SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Term 7- Sept 2025

Nanoelectronics and Technology (01.119/99.503)-Week 3 Class 2

Shubhakar K

Email: shubhakar@sutd.edu.sg

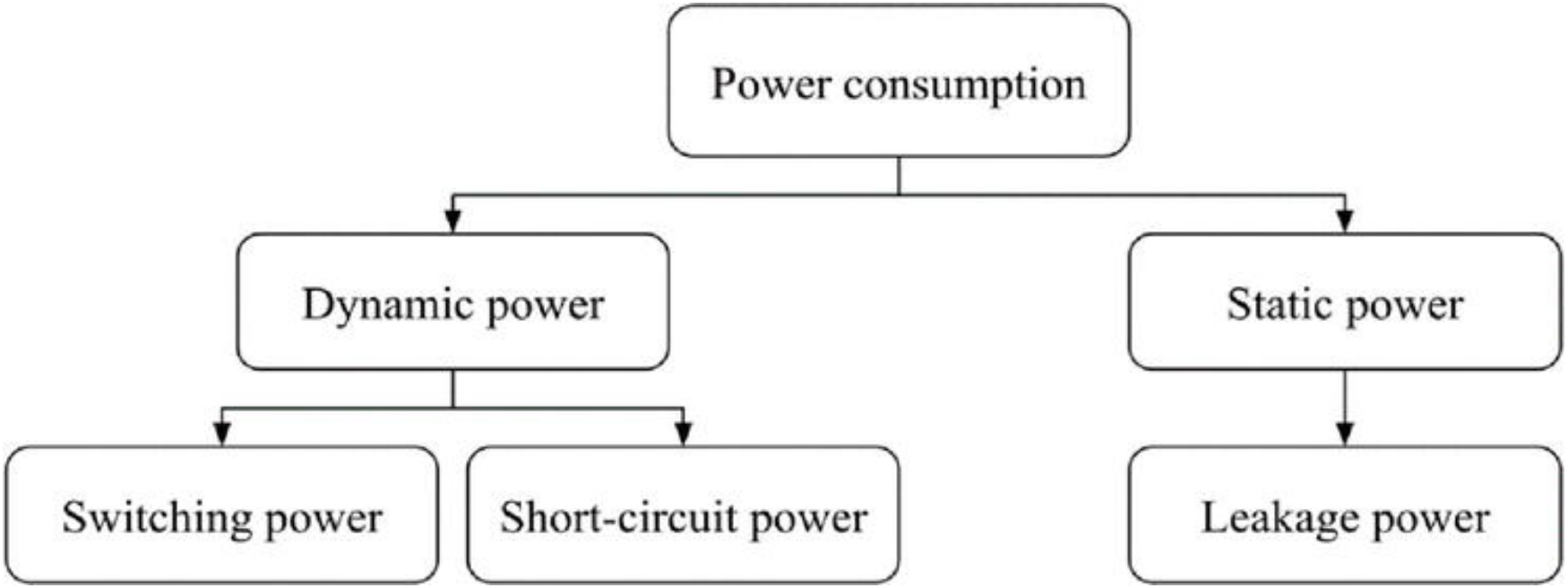
Office location: 1.602-26

30-Sept- 2025

Outline

- CMOS Inverter
- CMOS scaling
- Impact of Scaling on different CMOS parameters
- Scaling Challenges

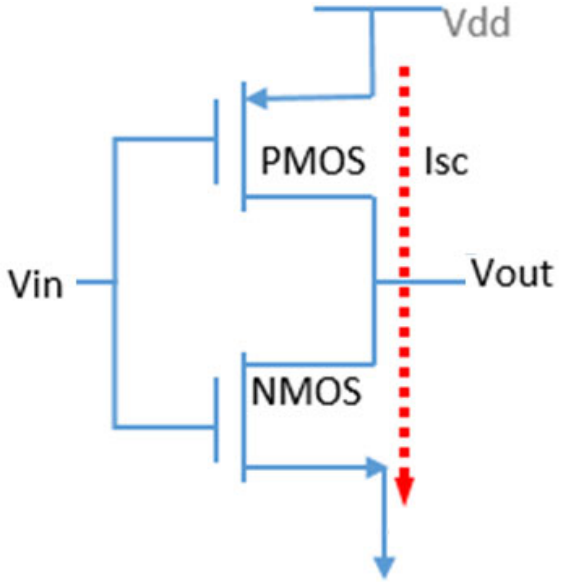
Power consumption CMOS Inverter



Power consumption CMOS Inverter

Static Power Consumption

- **Ideally very small** because in steady states (logic 0 or logic 1 at input), either NMOS or PMOS is OFF.
- In ideal CMOS, there's no direct path between VDD and GND when input is stable.
- **Main contributors:**
  - **Leakage currents** (subthreshold leakage, gate-oxide tunneling, junction leakage).
  - In modern nanometer technologies, leakage has become a dominant concern.



$$P_{static} = I_{static} V_{DD}$$

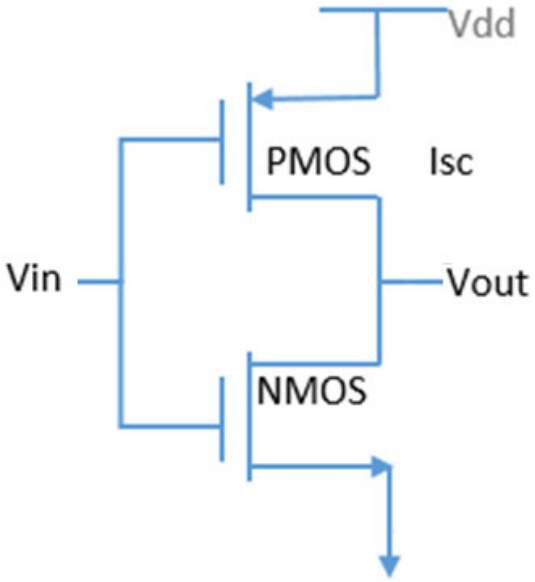
Power consumption CMOS Inverter

Short circuit power consumption

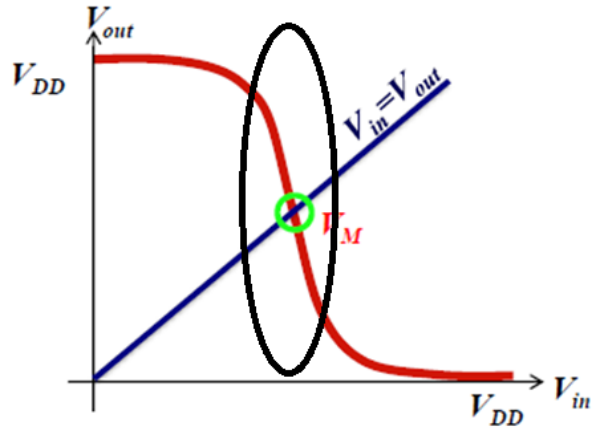
•During switching, for a brief time **both NMOS and PMOS conduct simultaneously**.

