EE-429
Fundamentals of VLSI Design

CMOS Power Consumption

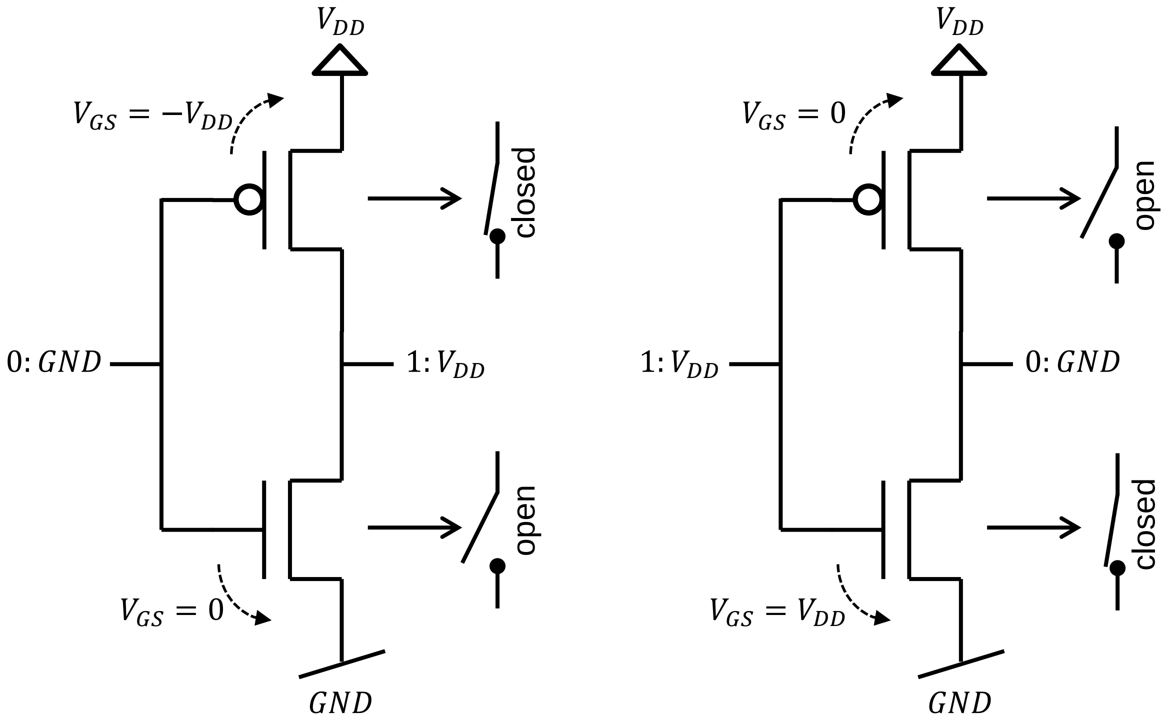
Andreas Burg, Alexandre Levisse

CMOS: Two Types of Power Consumption

- **CMOS circuit power/energy consumption comes in two forms:**
  - **Dynamic energy/power consumption**: depends on activity
    - Charging and discharging of capacitors
    - Cross (short-circuit) currents while pMOS and nMOS are on during switching
  - **Static power consumption**: independent from activity
    - Constant biasing currents (intentional)
    - Various types of leakage currents (parasitic)
    - Contention currents when driving opposite directions (mostly accidents)

Basic Inverter (Unloaded)

- **PMOS performs pull-up to V_{DD}, NMOS performs pull-down to GND**
  - **Complementary gate**
  - Static (steady state): **output connected to either V_{DD} or GND**

