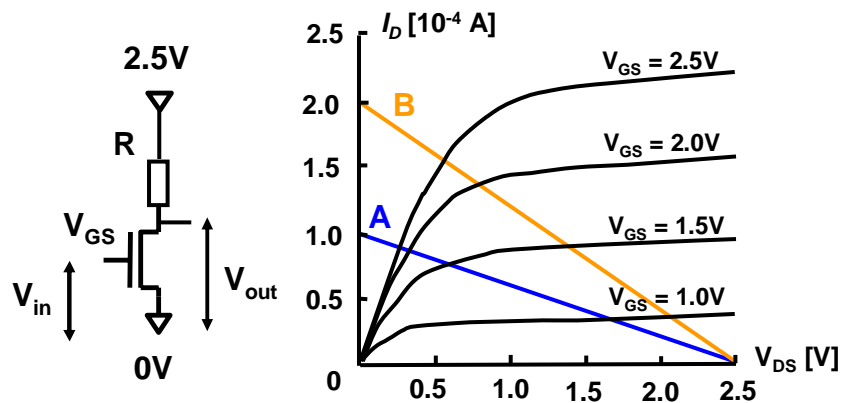
3/15/2009

8

## Voltage Transfer Characteristic (VTC)



## Load Line (Ckt Theory)



### Exercise:

The **blue** load line A corresponds to  $R =$

The **orange** load line B corresponds to  $R =$

With load line A and  $V_{GS} = 1V$ ,  $V_{out} =$

Draw a graph  $V_{out}(V_{in})$  for load line A and B

# PMOS Load Lines

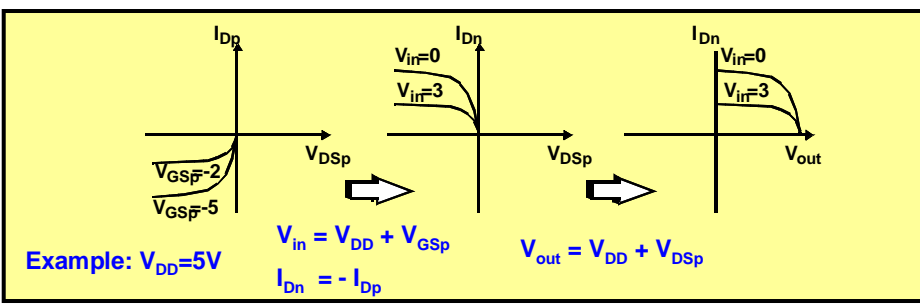
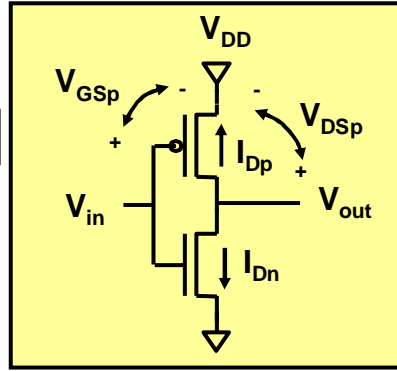
**Goal:** Combine  $I_{Dn}$  and  $I_{Dp}$  in one graph

**Kirchoff:**

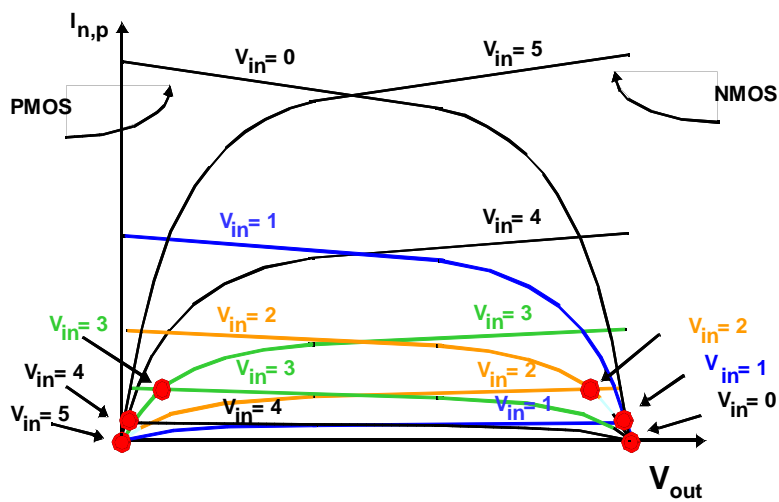
$$V_{in} = V_{DD} + V_{GSp}$$

$$I_{Dn} = -I_{Dp}$$

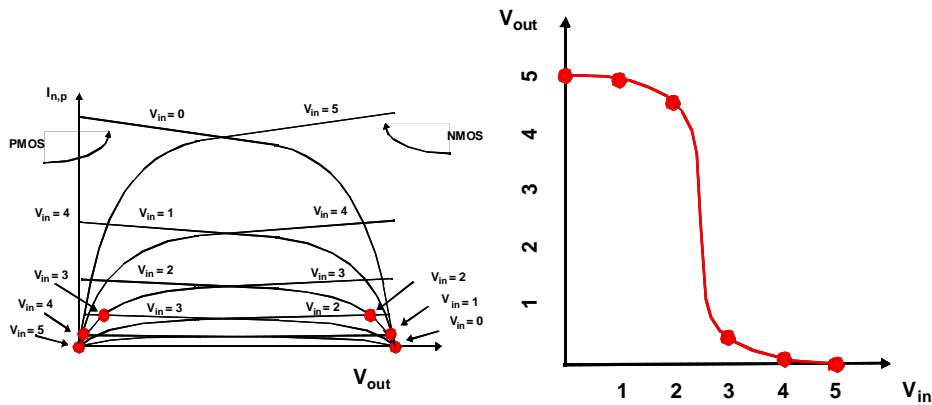
$$V_{out} = V_{DD} + V_{DSp}$$



# CMOS Inverter Load Characteristics



## CMOS Inverter VTC



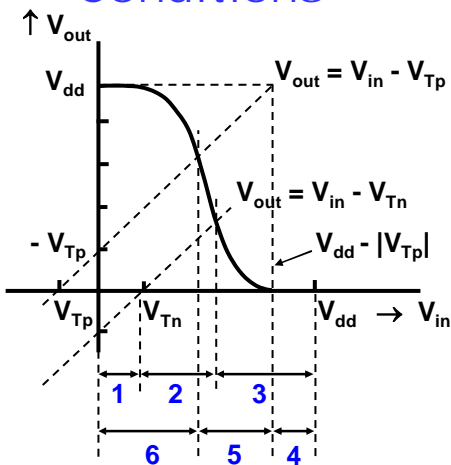
TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

13

## Operating Conditions

Need to know for proper dimensioning, analysis of noise margin, etc.



### NMOS

- 1  $V_{in} = V_{GS} < V_{Tn} \Rightarrow$  **off**
- 2  $V_{out} > V_{in} - V_{Tn}$   
 $V_{DS} > V_{GS} - V_{Tn}$   
 $V_{GD} < V_{Tn} \Rightarrow$  **saturation**
- 3  $V_{out} < V_{in} - V_{Tn} \Rightarrow$  **resistive**

### PMOS

- 4  $V_{in} > V_{DD} + V_{Tp} \Rightarrow$  **off**
- 5  $V_{out} < V_{in} - V_{Tp} \Rightarrow$  **saturation**
- 6  $V_{out} > V_{in} - V_{Tp} \Rightarrow$  **resistive**