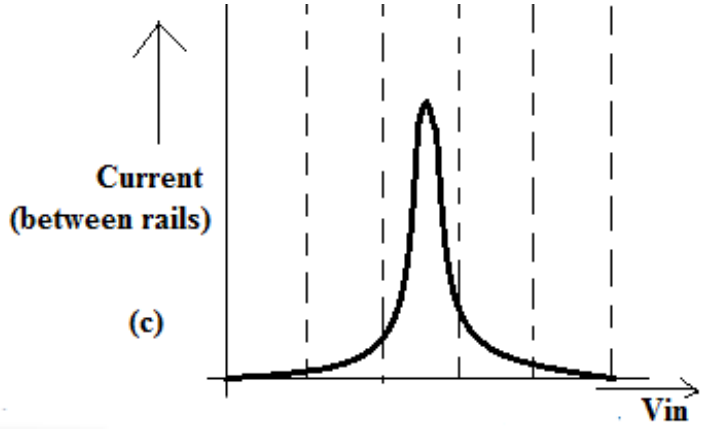
•Creates a **direct current path from VDD to GND** until output stabilizes

•Increases with slower input transitions and higher supply voltage.



$$P_{sc} = t_{sc} V_{DD} I_{peak} f$$



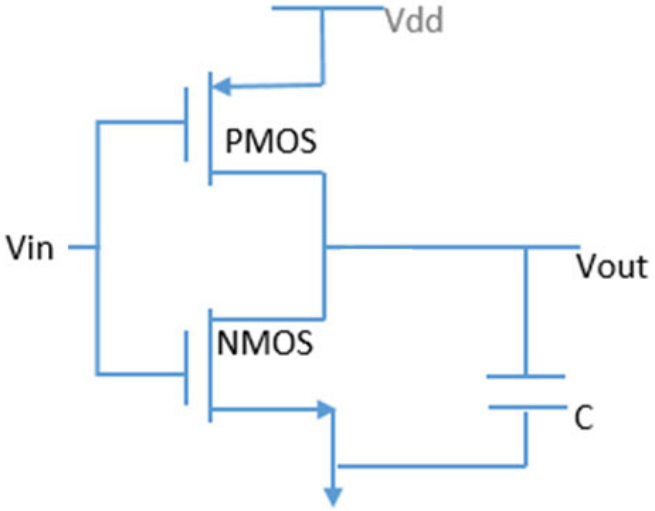


Power consumption CMOS Inverter

Dynamic power consumption

•Occurs when the input toggles and output transitions.

•Due to **charging and discharging of load capacitance C_L** at the output node.



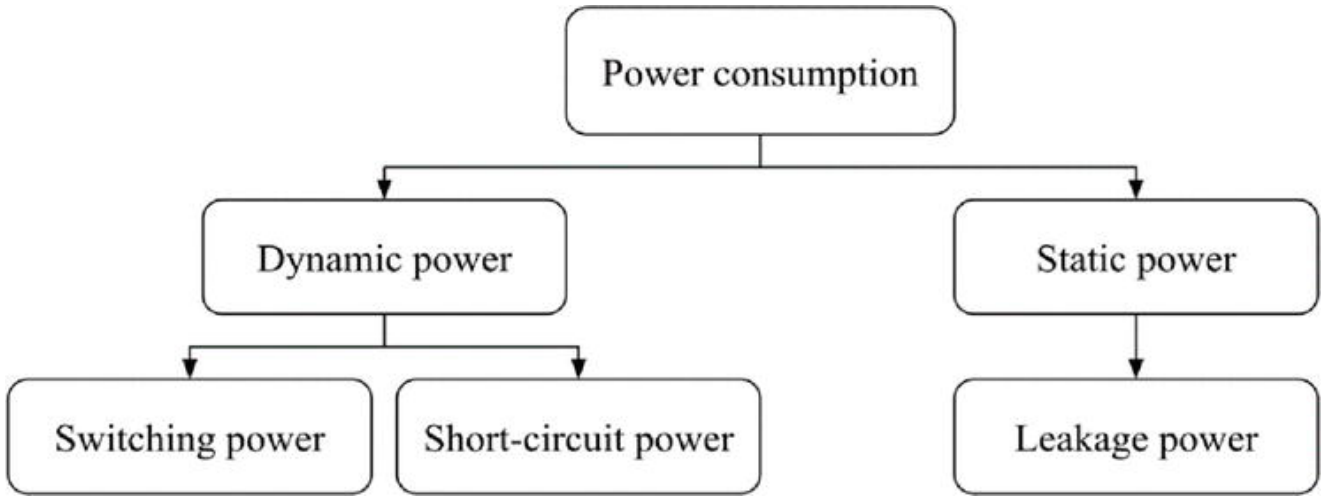
- **Cause:** Charging and discharging of the load capacitance (C_L) at the output node.
- When the inverter output switches from `0 → 1`, the PMOS charges the load capacitor to V_{DD}.
- When it switches from `1 → 0`, the NMOS discharges the capacitor to GND.
- Energy lost = $\frac{1}{2}C_L V_{DD}^2$ per transition, and since both charge and discharge happen, the average power is:

$$P_{dynamic} = \alpha\, C_L\, V_{DD}^2\, f$$

$$P_{dyn} = C_{eff} V_{DD}^{\ 2} f$$

- **Key point:** Purely capacitive effect, no direct current path from VDD to GND (except during transitions).

Power consumption CMOS Inverter



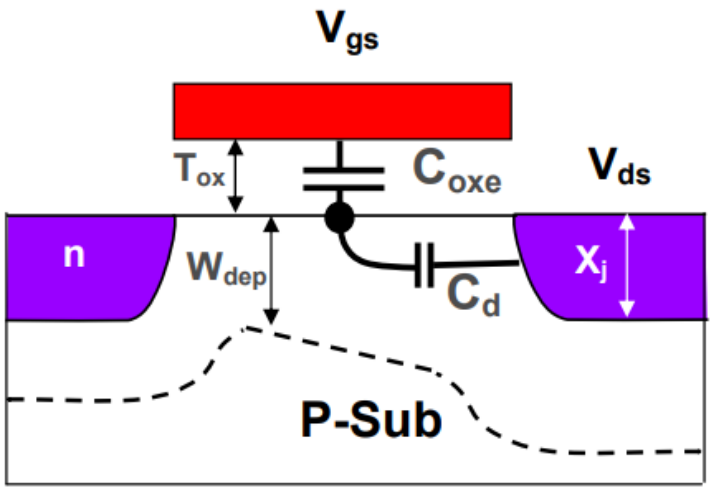
$$P_{dyn} = C_{eff} V_{DD}^{\ 2} f$$

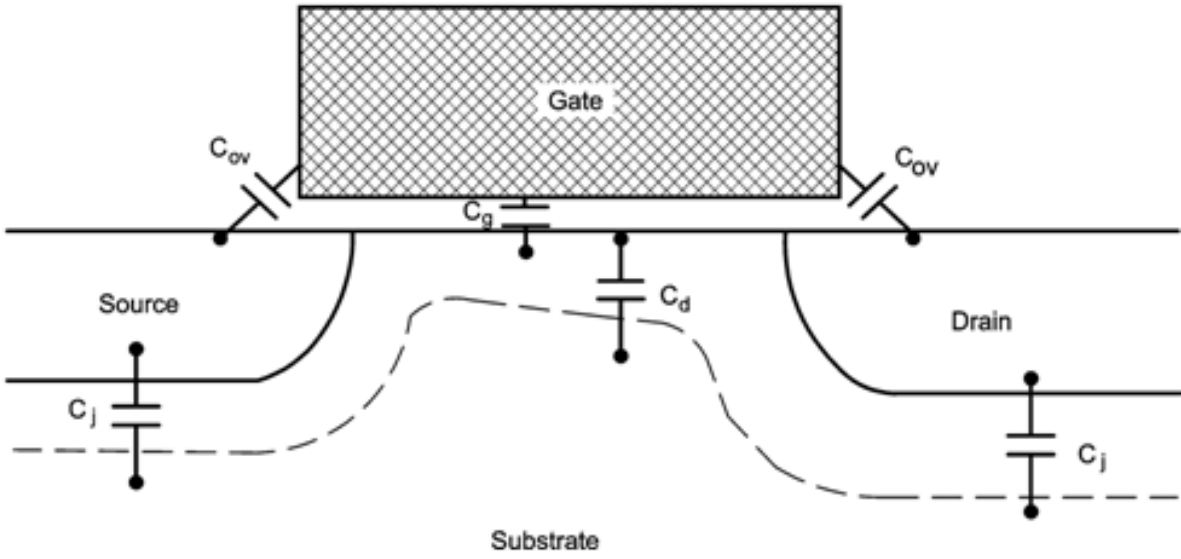
$$P_{sc} = t_{sc} V_{DD} I_{peak} f$$

$$P_{static} = I_{static} V_{DD}$$

- **Dynamic power** = "battery charging & draining the capacitor."
- **Short-circuit power** = "battery shorted briefly while capacitor is switching."