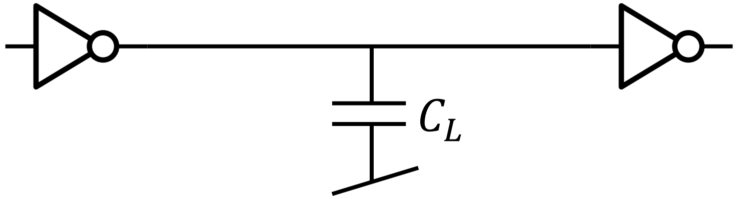


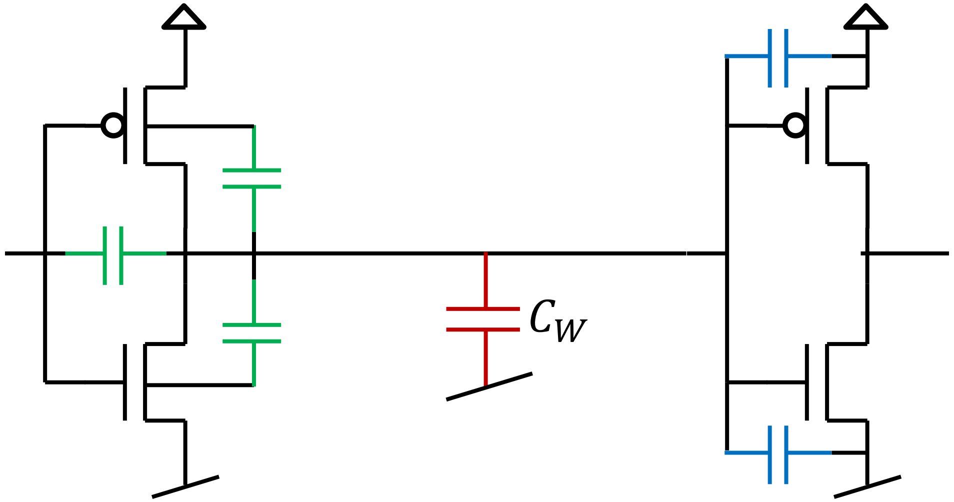
- **Ideally, no current path from V_{DD} to GND**: ideally, no static power consumption

CMOS Gates With Capacitive Load

- **Every CMOS gate sees a capacitive load from various sources**
  - Intrinsic MOS transistor capacitors (driver)
  - Extrinsig (fanout) MOS transistor capacitances
  - Interconnect capacitance

- **Various load capacitances are merged into a single load capacitor C_L**





  - Wider transistors increase the gain factor (drive) but also increase the load (capacitance)

Energy of an Inverter with Capacitive Load

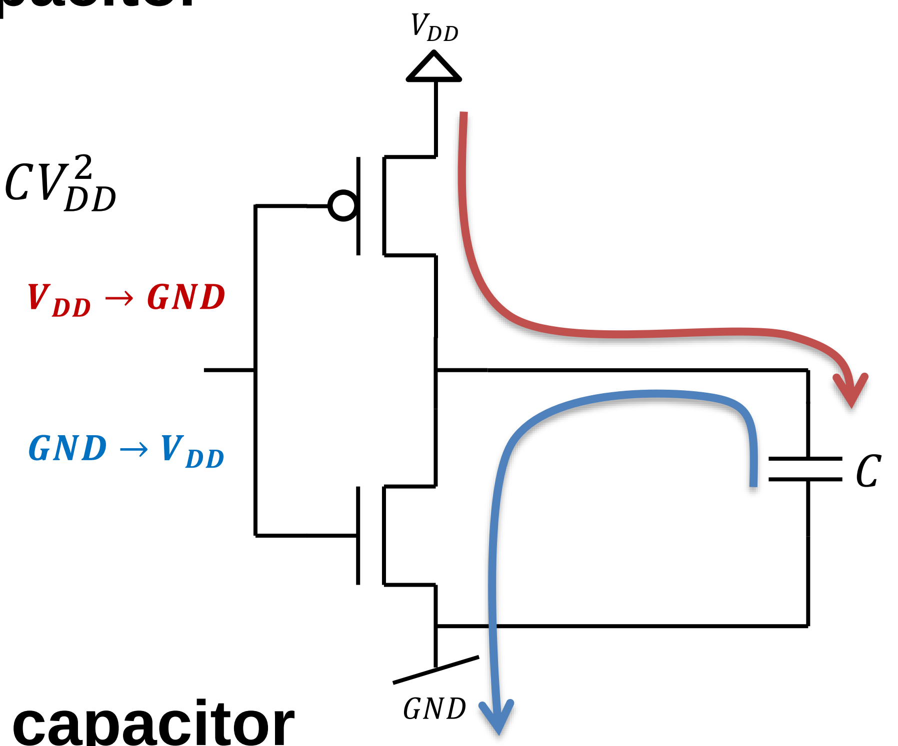
- **Switching** the **output from 0 to 1** **charges the capacitor**
  - Energy delivered by the power supply