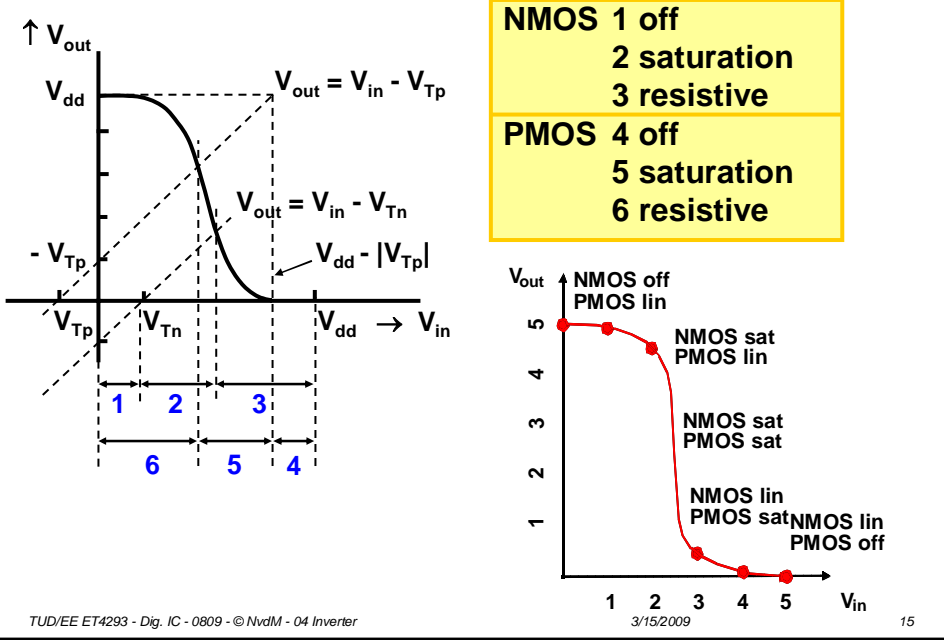
**Exercise: check results for PMOS**

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

14

## Operating Conditions



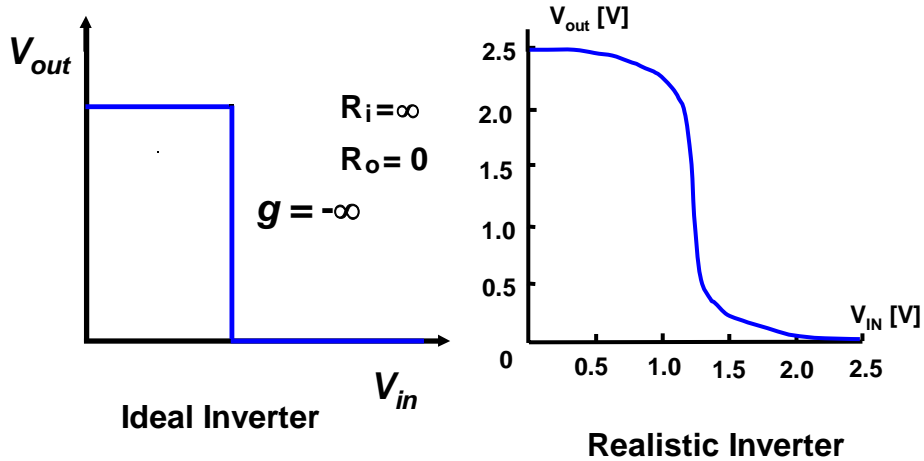
## Inverter Static Behavior

- Regeneration
- Noise margins
- Delay metrics





## The Realistic Inverter

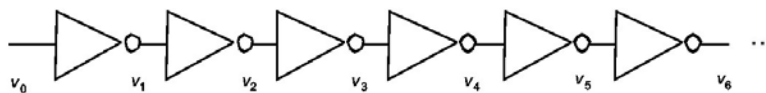


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

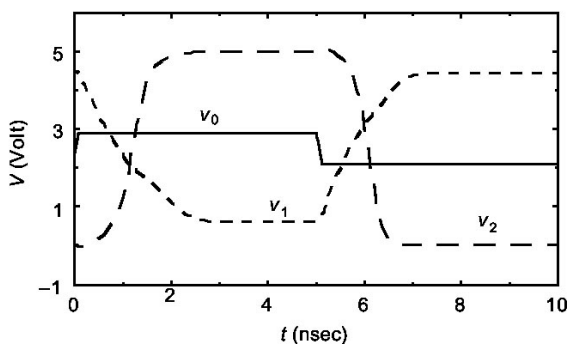
3/15/2009

17

## The Regenerative Property



A chain of inverters



■ **Regenerative Property:** ability to regenerate (repair) a weak signal in a chain of gates

The regenerative property

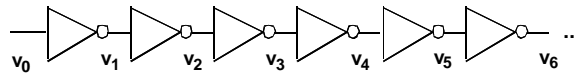


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

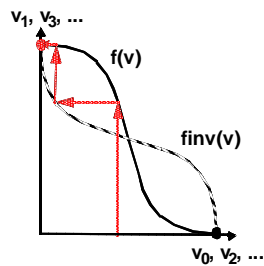
3/15/2009

18

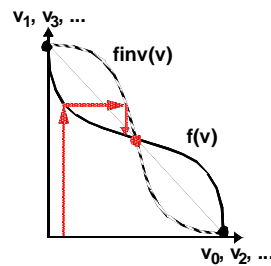
## The Regenerative Property (2)



(a) A chain of inverters.

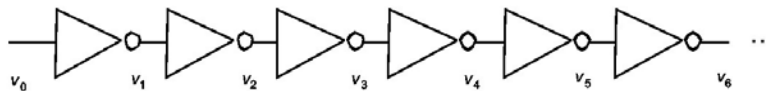


(b) Regenerative gate



(c) Non-regenerative gate

## The regenerative Property (3)

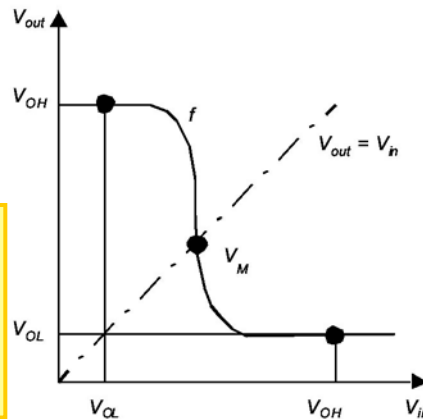
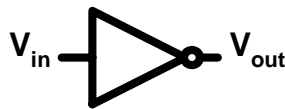


**Exercise:** what is the output voltage of a chain of 4 inverters with a piece-wise linear VTC passing through (0, 10), (3, 7), (7, 1) and (10, 0) [Volt], as the result of an input voltage of 6 [Volt].

**Exercise:** discuss the behavior for an input of 5 [Volt]

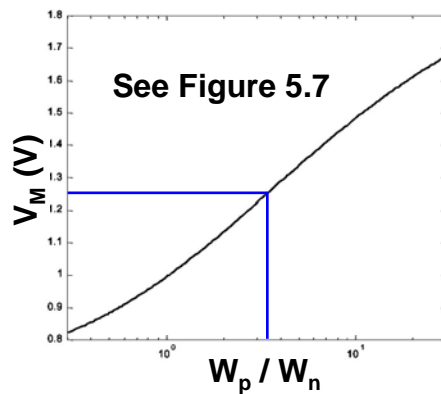
## Inverter Switching Treshold

- Not the device threshold  $V_m = f(R_{onn}, R_{onp})$
- Point of  $V_{in} = V_{out}$



- Try to set  $W_n, L_n, W_p, L_p$  so that VTC is symmetric as this will improve noise margins
- optimize NMOS-PMOS ratio

## Simulated Gate Switching Threshold



**Electrical Design Rule**  
 $W_p \approx 2.5 W_n$

- Assumes  $L_p = L_n$
- Should be applied consistently

- Symmetrical VTC  $\Rightarrow V_m \approx \frac{1}{2} V_{DD} \Rightarrow W_p/W_n \approx$
- In practice: somewhat smaller
- Why?