Impact of scaling on parasitic C and R



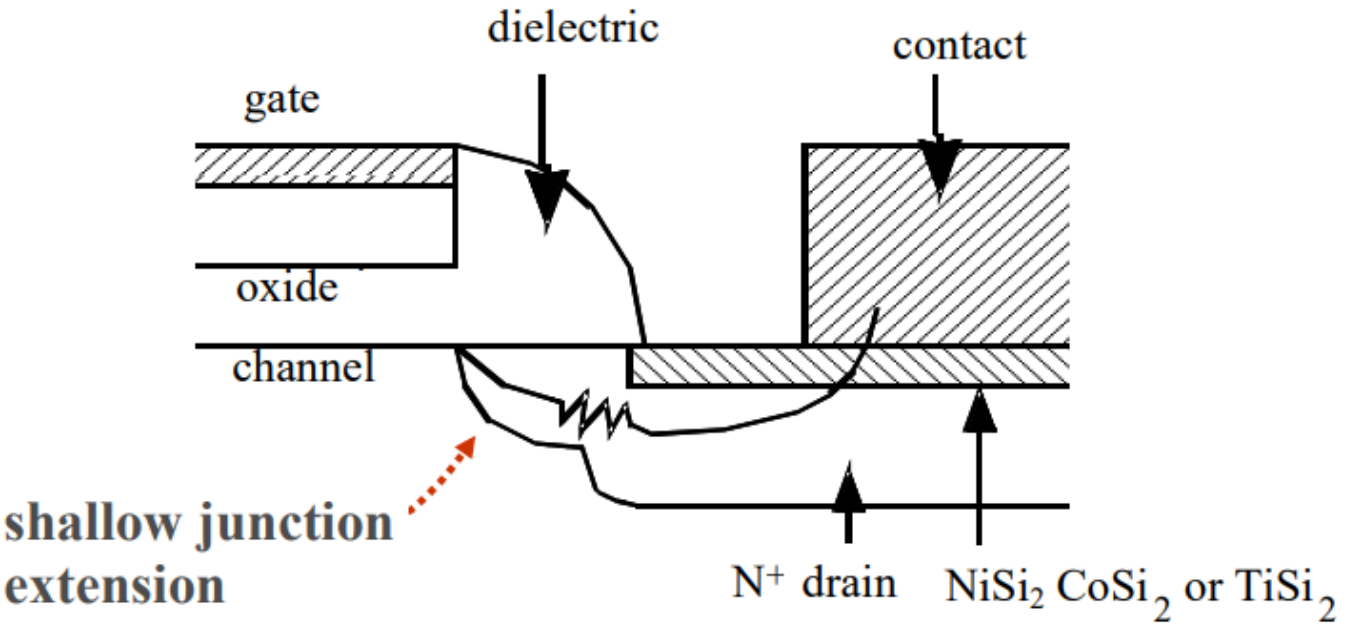


MOSFET parasitic capacitances

https://inst.eecs.berkeley.edu/~ee130/sp06/chp7full.pdf

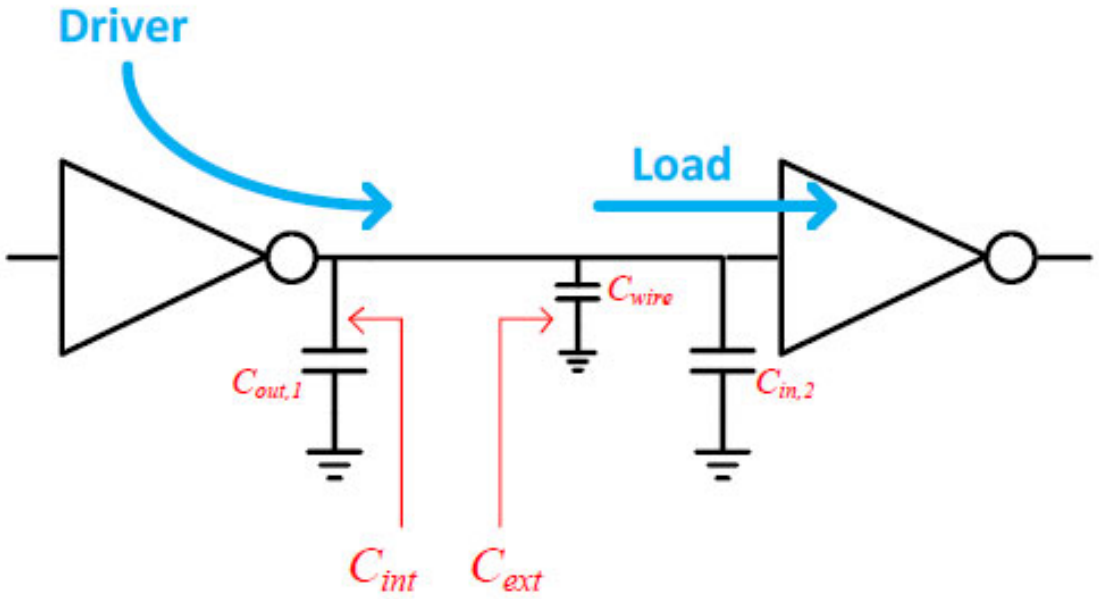
https://www.electronics-tutorial.net/Analog-CMOS-Design/MOSFET-Parasitics/Parasitic-Capacitances-MOSFETS/