    $$E_{VDD} = \int_0^\infty I(t)V_{DD}dt = \int_0^\infty C\frac{dV}{dt}V_{DD}dt = CV_{DD}\int_0^{V_{DD}} dV = CV_{DD}^2$$

  - Once the output transition is complete, the energy stored on the capacitor is given by

    $$E_C = \frac{1}{2}CV_{DD}^2$$

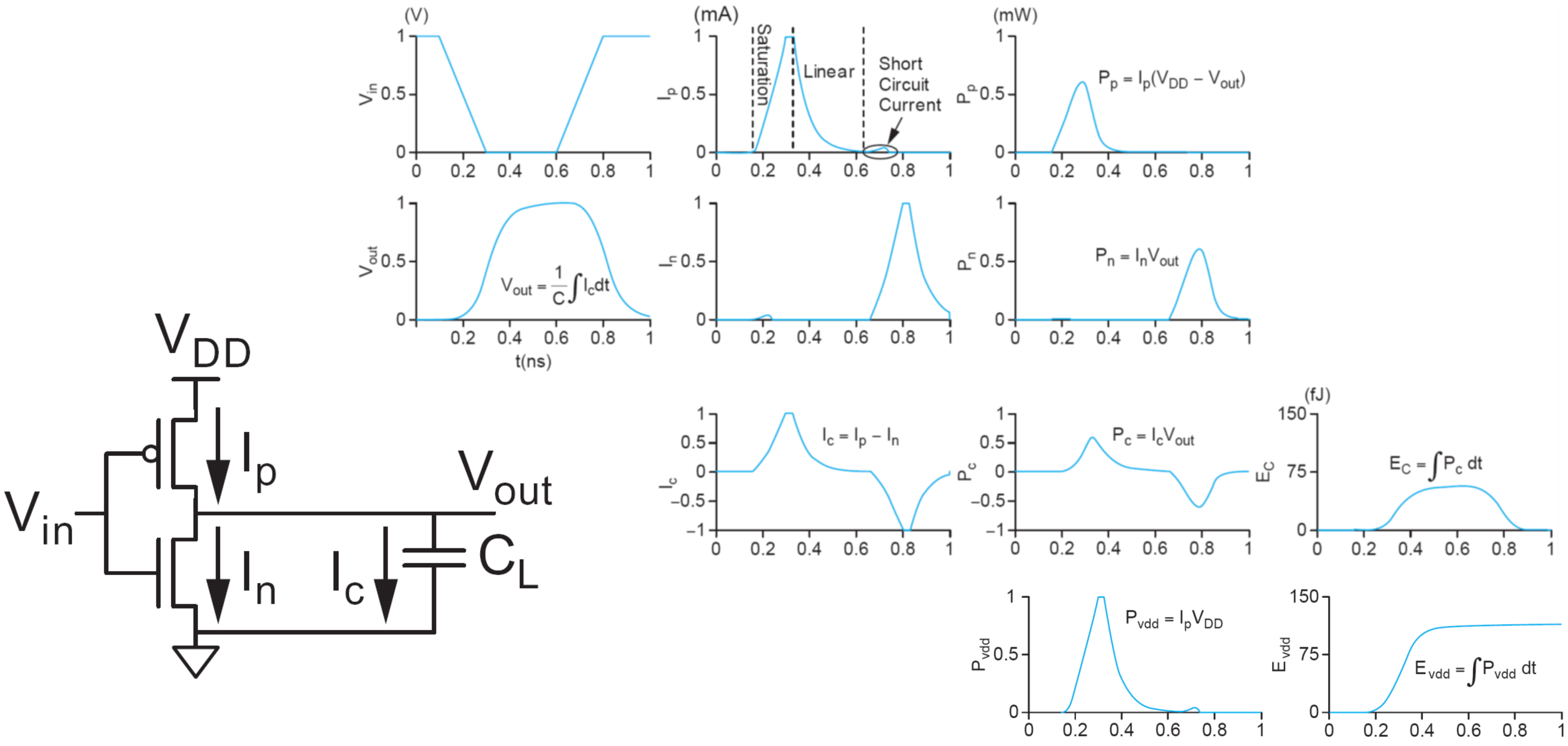
  - Energy difference is dissipated to heat in pMOS