## Inverter Switching Threshold Analytical Derivation



- $V_M$  is  $V_{in}$  such that  $V_{in} = V_{out}$
- $V_{DS} = V_{GS} \Leftrightarrow V_{GD} = 0 \Rightarrow$  saturation
  - Assume  $V_{DSAT} < V_M - V_T$
  - (velocity saturation)
  - Ignore channel length modulation
- $V_M$  follows from
  - $I_{DSATn}(V_M) = -I_{DSATp}(V_M)$

## Inverter Switching Threshold Analytical Derivation (ctd)

$$I_{DSATn}(V_M) = -I_{DSATp}(V_M) \quad I_D = kV_{DSAT}(V_{GS} - V_T - V_{DSAT}/2)$$

$$\Leftrightarrow k_n V_{DSATn} (V_M - V_{Tn} - V_{DSATn}/2) = -k_p V_{DSATp} (V_M - V_{DD} - V_{Tp} - V_{DSATp}/2)$$

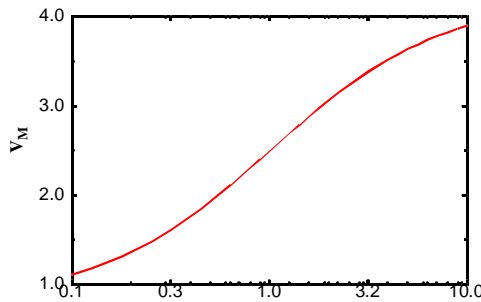
$$\Leftrightarrow \frac{k_p}{k_n} = \frac{-V_{DSATn} (V_M - V_{Tn} - V_{DSATn}/2)}{V_{DSATp} (V_M - V_{DD} - V_{Tp} - V_{DSATp}/2)} \quad k = \frac{W}{L} k'$$

$$\Rightarrow \frac{(W/L)_p}{(W/L)_n} = \frac{k'_n V_{DSATn} (V_M - V_{Tn} - V_{DSATn}/2)}{k'_p V_{DSATp} (V_M - V_{DD} - V_{Tp} - V_{DSATp}/2)}$$

- See Example 5.1:
- $(W/L)_p = 3.5 (W/L)_n$  for typical conditions and  $V_M = \frac{1}{2} V_{DD}$
- Usually:  $L_n = L_p$

## Gate Switching Threshold w/o Velocity Saturation

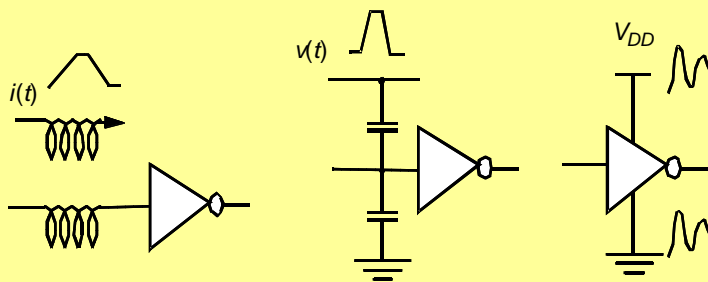
- Long channel approximation
- Applicable with low  $V_{DD}$



**Exercise (Problem 5.1):**  
derive  $V_M$  for long-channel approximation as shown below

$$V_M = \frac{r(V_{DD} - V_{Tp} + V_{Tn})}{1+r} \text{ with } r = \sqrt{\frac{-k_p}{k_n}}$$

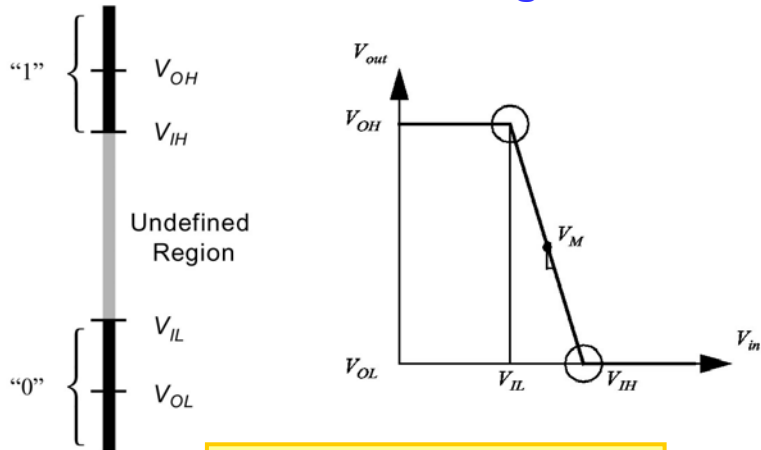
## Noise in Digital Integrated Circuits



(a) Inductive coupling    (b) Capacitive coupling    (c) Power and ground noise

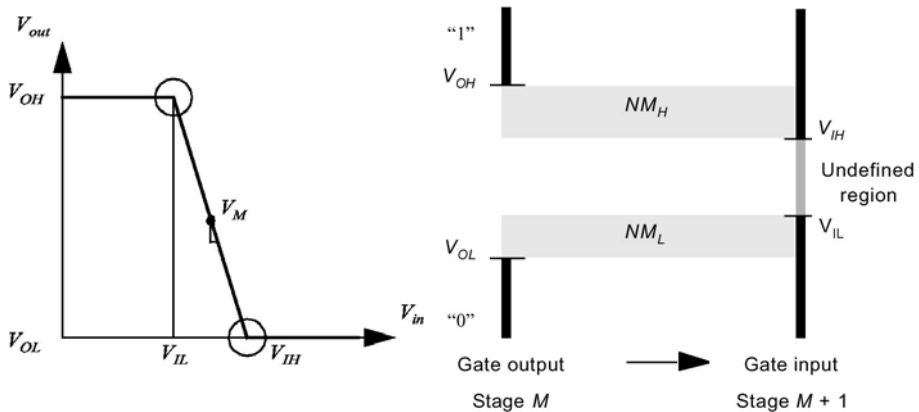
- Study behavior of static CMOS Gates with **noisy signals**

## Noise Margins



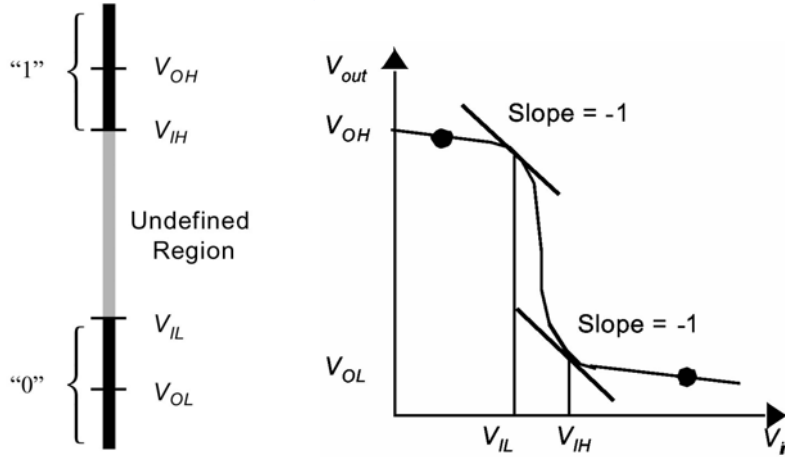
- $V_{OL}$  = Output Low Voltage
- $V_{IL}$  = Input Low Voltage
- $V_{OH}, V_{IH} = \dots$

## Noise Margins



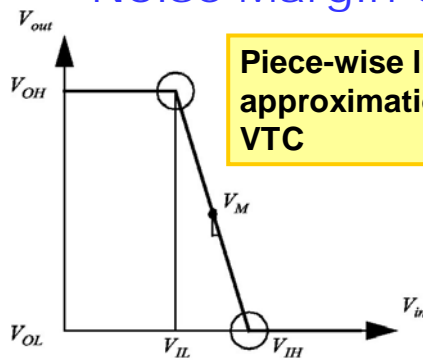
- $NM_H = V_{OH} - V_{IH} = \text{High Noise Margin}$
- $NM_L = V_{IL} - V_{OL} = \text{Low Noise Margin}$

## Noise Margin for Realistic Gates



**Exercise:** explain significance of slope = -1 for noise margin

## Noise Margin Calculation



**Piece-wise linear approximation of VTC**

**g = gain factor (slope of VTC)**

**We know how to compute  $V_M$**   
**Next: how to compute g**

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

$$V_{IH} = V_M - \frac{V_M}{g} \quad V_{IL} = V_M + \frac{V_{DD} - V_M}{g}$$

$$NM_H = V_{DD} - V_{IH} \quad NM_L = V_{IL}$$



## Noise Margin Calculation (2)

- Approximate  $g$  as the slope in  $V_{out}$  vs.  $V_{in}$  at  $V_{in} = V_M$

$$k_n V_{DSATn} (V_{in} - V_{Tn} - V_{DSATn} / 2)(1 + \lambda_n V_{out}) +$$

$$k_p V_{DSATp} (V_{in} - V_{DD} - V_{Tp} - V_{DSATp} / 2)(1 + \lambda_p V_{out} - \lambda_p V_{DD}) = 0$$

$$g = \left. \frac{dV_{out}}{dV_{in}} \right|_{V_{in}=V_M} \approx \frac{1+r}{(V_M - V_T - V_{DSATp} / 2)(\lambda_n - \lambda_p)} \quad r = \frac{k_p V_{DSATp}}{k_n V_{DSATn}}$$