Impact of scaling on parasitic C and R



https://inst.eecs.berkeley.edu/~ee130/sp06/chp7full.pdf

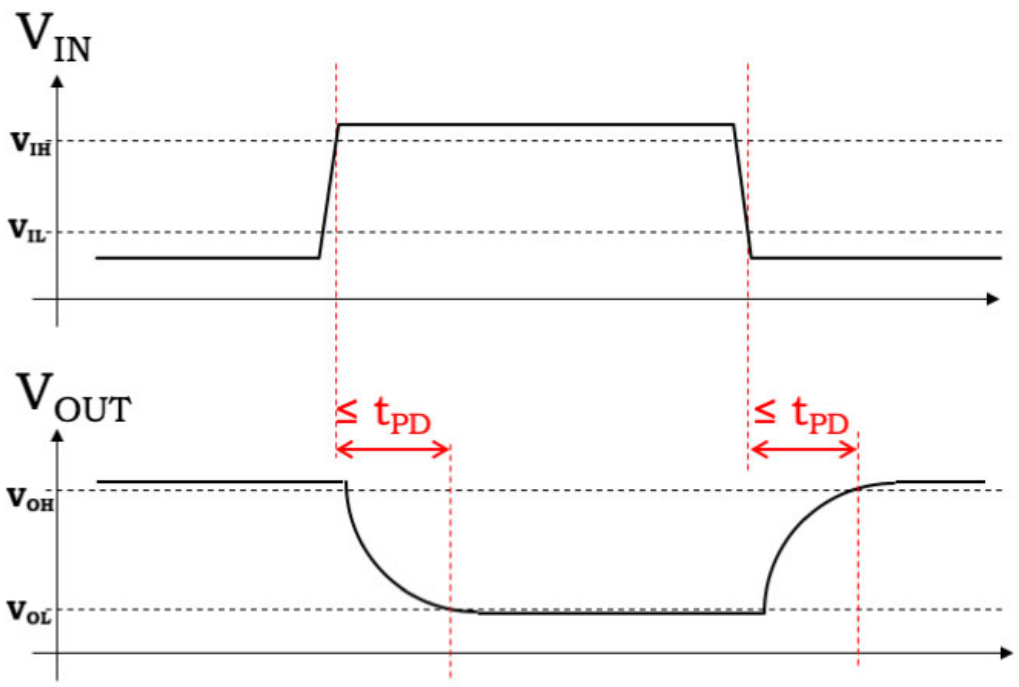
Parasitic Capacitances



$$C_{load} = C_{out} + C_{wire} + N \cdot C_{in}$$

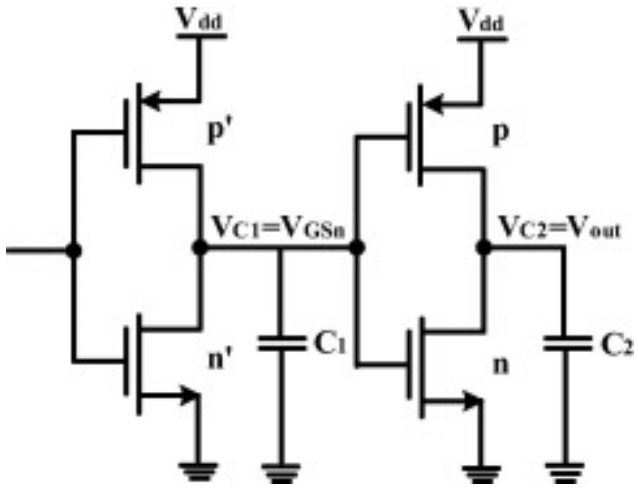
Propagation delay in CMOS Inverter

Propagation delay (t_{PD}): An UPPER BOUND on the delay from valid inputs to valid outputs.

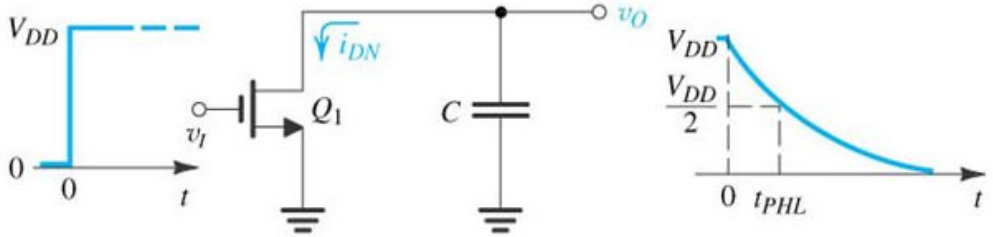


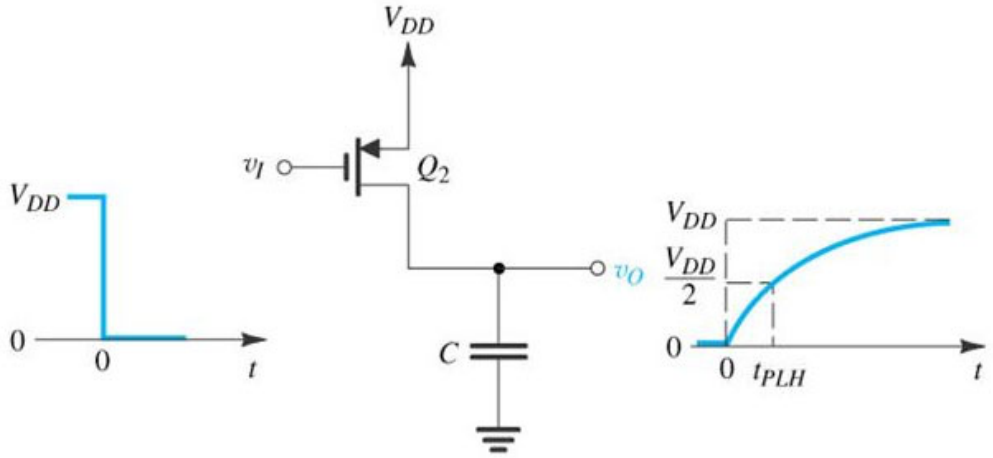
GOAL:
minimize propagation delay!

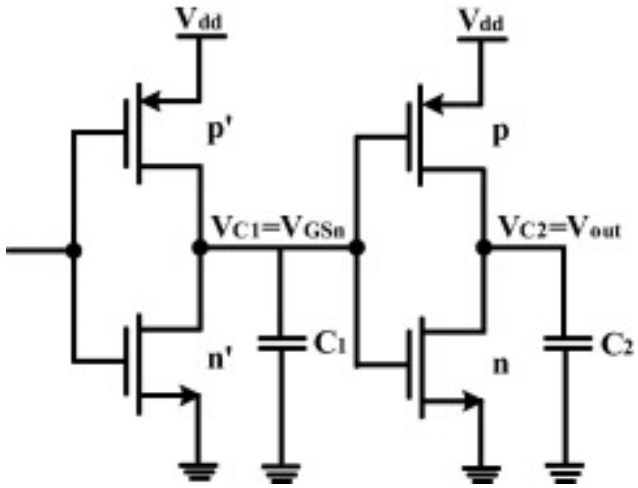
ISSUE:
keep capacitances low and transistors fast

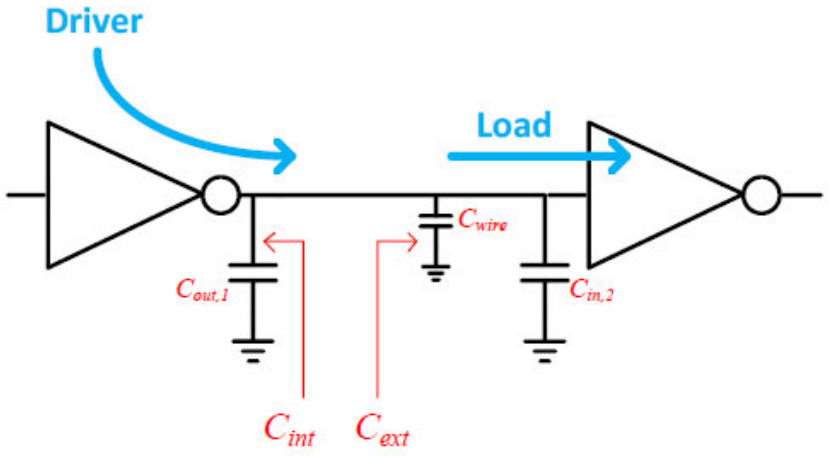


Propagation delay in CMOS Inverter









CMOS scaling Issues:

- Short channel effects
- Threshold voltage variation
- DIBL (Drain Induced Barrier Lowering)
- Gate leakage current
- GIDL (Gate Induced Drain Leakage)
- Shallow S/D – Parasitic resistance
- Mobility issues/Velocity saturation/Hot carrier effect

HK-MG technology at 45nm

High –k Dielectric

a. Increases the gate field effect.

b. Allows use of thicker dielectric layer to reduce gate leakage .

c. Approximately 10 atomic layers thick deposited by ALD .