- **Switching** the **output from 1 to 0** **discharges the capacitor**
  - Energy on capacitor is dissipated to heat in nMOS

- **NOTE: energy consumption is independent of the waveform**

Active Energy Consumption Waveforms



Active Power/Energy Consumption

- **Energy consumed during one pair of transitions $E_{\downarrow\uparrow}$:**

$$E_{\downarrow\uparrow} = (CV_{dd})V_{dd} = CV_{dd}^2$$

- **Energy/transition**

$$E_t = CV_{dd}^2/2$$

- **Average power consumption:** energy per time T
  - Depends on the switching frequency f_{sw} of the output $P_{sw} = \frac{E_t}{T} = \frac{2f_{sw}TCV_{dd}^2/2}{T} = f_{sw}CV_{dd}^2$
  - **Activity factor α : average number of transitions per cycle**
    - Relates activity of a node to the clock frequency f_{clk}
    - Energy/transition * average-transition/cycle (α) * clock frequency (f_{clk})

$$P = \frac{\alpha}{2}CV_{dd}^2 f_{clk}$$

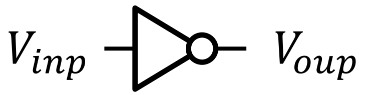
Short Circuit Currents

- **CMOS gates have a** large, but **finite gain** in the transition region
  - **Cross-over currents lead to power consumption** during transients

range	applies when	n-channel ▼	p-channel ▲
A	$0 \leq U_{inp} \leq U_{th\,n}$	subthreshold	linear
B	$U_{th\,n} < U_{inp} < U_{inv}$	saturation	linear
C	$U_{inp} \approx U_{inv}$	saturation	saturation
D	$U_{inv} < U_{inp} < U_{dd} + U_{th\,p}$	linear	saturation
E	$U_{dd} + U_{th\,p} \leq U_{inp} \leq U_{dd}$	linear	subthreshold