- Mostly determined by technology
- See example 5.2

- **Exercise:** verify calculation
- **Exercise:** explain why we add channel length modulation to the  $I_D$  expressions (we did not do this to determine  $V_M$ )

## CMOS INVERTER dynamic behavior (performance)

- Capacitances
- (Dis)charge times
- Delay





■ Before: propagation delay analysis

$$t_p \approx 0.69 \times \frac{3}{4} \frac{V_{DD}}{I_{DSAT}} \left( 1 - \frac{5}{6} \lambda V_{DD} \right) C_L$$

■ Next: propagation delay from a design perspective inverter sizing



Reducing  $t_p$

$$\left. \begin{aligned} t_{pHL} &\approx 0.69 \times \frac{3}{4} \frac{V_{DD}}{I_{DSATn}} \left( 1 - \frac{5}{6} \lambda V_{DD} \right) C_L \\ I_{DSAT} &= k' \frac{W}{L} \left[ (V_{GS} - V_T) V_{DSAT} - \frac{1}{2} V_{DSAT}^2 \right] \end{aligned} \right\} \lambda=0 \rightarrow$$

$$t_{pHL} \approx 0.52 \frac{C_L V_{DD}}{(W/L)_n k_n' V_{DSATn} (V_{DD} - V_{Tn} - V_{DSATn}/2)}$$

Propagation Delay  $t_p$  can be reduced by

- Increasing  $V_{DD}$  (until  $V_{DD} \gg V_T + V_{DSAT}/2$ )
- Increasing  $W$
- Reducing  $C_L$

## Delay as a function of $V_{DD}$

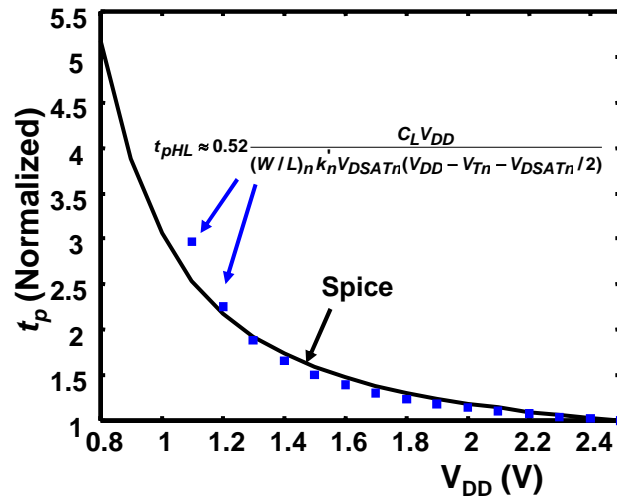


Fig.5.17

## Sizing

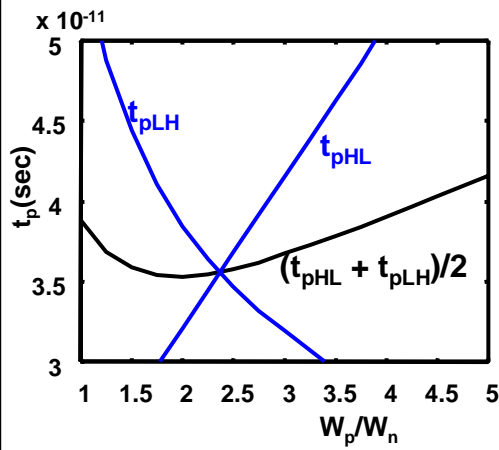
Propagation Delay  $t_p$  can be reduced by

- Increasing  $V_{DD}$  (until  $V_{DD} \gg V_T + V_{DSAT}/2$ )
- Increasing  $W$
- Reducing  $C_L$

$$t_{pHL} = 0.52 \frac{C_L V_{DD}}{(W/L)_n k_n V_{DSATn} (V_{DD} - V_{Tn} - V_{DSATn}/2)} \propto \frac{C_L}{W}$$

- $C_L$  can be reduced by good layout design
- But part of  $C_L$  depends on  $W$ !

## $t_p$ as a function of $W_p/W_n$

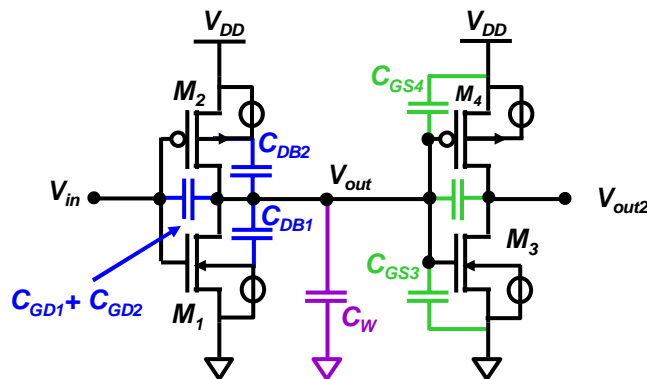


$$W_n \uparrow \begin{cases} \Rightarrow I_{pd} \uparrow \Rightarrow t_{pHL} \downarrow \\ \Rightarrow C_D \uparrow \Rightarrow t_{pLH} \uparrow \end{cases}$$

$$W_p \uparrow \begin{cases} \dots \\ \dots \end{cases}$$

Min  $t_p$  in general not when  $t_{pLH} = t_{pHL}$   
Save area, time at expense of robustness

## Intrinsic vs Extrinsic vs Parasitic Load Cap



$$C_{int} = C_{DB1} + C_{DB2} + 2(C_{GD1} + C_{GD2})$$

Intrinsic load

$$C_{ext} = C_{GS3} + C_{GS4} + C_{GD3} + C_{GD4}$$

Extrinsic / fan-out load

$$C_{par} = C_w$$

Parasitic load

## Isolated Inverter Sizing

$$t_p = 0.69 R_{eq} (C_{int} + C_{ext}) = 0.69 R_{eq} C_{int} \left( 1 + \frac{C_{ext}}{C_{int}} \right)$$

Assume  $C_{par}$  can be ignored or its effect can be absorbed in other C

$R_0$ : resistance of minimum size inverter (assume proper  $\beta = W_p / W_n$  ratio)

$C_0$ : intrinsic load (output, drain) cap of min. size inverter

$$t_{p0} = 0.69 R_0 C_0$$

intrinsic or unloaded delay

basic time constant for technology

minimum delay possible in technology given  $V_{DD}$

$S$ : sizing factor for  $W_n, W_p$  of driving inverter

$$W_n = S W_{min}, W_p = S \beta W_{min}$$

$$\rightarrow R_{eq} = R_0 / S$$

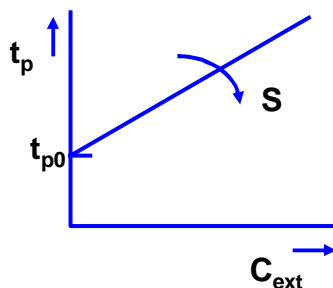
$$C_{int} = S C_0$$

$$\Rightarrow t_p = t_{p0} \left( 1 + \frac{C_{ext}}{S C_0} \right)$$

## Isolated Inverter Sizing

$$t_p = t_{p0} \left( 1 + \frac{C_{ext}}{S C_0} \right)$$

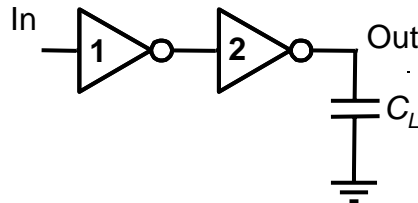
Increasing  $S$  reduces delay until  $S C_0 \gg C_{ext}$



## Inverter Chain

Assume size of inverter 1 is fixed.