Advantages of HK Dielectrics + Metal Gates

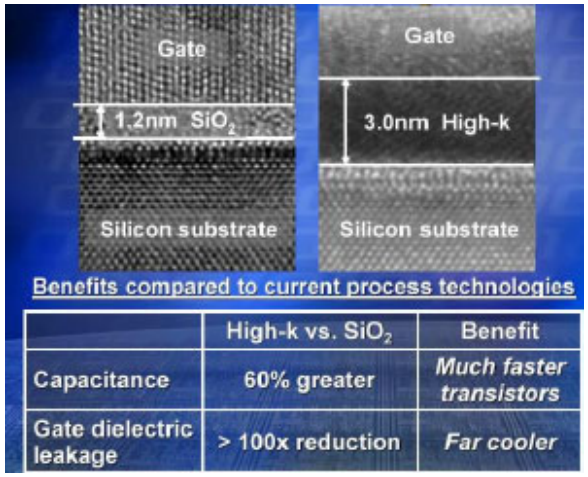
- Drive current increase >20%
- Source-Drain leakage reduced >5x
- Gate oxide leakage reduced >10x
- Lower power dissipation that creates less heat

Metal Gates (titanium nitride “TiN” for PMOS; titanium nitride & with an aluminum alloy ”TiAlN” for NMOS).

Increases the gate effect

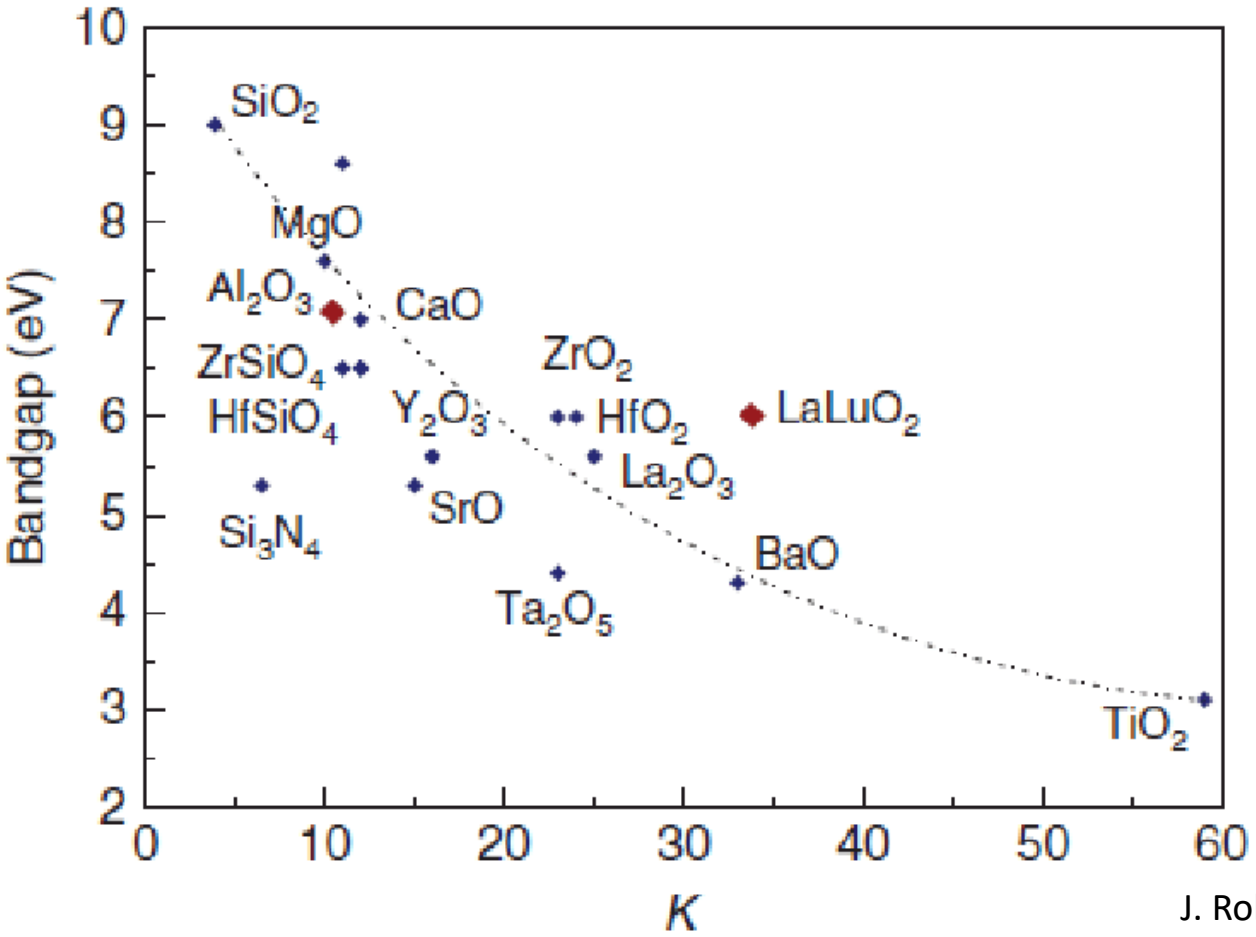
CMOS scaling: HK-MG technology

- Challenges:
  - High density of defects
  - Interaction with the polysilicon gate
  - Interaction with the substrate
  - Polycrystallization structure
  - Increased trap-assisted tunneling
  - Increased leakage current
  - Reliability and failure of high-k dielectric



***Source*: Intel**

CMOS scaling: HK-MG technolgy

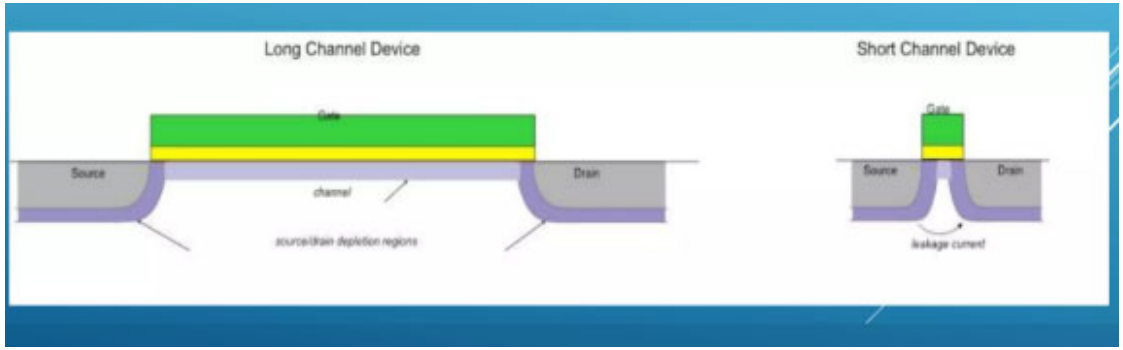


$$\mathrm{EOT} = \left(\frac{\kappa_{\mathrm{ox}}}{\kappa_{\text{hi-}k}} \right) t_{\text{hi-}k}$$

J. Robertson, J. Vac. Sci. Technol. B 18, 1785 (2000)

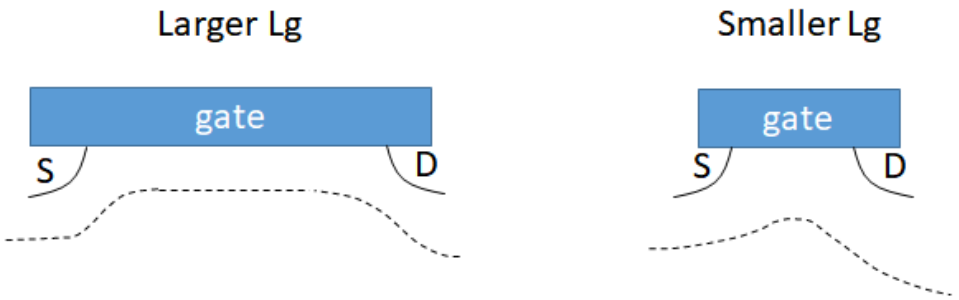
CMOS scaling:

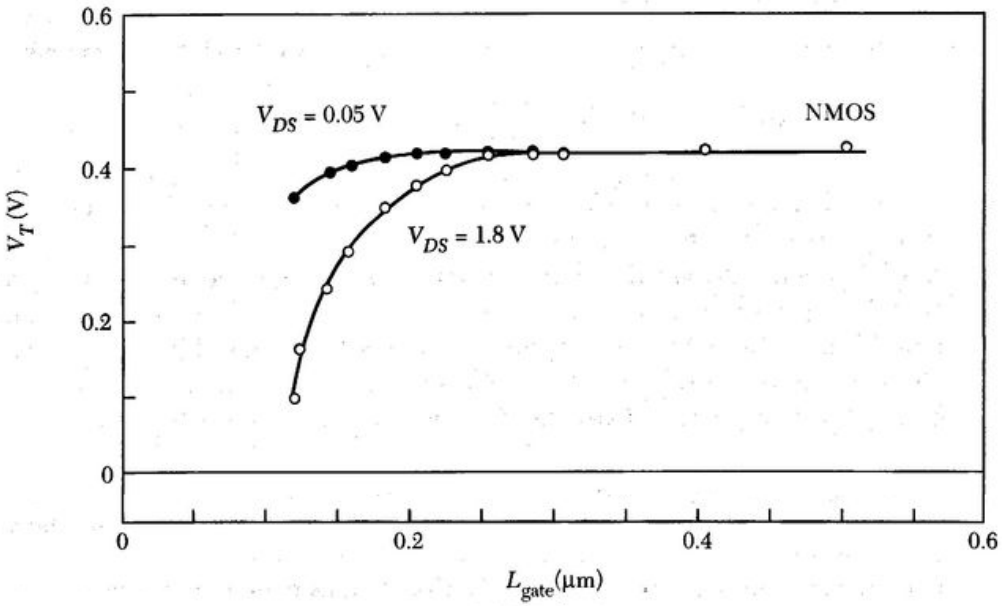
- Gate length scaling



https://www.slideshare.net/slideshow/short-channel-effects/50101267

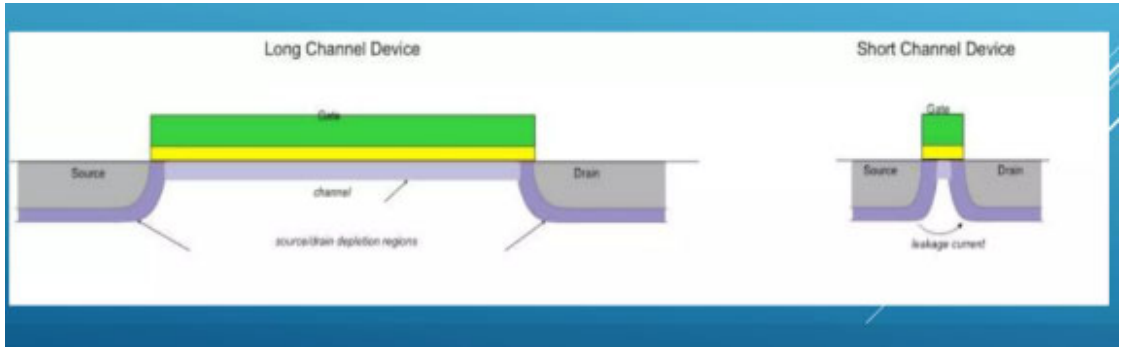
Threshold Voltage: (V_{th}) Variation





CMOS scaling:

- Gate length scaling



https://www.slideshare.net/slideshow/short-channel-effects/50101267

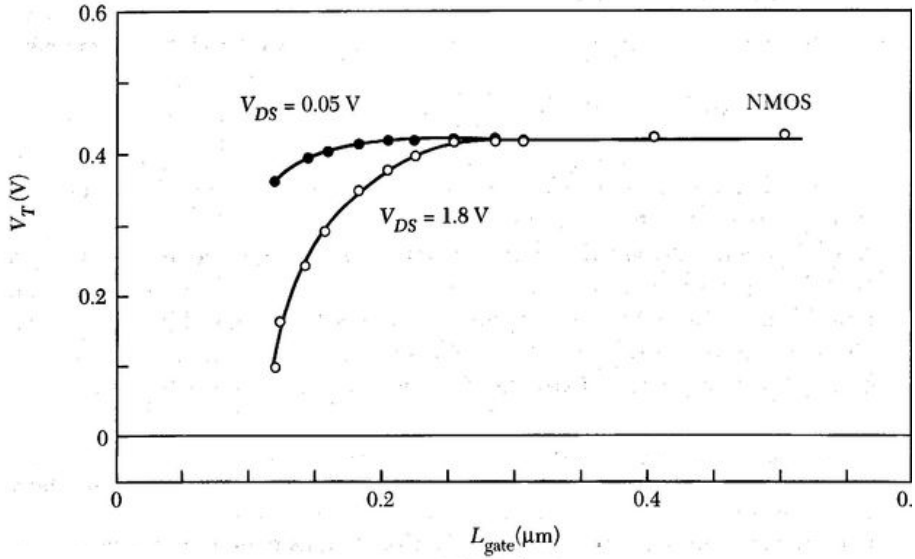
Threshold Voltage: (V_{th}) Variation

Long Channel Devices (Ideal Behavior)

- For large channel lengths (e.g., > 1 µm in older technologies), V_t is almost constant and determined by:

$$V_t = V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_s q N_A (2\phi_F)}}{C_{ox}}$$

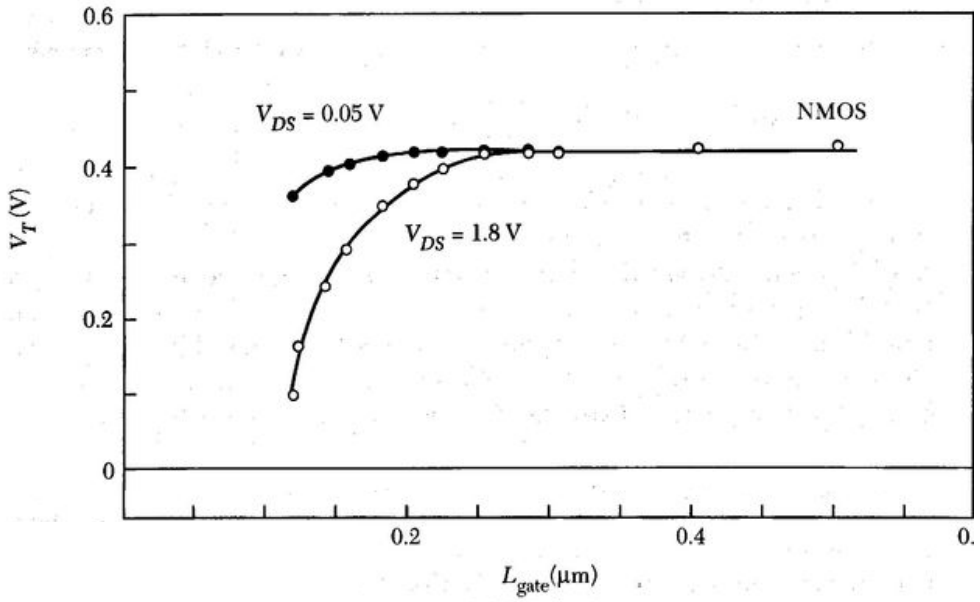
- Where V_{FB} is flat-band voltage, $2\phi_F$ is surface potential, C_{ox} is oxide capacitance.
- In this regime, channel is well controlled by gate → little variation with L.