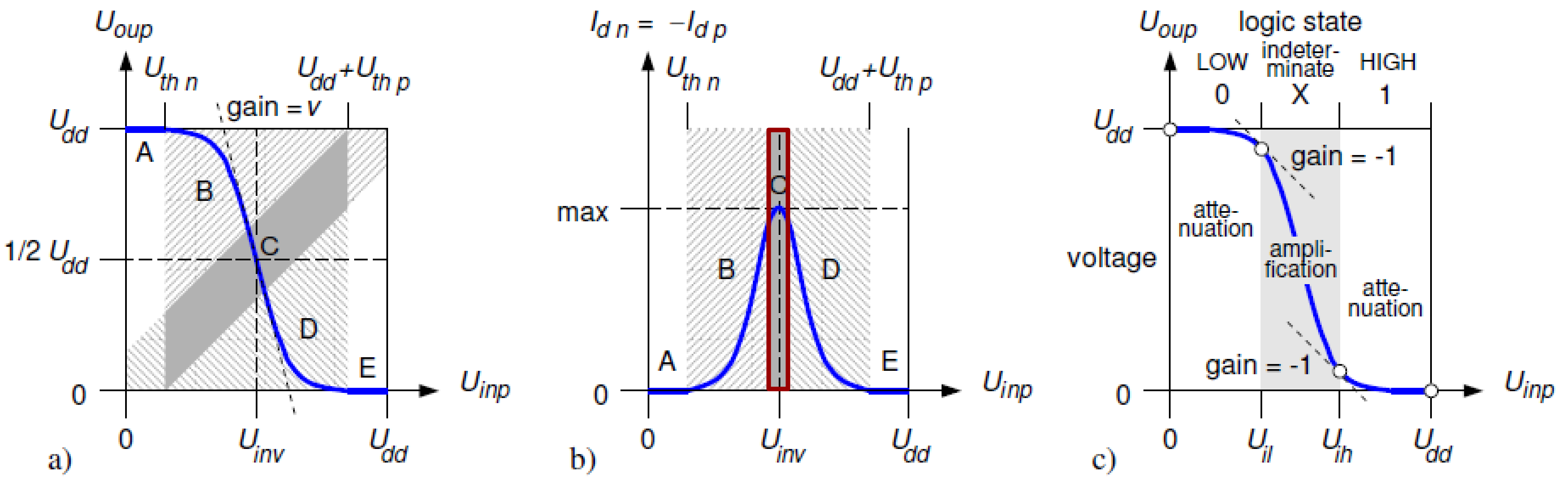
PROBLEMATIC

Dominant during transition region:
rapid opening of the driver for the new level



- **Note:** short-circuit power is irrelevant if

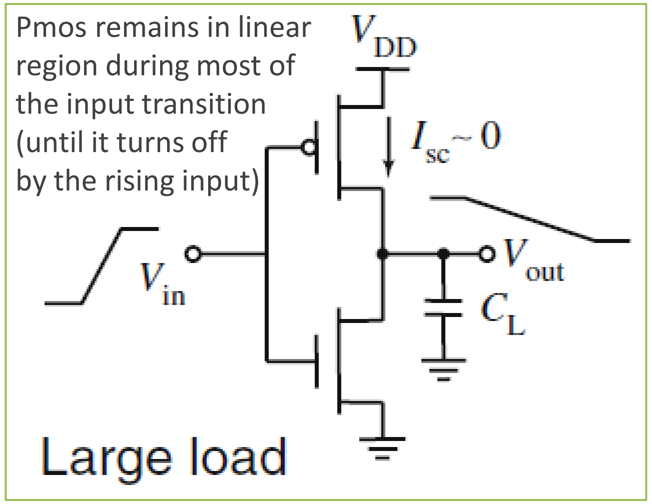
$$V_T < \frac{V_{DD}}{2}$$

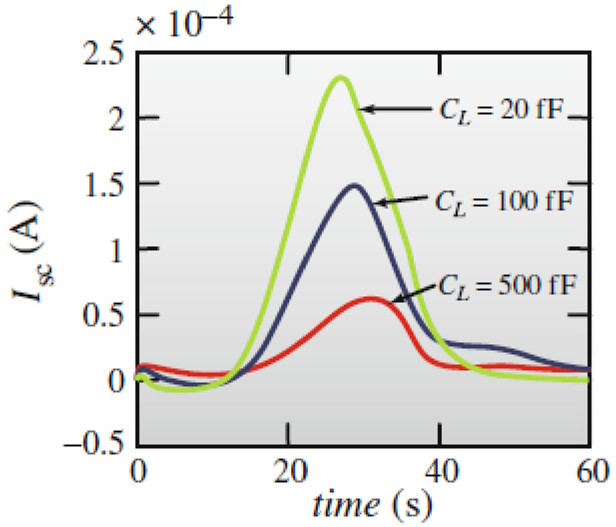


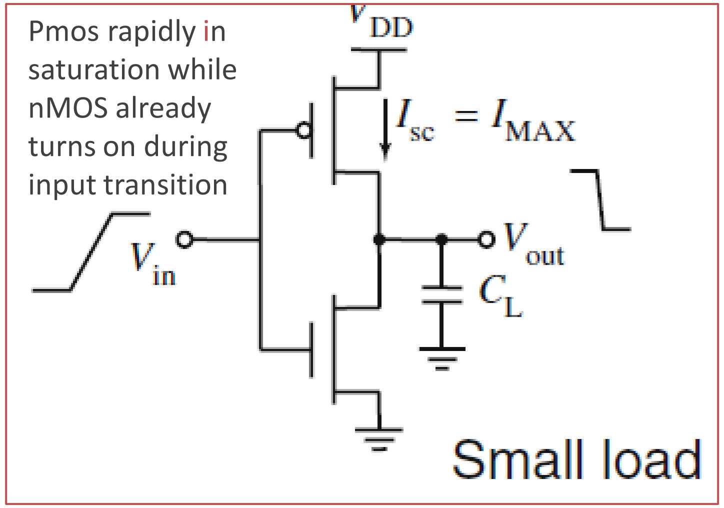
(a) Transfer characteristic (b) Crossover current (c) Logic states