- ☺ Increasing  $S$  of inverter 2 reduces  $t_p$  of inverter 2
- ☹ But it increases  $t_p$  of inverter 1 (higher load cap)
- ☺ Expect an optimum!



If  $C_L$  is given:

- How many stages are needed to minimize the delay?
- How to size the inverters?

## Delay Formula

$$t_p = t_{p0} \left( 1 + \frac{C_{ext}}{SC_0} \right) = t_{p0} \left( 1 + \frac{f}{\gamma} \right)$$

$C_{gin}$  input gate capacitance

$$\gamma = C_{int}/C_{gin} = SC_0/C_{gin}$$

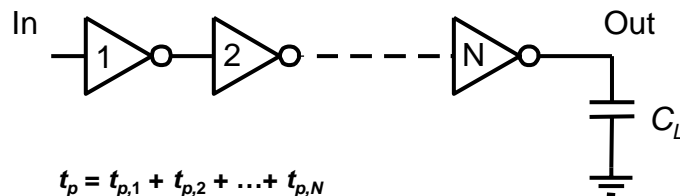
self loading coefficient

property of technology, typically  $\gamma \approx 1$

$f = C_{ext}/C_{gin}$  effective fanout

$$\frac{f}{\gamma} = \frac{C_{ext}}{C_{gin}} \times \frac{C_{gin}}{SC_0}$$

## Apply to Inverter Chain



$$t_p = t_{p,1} + t_{p,2} + \dots + t_{p,N}$$

$$t_{p,j} = t_{p0} \left( 1 + \frac{f_j}{\gamma} \right) = t_{p0} \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right)$$

$$t_p = \sum_{j=1}^N t_{p,j} = t_{p0} \sum_{j=1}^N \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right), \quad C_{gin,N+1} = C_L$$

## Apply to Inverter Chain

Delay equation has  $N-1$  unknowns,  $C_{gin,2} \dots C_{gin,N}$

$$t_p = \sum_{j=1}^N t_{p,j} = t_{p0} \sum_{j=1}^N \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right), \quad C_{gin,N+1} = C_L$$

Make  $N-1$  partial derivatives for  $C_{gin,j}$  zero for minimization:

$$\frac{\partial t_p}{\partial C_{gin,j}} = t_{p0} \left( \frac{1}{\gamma C_{gin,j-1}} - \frac{C_{gin,j+1}}{\gamma (C_{gin,j})^2} \right) = 0, \quad j = 2 \dots N-1$$

Size of each stage is geometric mean of 2 neighbors:

$$C_{gin,j} = \sqrt{C_{gin,j-1} \times C_{gin,j+1}}, \quad j = 2 \dots N-1$$

## Optimal Tapering for Given N

Size of each stage is geometric mean of 2 neighbors:

$$C_{gin,j} = \sqrt{C_{gin,j-1} \times C_{gin,j+1}}, \quad j = 2 \dots N-1$$

$$\Rightarrow C_{gin,j}^2 = C_{gin,j-1} \times C_{gin,j+1}$$

$$\Rightarrow \frac{C_{gin,j}}{C_{gin,j-1}} = \frac{C_{gin,j+1}}{C_{gin,j}}$$

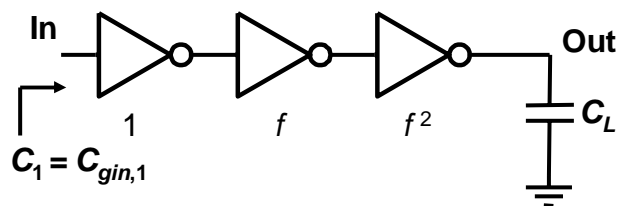
Load cap / input cap ratio  
same for each stage

$$\Rightarrow f_j = \frac{C_{gin,j+1}}{C_{gin,j}} = N \sqrt[N]{\frac{C_L}{C_{gin,1}}} = N \sqrt[N]{F}$$

$F = C_L / C_{gin,1}$ : path fan-out.  
Same fan-out, same delay for  
each stage.

$$\Rightarrow t_{p,j} = t_{p0} \left( 1 + \frac{f_j}{\gamma} \right) = t_{p0} \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right) = t_{p0} \left( 1 + \frac{N \sqrt[N]{F}}{\gamma} \right)$$

## Optimal Tapering for Fixed-N Summary



Delay per stage and total Path Delay

$$t_{p,j} = t_{p0} \left( 1 + \frac{f_j}{\gamma} \right) = t_{p0} \left( 1 + \frac{C_{gin,j+1}}{\gamma C_{gin,j}} \right) = t_{p0} \left( 1 + \frac{N \sqrt[N]{F}}{\gamma} \right)$$

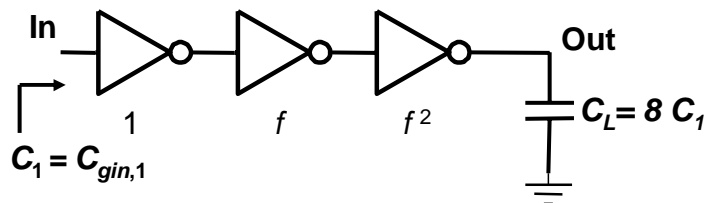
$$t_p = N t_{p0} \left( 1 + \frac{f_j}{\gamma} \right)$$

$$f_1 = f_2 = f_3 = \dots = F^{1/N}$$

$$f_1 \times f_2 \times f_3 \times \dots = F$$

$$F = C_L / C_{gin,1}$$

## Example



$C_L/C_1$  has to be evenly distributed across  $N = 3$  stages:

$$f = \sqrt[3]{8} = 2 \quad t_p = 3t_{p0} \left( 1 + \frac{\sqrt[3]{8}}{\gamma} \right) = 9t_{p0} \quad \text{for } \gamma = 1$$

## Optimum Number of Stages

For a given load,  $C_L$  and given input capacitance  $C_{in}$   
 find optimal  $f$  if  $N$  is free (and possibly non-integer)

$$C_L = F \cdot C_{in} = f^N C_{in} \quad \text{with } N = \frac{\ln F}{\ln f}$$

$$t_p = N t_{p0} \left( 1 + \frac{f}{\gamma} \right) = \frac{t_{p0} \ln F}{\gamma} \left( \frac{\gamma}{\ln f} + \frac{f}{\ln f} \right)$$

$$\frac{\partial t_p}{\partial f} = \frac{t_{p0} \ln F}{\gamma} \cdot \frac{-\gamma/f + \ln f - 1}{\ln^2 f} = 0$$

$$\ln f = 1 + \frac{\gamma}{f}$$

$$f = \exp \left( 1 + \frac{\gamma}{f} \right)$$

Closed-form solution  
 only for  $\gamma = 0$

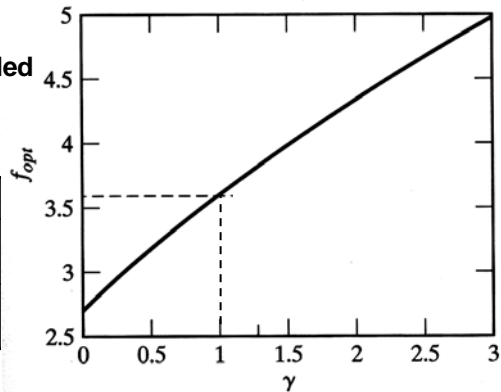


## Optimum Effective Fanout $f$ Optimum $f$ for given process defined by $\gamma$

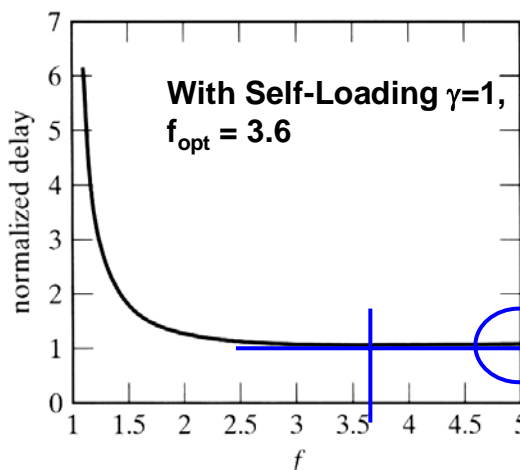
$$f = \exp\left(1 + \frac{\gamma}{f}\right) \quad N = \frac{\ln F}{\ln f}$$