CMOS scaling:

- Gate length scaling

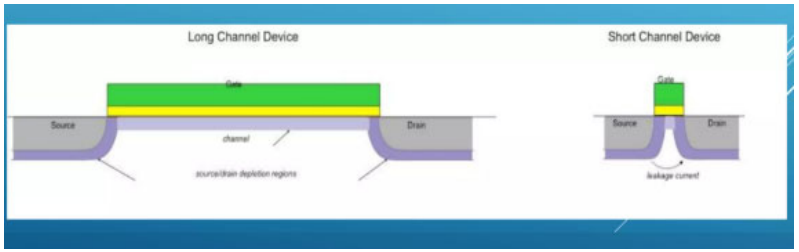
Threshold Voltage: (V_{th}) Variation



Short-Channel Devices

As L decreases (deep submicron), **2D electrostatics** start affecting threshold voltage:

(a) Threshold Voltage Roll-Off



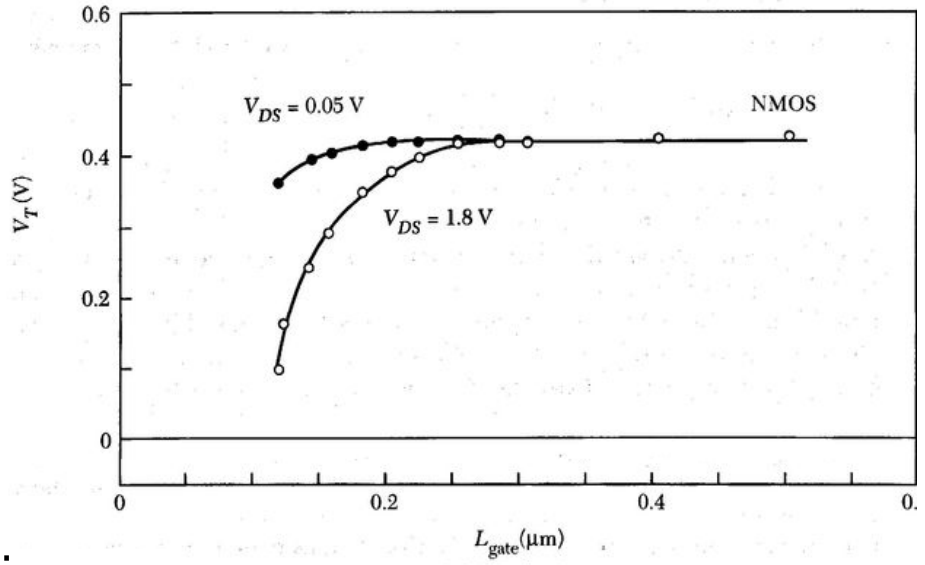
- When channel length is **short**, depletion regions of source and drain penetrate into the channel.
- This lowers the effective gate control over the channel.
- Result: **Vt decreases with decreasing L**.
- Called "**Vt roll-off.**"

https://www.slideshare.net/slideshow/short-channel-effects/50101267

CMOS scaling:

- Gate length scaling

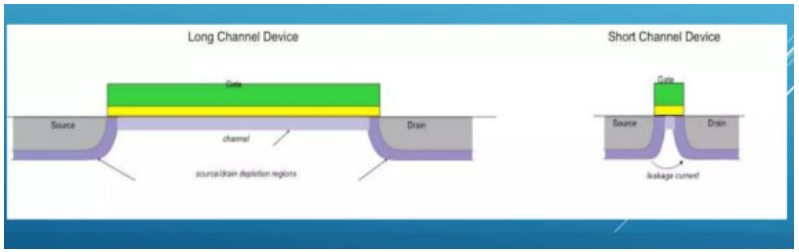
Threshold Voltage: (V_{th})



- At nanometer nodes, controlling **Vt roll-off** is critical to reduce leakage power.

Techniques:

- •Halo/pocket implants
- •High-k dielectrics
- •FinFETs / GAA FETs for better gate control



https://www.slideshare.net/slideshow/short-channel-effects/50101267

CMOS scaling: • Gate length scaling

Channel Length Modulation

- It occurs when transistor is in *Saturation region*.

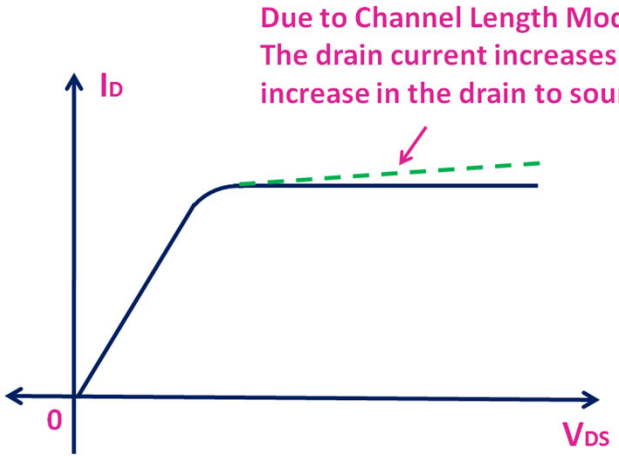
  i.e. **Saturation region**,

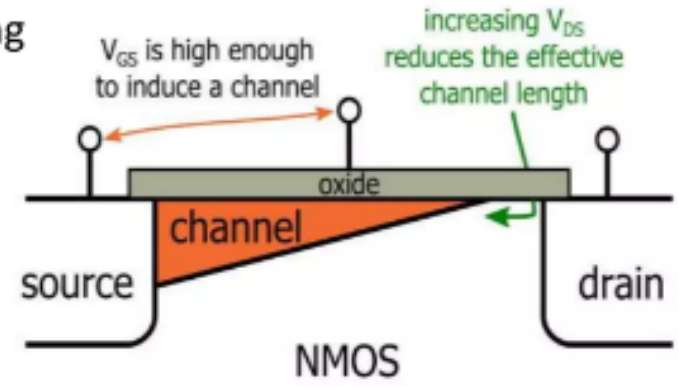
  $V_{GS} > V_{th}$ and $V_{DS} > V_{GS} - V_{th}$

  I_D increases slightly with increasing V_{DS} .

- The pinch-off point moves toward the source as V_{DS} increases.

- The length of the channel becomes shorter with increasing V_{DS}.



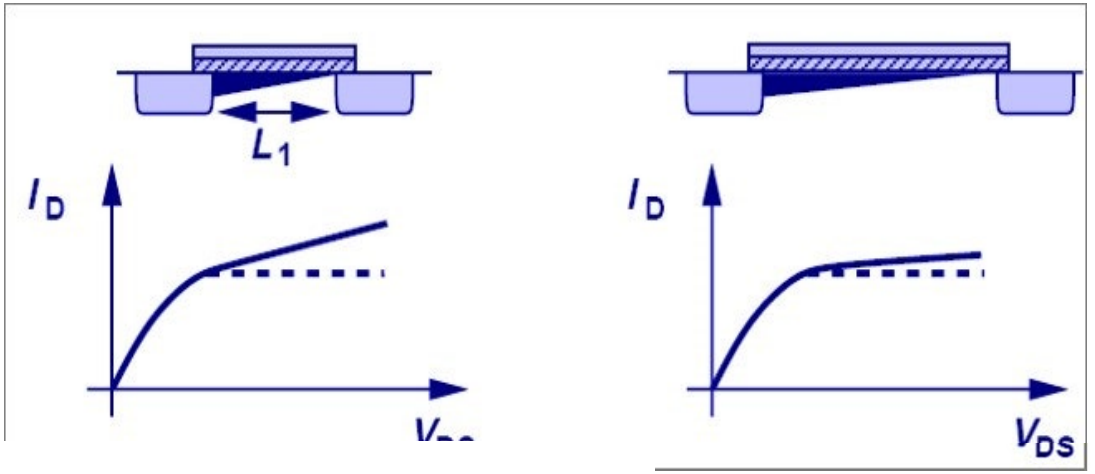


- Since drain current is proportional to $1/L$, reducing L increases current.