Minimizing Short-Circuit Currents

- **Obvious approach:** reduce input transition time
  - However, short input transitions require stronger driver => increase transition time (and load) on the driver of the driver
- **Control short circuit current by controlling the output slope:**
  - Fast input slow output: driving device mostly in linear regime => good for low power
  - Slow input fast output: driving device remains long in saturation => bad for low power

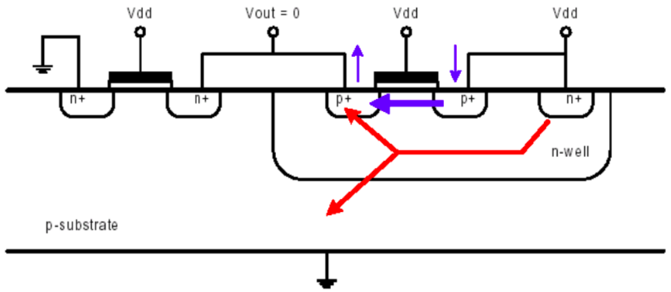




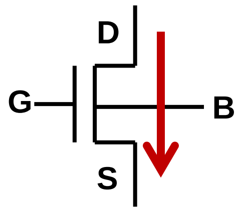


- **Best compromise: balance input slope and output slope**

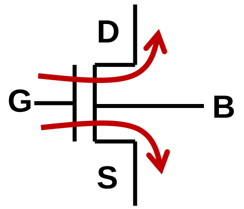
Leakage Power



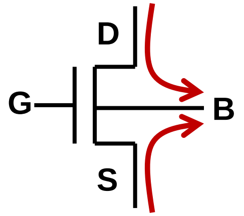
- **Transistors leak currents even when in off-state**
- **Sources for leakage**
  - Sub-threshold leakage
    - Dominant component in most circuits



  - Gate tunneling
    - Generally low, even in modern technologies due to high-k gate dielectrics
    - Decreases very rapidly with decreasing V_{dd}



  - Junction current
    - Generally low
    - Decreases very rapidly with decreasing V_{dd}