(In practice, N must be rounded up or down to integer value)

	$\gamma=0$	$\gamma=1$
$f_{opt}$	$e=2.72$	3.6
$N_{opt}$	$\ln F$	$0.78 \ln F$



## Normalized $t_p$ vs. $f$



- Slight increase of  $t_p$  for  $f > f_{opt}$
- Choosing too few stages ( $f > f_{opt}$ ) is relatively harmless for delay and saves area
- Too many stages is expensive in terms of delay

**Fan-out of 4 (FO4) is safe common practice**

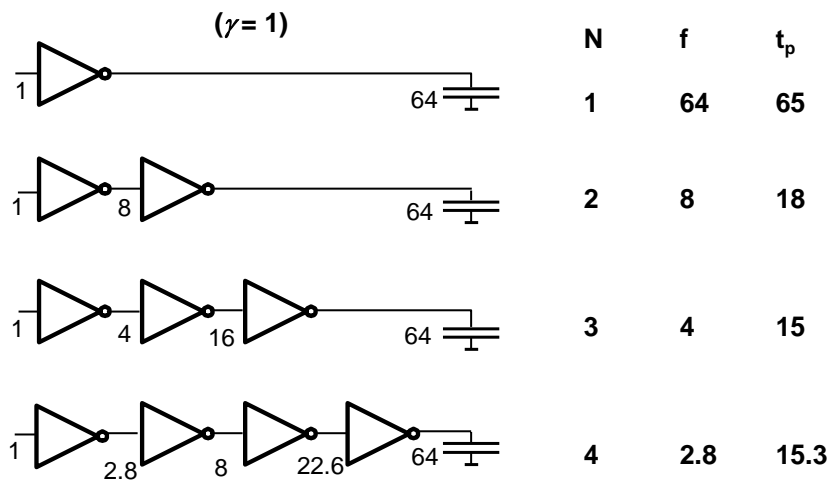
<http://en.wikipedia.org/wiki/FO4>

## Normalized delay function of $F$

$$t_p = Nt_{p0} \left( 1 + \frac{N\sqrt{F}}{\gamma} \right) \quad (\gamma = 1)$$

F	Unbuffered	Two Stage	Inverter Chain
10	11	8.3	8.3
100	101	22	16.5
1000	1001	65	24.8
10,000	10,001	202	33.1

## Buffer Design



## Power

- Dynamic Power
- Static Power
- Metrics



[www.quitpc.com](http://www.quitpc.com)



\$ 5.5



24 hours audio  
playback time

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

53

## CMOS Power Dissipation

- Power dissipation is a **very important** circuit characteristic
- CMOS has relatively low static dissipation
- Power dissipation was the reason that CMOS technology won over bipolar and NMOS technology for digital IC's
- (Extremely) high clock frequencies increase dynamic dissipation
- Low  $V_T$  increase leakage
- Advanced IC design is a continuous struggle to contain the power requirements!



\$ 1.3.4



\$ 5.5

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

54

## Power Density



### Estimate

- Furnace: 2000 Watt,  $r=10\text{cm}$  →  $P \approx 6\text{Watt/cm}^2$
- Processor chip: 100 Watt,  $3\text{cm}^2$  →  $P \approx 33\text{Watt/cm}^2$

**Power-aware design, design for low power, is blossoming subfield of VLSI Design**

## Where Does Power Go in CMOS

- **Dynamic Power Consumption**  
Charging and discharging capacitors
- **Short Circuit Currents**  
Short circuit path between supply rails during switching (NMOS and PMOS on together)
- **Leakage**  
Leaking diodes and transistors  
May be important for battery-operated equipment

## Dynamic Power

### Dynamic Power

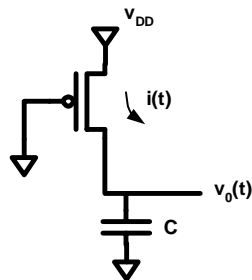
- $E_i$  = energy of switching event  $i$ 
  - independent of switching speed
  - depends on process, layout

- Power = Energy/Time

$$P = \frac{1}{T} \sum_i E_i$$

- $E_i$  = Power-Delay-Product P-D
  - important quality measure
- Energy-Delay-Product E-D
  - combines power\*speed performance

## Low-to-High Transition Energy



Equivalent circuit for **low-to-high transition**

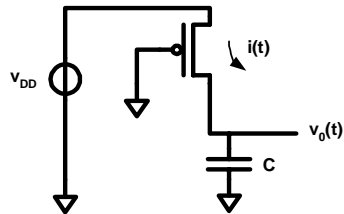
$E_C$  - Energy stored on C

$$E_C = \int_0^{\infty} i v_0 dt \quad v_0 = v_0(t) \quad i = i(t) = C \frac{dv_0}{dt}$$

$$= \int_0^{\infty} C v_0 \frac{dv_0}{dt} dt$$

$$= \int_0^{V_{DD}} C v_0 dv_0 = \frac{1}{2} C v_0^2 \Big|_0^{V_{DD}} = \frac{1}{2} C V_{DD}^2$$

## Low-to-High Transition Energy



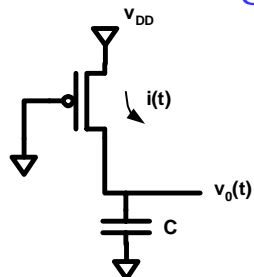
$E_{V_{DD}}$  Energy delivered by supply

$$E_{V_{DD}} = \int_0^{\infty} i(t) V_{DD} dt = \int_0^{V_{DD}} C V_{DD} \frac{dv_0}{dt} dt = C V_{DD}^2$$

$$E_{V_{DD}} = C V_{DD}^2 \quad E_c = \frac{1}{2} C V_{DD}^2$$

Where is the rest?

## Low-to-High Transition Energy

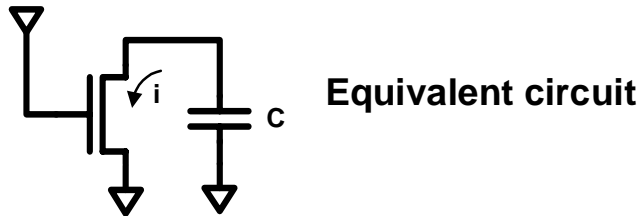


$E_{diss}$  Energy dissipated in transistor

$$\begin{aligned} E_{diss} &= \int_0^{\infty} i(V_{DD} - v_0) dt \\ &= \int_0^{\infty} i V_{DD} dt - \int_0^{\infty} i v_0 dt \\ &= E_{V_{DD}} - E_c \end{aligned}$$



## High-to-Low Transition Energy



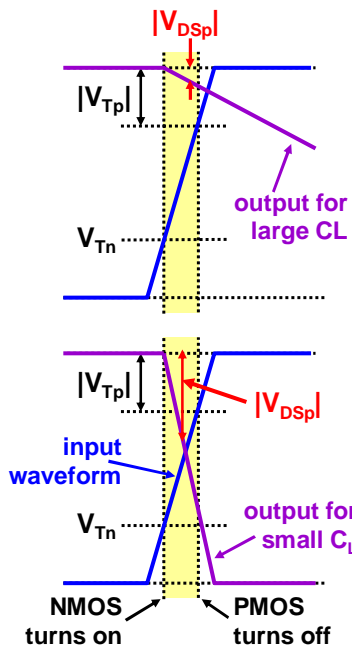
**Exercise:** Show that the energy that is dissipated in the transistor upon discharging C from  $V_{DD}$  to 0 equals  $E_{\text{diss}} = \frac{1}{2}CV_{DD}^2$

## CMOS Dynamic Power Dissipation

$$\begin{aligned} \text{Power} &= \frac{\text{Energy}}{\text{Time}} = \frac{\text{Energy}}{\text{transition}} \times \frac{\#\text{transitions}}{\text{time}} \\ &= CV_{DD}^2 \times f \end{aligned}$$

- Independent of transistor on-resistances
- Can only reduce C,  $V_{DD}$  or  $f$  to reduce power