Leakage Power

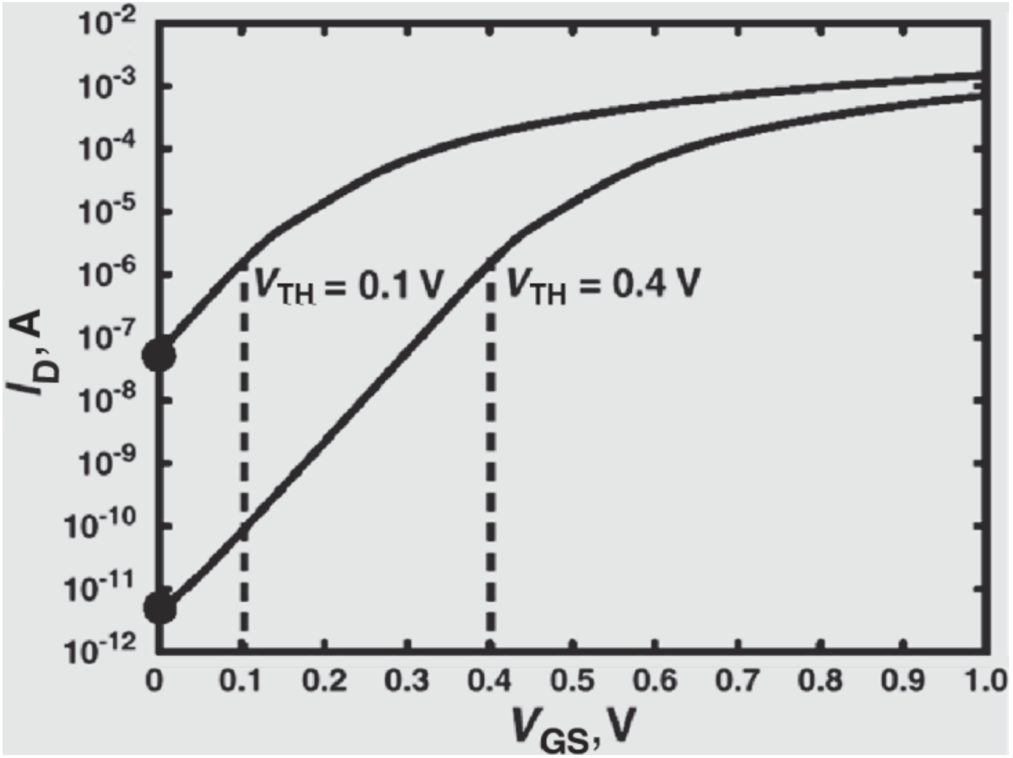
- **Long channel deices (>130nm)**

$$I_{DS} = I_0 e^{\frac{V_{GS}-V_T}{v_t n}}$$

v_t : thermal voltage
n : constant

- I_{DS} mostly independent from Drain-Source voltage
- Leakage current depends strongly on $V_{GS} - V_T$
  - **Lower threshold voltage increases leakage**
  - **Higher threshold voltage decreases leakage**

  - **Subthreshold slope**: slope of the logarithmic leakage current for $V_{GS} - V_T < 0$



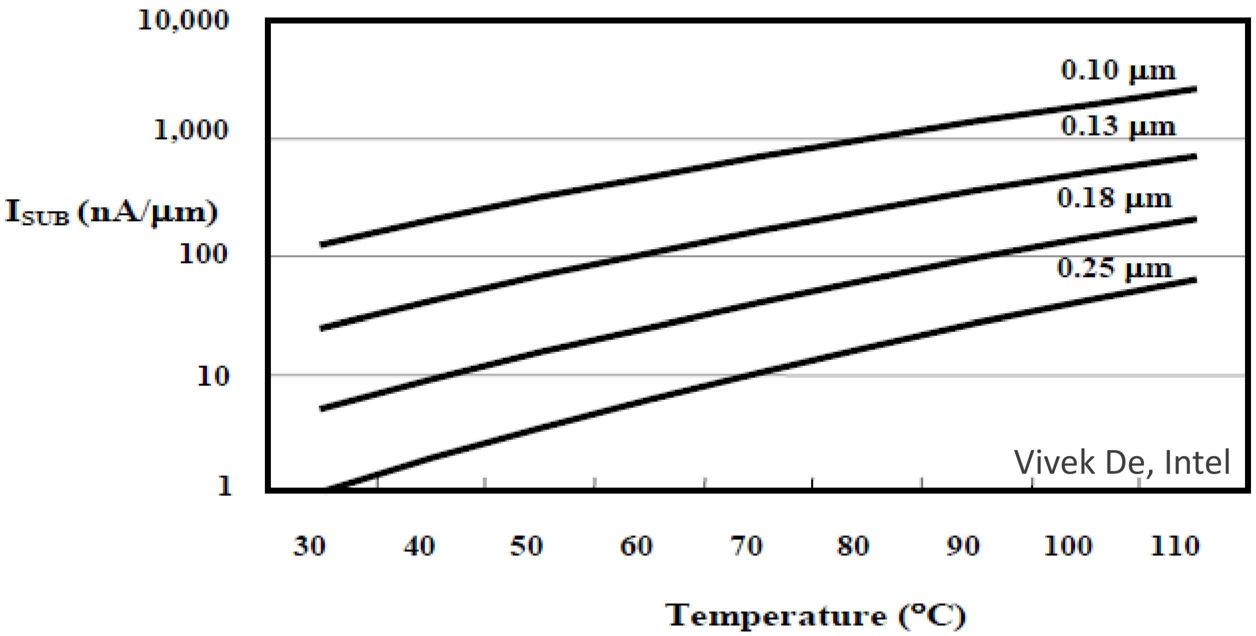
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