## Short Circuit Current



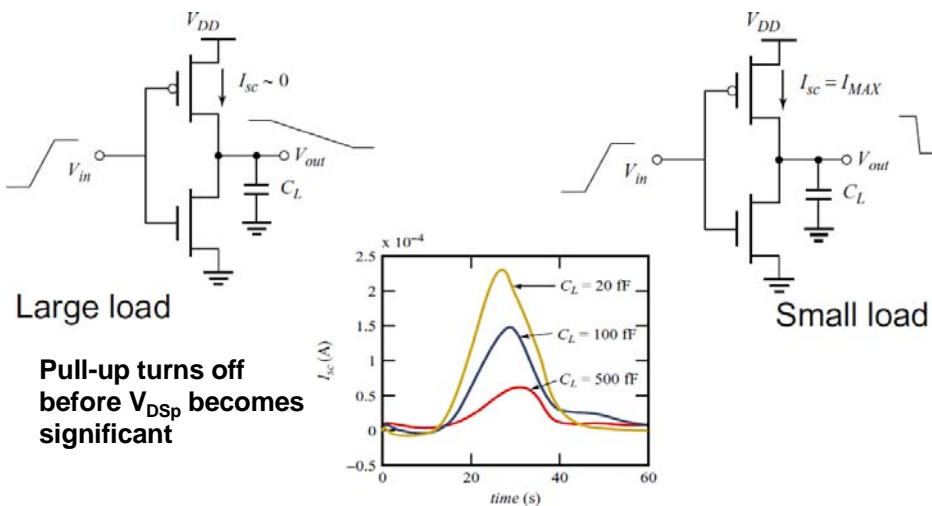
- **Input** and **output** waveforms of inverter loaded with a large capacitance (top) and with a small capacitance (bottom).
- **Shaded area** is where both pull-up and pull-down transistors are on (this is when short-circuit current can exist). This region is determined by crossings of input waveform with  $V_{Tn}$  and  $V_{DD}-|V_{Tpl}|$ .
- Short-circuit current increases with  $|V_{DSpl}|$ . This is clearly much larger for small  $C_L$  compared to large  $C_L$ .
- Similarly, short-circuit current can exist for low-to-high transition at output.

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

63

## Short Circuit Current



Large load

Pull-up turns off before  $V_{Dsp}$  becomes significant

Small load

Best to maintain approximately equal input/output slopes

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

64

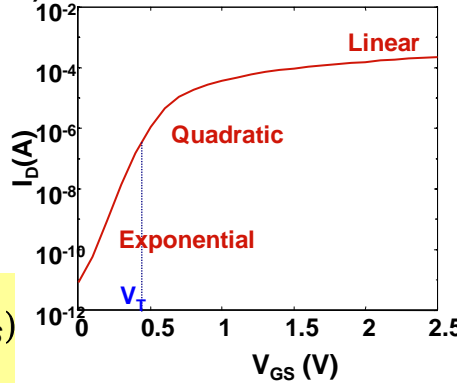


# Leakage

- Leakage current of reverse biased S/D junctions
- Sub-threshold current of MOS devices
- no channel → parasitic bipolar device:  
n+ (source) – p (bulk) – n+ (drain)
- Important source of leakage



$$I_D = I_0 e^{\frac{qV_{GS}}{nkT}} \left( 1 - e^{-\frac{qV_{DS}}{kT}} \right) (1 + \lambda \cdot V_{DS})$$



# Sub-Threshold Current

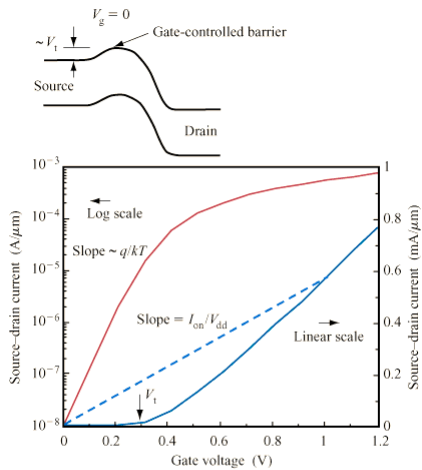


Figure 2

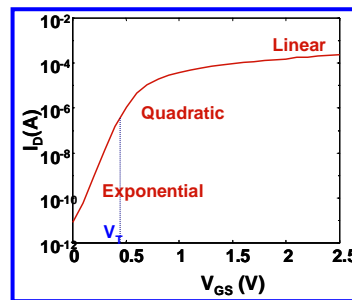
MOSFET current in both logarithmic (left) and linear (right) scales vs. gate voltage. The slope of the dotted line represents the large-signal transconductance for a digital circuit. Inset shows the band diagram of an n-MOSFET. The barrier height at  $V_g = 0$  is proportional to  $V_T$ .

- Rapidly becomes bottleneck with lowering threshold voltages
- Modern technologies offer low- $V_t$  and hi- $V_t$  devices  
Balance speed and power

Y. Taur, CMOS design near the limit of scaling, IBMJRD, Volume 46, Numbers 2/3, 2002

## Sub-Threshold Current

- Rapidly becomes bottleneck with lowering threshold voltages
- Modern technologies offer low- $V_t$  and hi- $V_t$  devices  
Balance speed and power

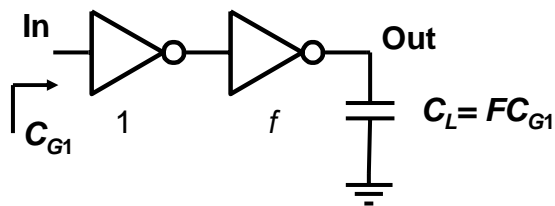


TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

67

## Transistor Sizing for Minimum Energy



$$t_{pHL} \approx 0.52 \frac{C_L V_{DD}}{(W/L)_n k_n V_{DSATn} (V_{DD} - V_{Tn} - V_{DSATn}/2)}$$



$$t_p = t_{p0} \left( \left( 1 + \frac{f}{\gamma} \right) + \left( 1 + \frac{F}{f\gamma} \right) \right)$$

$$t_{p0} \propto \frac{V_{DD}}{V_{DD} - V_{TE}}$$

See Eq. 5.21

$V_{TE} = V_T - \frac{1}{2} V_{DSAT}$  : Effective  $V_T$

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

68

## Transistor Sizing (2)

$$t_p = t_{p0} \left( \left( 1 + \frac{f}{\gamma} \right) + \left( 1 + \frac{F}{f\gamma} \right) \right) \quad t_{p0} \propto \frac{V_{DD}}{V_{DD} - V_{TE}}$$

Performance Constraint (with  $\gamma = 1$ ):

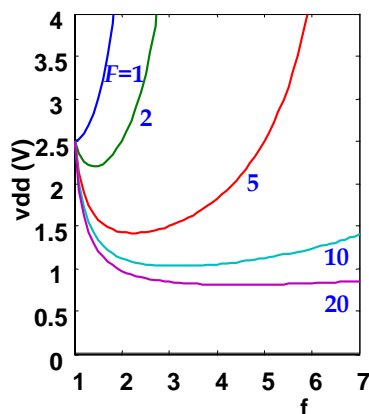
$$\frac{t_p}{t_{pref}} = \frac{t_{p0}}{t_{p0ref}} \frac{\left( 2 + f + \frac{F}{f} \right)}{(3 + F)} = \left( \frac{V_{DD}}{V_{ref}} \right) \left( \frac{V_{ref} - V_{TE}}{V_{DD} - V_{TE}} \right) \frac{\left( 2 + f + \frac{F}{f} \right)}{(3 + F)} = 1$$