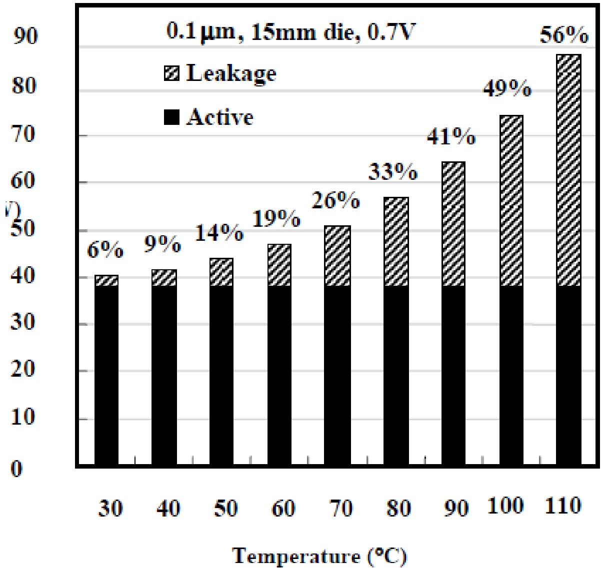
Leakage Power over Temperature

- **Sub-threshold current depends exponentially on thermal voltage $v_t = kT/q$**

$$I_{DS} = I_0 e^{\frac{V_{GS}-V_{th}}{kTn/q}}$$

- **Exponential sub-threshold leakage (I_{DS}) increase with temperature**

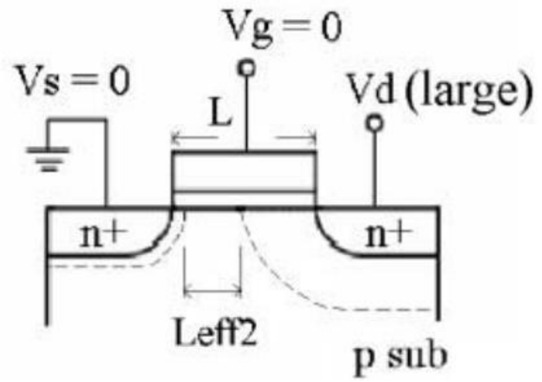




Example: 0.7V, 100nm process, 15mm2 die

Leakage Power (DIBL)

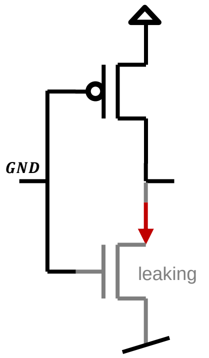
- **Impact of technology scaling on sub-threshold leakage (<130nm)**
  - Drain-Induced Barrier Lowering (DIBL): V_{DS} modulates threshold voltage
  - I_{DS} becomes a function of V_{DS}



$$I_{DS} = I_0 e^{\frac{V_{GS} - V_{th} + \lambda_{DS} V_{DS}}{v_t n}}$$

λ_{DS} : DIBL coefficient
v_t : thermal voltage
n : constant

- **Impact on inverter leakage:** no longer supply independent



$$I_{leak} = I_0 e^{\frac{-V_{th} + \lambda_{DS} V_{DD}}{v_t n}}$$

Reducing voltage reduces leakage

Leakage in Transistor Stacks (Short Channel)

- **Stacking occurs**
  - In many logic gates (> 1 input)
  - When introduced intentionally for leakage reduction

$$I_{leak,M1} = I_0 e^{\frac{-V_M - V_{th} + \lambda_{DS}(V_{dd} - V_M)}{v_t n}}$$

$$I_{leak,M2} = I_0 e^{\frac{-V_{th} + \lambda_{DS} V_M}{v_t n}}$$

Leakage Reduction	
2 NMOS	9
3 NMOS	17
4 NMOS	24
2 PMOS	8
3 PMOS	12
4 PMOS	16

Small speed penalty: ~25%