- $V_{TE}$ : technology (0.5 V),  $V_{ref}$ : standard supply (2.5 V)
- F: fanout
- $V_{DD}$ , f: design parameters
- $V_{DD}$  is a function of f, given a fixed performance

$$V_{DD} = \frac{f(15 + 5F)}{(14f - 8f^2 - 8F + 10fF)}$$

## Transistor Sizing (3)

$V_{DD} = f(f)$



- Supply voltage needed as a function of f to maintain reference performance
- Lowest supply voltage needed for  $f = F^{0.5}$

$$V_{DD} = \frac{f(15 + 5F)}{(14f - 8f^2 - 8F + 10fF)}$$

## Transistor Sizing (4)

Energy for single Transition:

$$E = V_{DD}^2 C_{g1} [(1+\gamma)(1+f) + F]$$

$$\frac{E}{E_{ref}} = \left( \frac{V_{DD}}{V_{ref}} \right)^2 \left( \frac{2+2f+F}{4+F} \right)$$

$$V_{DD} = \frac{f(15+5F)}{(14f-8f^2-8F+10fF)}$$

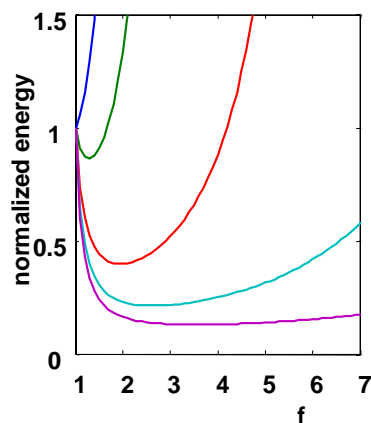
$C_{g1} + C_{int}$

Size of 1<sup>st</sup> + 2<sup>nd</sup> inverter

$C_L = FC_{G1}$

## Transistor Sizing (5)

$E/E_{ref} = f(f)$



$$\frac{E}{E_{ref}} = \left( \frac{V_{DD}}{V_{ref}} \right)^2 \left( \frac{2+2f+F}{4+F} \right)$$

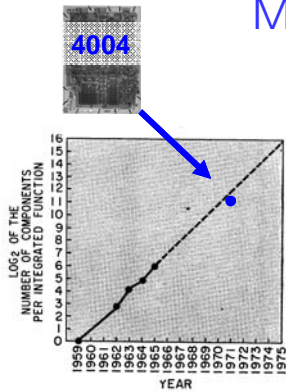
- Device sizing is effective
- Oversizing is expensive for power
- Optimal sizing for energy slightly different from sizing for performance

# Technology Scaling

Also see: IBM JRD, Vol 46, no 2/3, 2002  
Scaling CMOS to the limit  
<http://www.research.ibm.com/journal/rd46-23.html>



# Moore's Law



The number of transistors that can be integrated on a single chip will double every 18 months

Gordon Moore, co-founder of Intel  
[Electronics, Vol 38, No. 8, 1965]



## Why Scaling

- Reduce **price per function**:
  - Want to sell more functions (transistors) per chip for the same money → **better products**
  - Build same products cheaper, sell the same part for less money → **larger market**
  - Price of a transistor has to be reduced
- But also want to be **faster, smaller, lower power**

## Scaling Models

### Fixed Voltage Scaling

- most common model until 1990's
- only dimensions scale, voltages remain constant

### Full Scaling (Constant Electrical Field)

- ideal model — dimensions and voltage scale together by the same factor  $S$

### General Scaling

- most realistic for today's situation —
- voltages and dimensions scale with different factors

## Scaling for Velocity Saturated Devices

Constant Field Scaling:  $S = U$

Parameter	Relation	General Scaling
$W, L, t_{ox}$		$1/S$
$V_{DD}, V_T$		$1/U$
$N_{SUB}$	$V / W_{depl}^2$	$S^2/U$
<b>Area / Device</b>	<b>WL</b>	<b><math>1/S^2</math></b>
$C_{ox}$	$1/t_{ox}$	$S$
$C_{gate}$	$C_{ox} WL$	$1/S$
$k_n, k_p$	$C_{ox} W/L$	$S$
$I_{sat}$	$C_{ox} W V$	$1/U$
<b>Current Density</b>	<b><math>I_{sat} / \text{Area}</math></b>	<b><math>S^2/U</math></b>
$R_{on}$	$V / I_{sat}$	$1$
<b>Intrinsic Delay</b>	<b><math>R_{on} C_{gate}</math></b>	<b><math>1/S</math></b>
<b>Power / Device</b>	<b><math>I_{sat} V</math></b>	<b><math>1/U^2</math></b>
<b>Power Density</b>	<b><math>P / \text{Area}</math></b>	<b><math>S^2/U^2</math></b>

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

77

## IC Technology Scaling

Scaling improves density and performance

■ First order scaling theory		2008 / 1971
■ dimensions,	$1/S$	0.007
■ voltages	$1/S$	0.007
■ intrinsic delay	$1/S$	0.007
■ power per transistor	$1/S^2$	0.00004

### ■ Scaling trend



TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 04 Inverter

3/15/2009

78

# Technology Practice & ITRS

- **Scaling – Technology Generations**
- $S \approx 1.4 \approx 2^{0.5}$  per generation
- ... – 250 – 180 – 130 – 90 – 65 – 45 – 35 – 22 – ... nm
- **ITRS: International Technology Roadmap for Semiconductors**  
Industry-wide organization for forecasting technology developments – and (planning) requirements



<http://www.itrs.net/home.html>

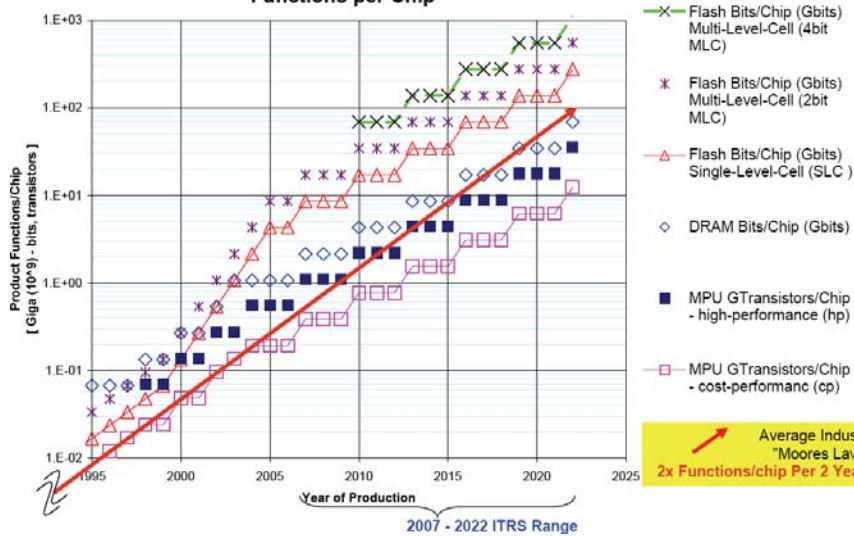


Not really – it is more like science  
(and a self-fulfilling prophecy at the same time)

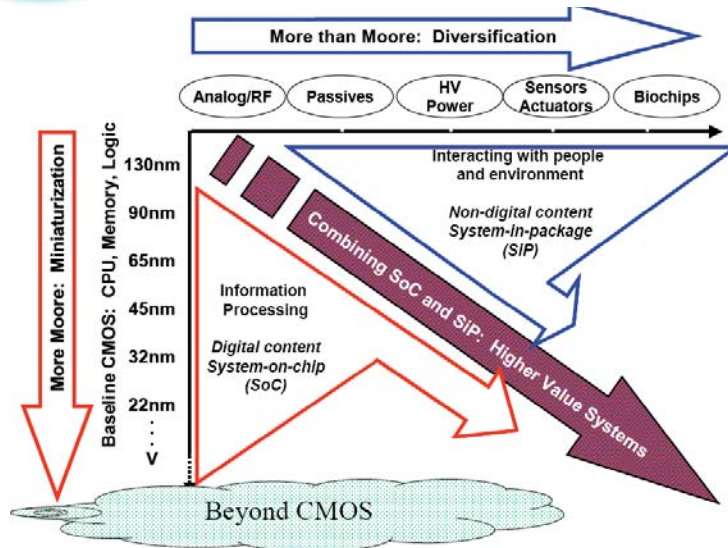


International Technology Roadmap for Semiconductors

2007 ITRS Product Technology Trends -  
Functions per Chip







## Summary

- Digital Gate Characterization (§ 1.3)
- Static Behavior (Robustness) (§ 5.3)
  - VTC
  - Switching Threshold
  - Noise Margins
- Dynamic Behavior (Performance) (§ 5.4)
  - Capacitances
  - Delay
- Power (§ 5.5)
  - Dynamic Power, Static Power, Metrics
- Scaling (§ 5.6)