CMOS scaling:

- Gate length scaling

Channel length Modulation



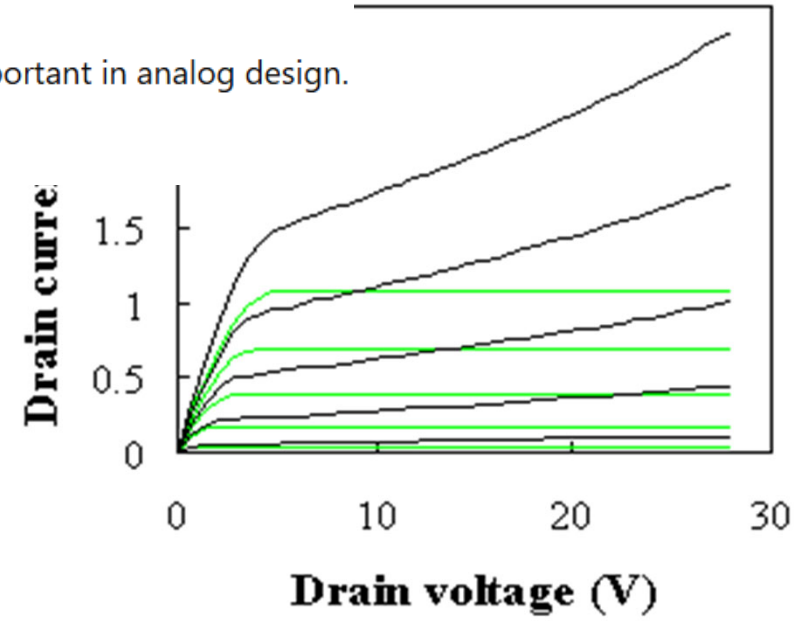
Practical Implications

- Higher drain current than expected in saturation.
- Reduced output resistance ($r_o = 1/(\lambda I_D)$) → important in analog design.
- Impacts gain of MOS amplifiers.

$$I_D = \frac{1}{2}$$

Where:

- λ = channel length modulation coefficient (depends on technology & channel length).
- Larger λ = stronger dependence on V_{DS}.
- For **long-channel MOSFETs**, $\lambda \approx 0$ (negligible).
- For **short-channel MOSFETs**, λ is significant.

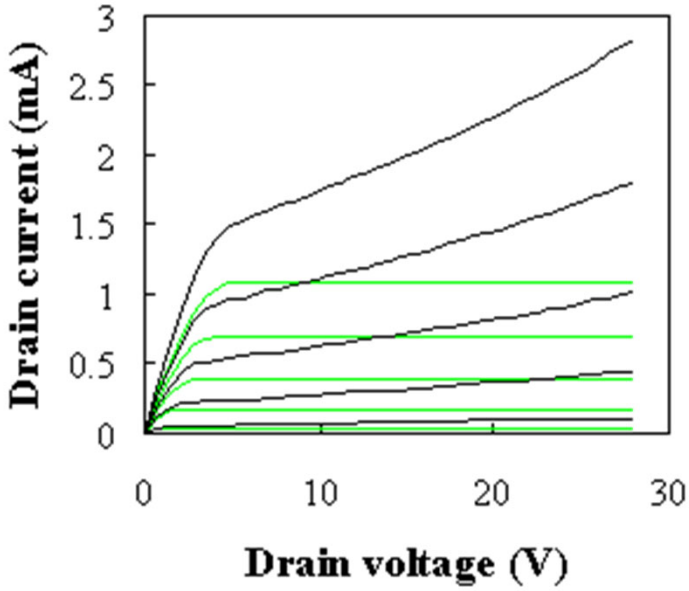


CMOS scaling:

- Gate length scaling

Channel length Modulation

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_t)^2\,(1 + \lambda V_{DS})$$



Where:

- λ = channel length modulation coefficient (depends on technology & channel length).
- Larger λ = stronger dependence on V_{DS}.
- For **long-channel MOSFETs**, $\lambda \approx 0$ (negligible).
- For **short-channel MOSFETs**, λ is significant.

Practical Implications

- Higher drain current than expected in saturation.
- Reduced output resistance ($r_o = 1/(\lambda I_D)$) → important in analog design.
- Impacts gain of MOS amplifiers.

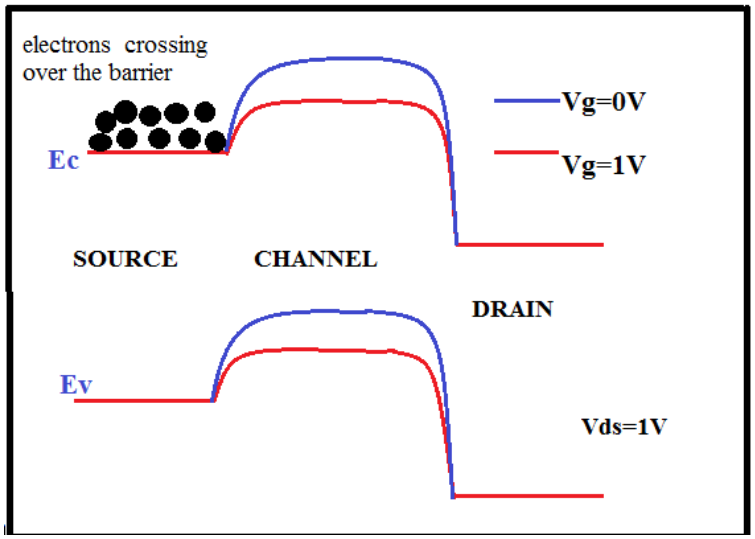
CMOS scaling:

- Gate length scaling

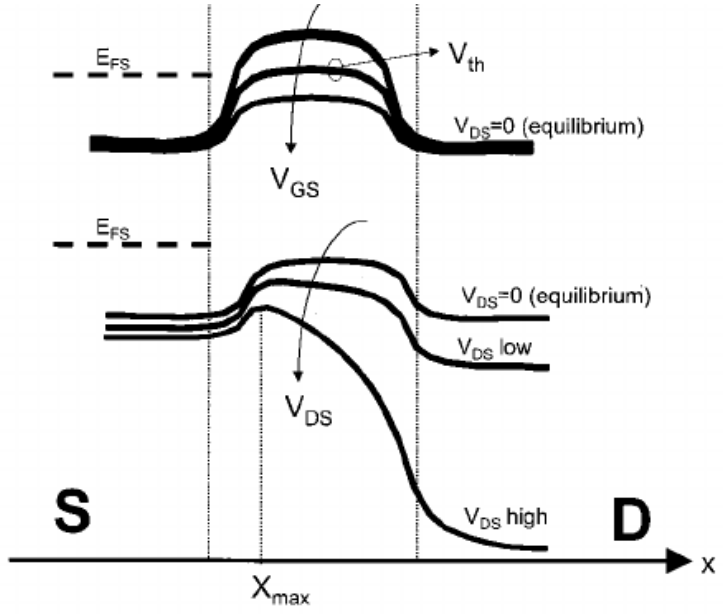
Drain Induced Barrier Lowering (DIBL)

As the source and drain get closer, they become electrostatically coupled, so that the drain bias can affect the potential barrier to carrier diffusion at the source junction

→ V_T decreases (*i.e.* OFF state leakage current increases)



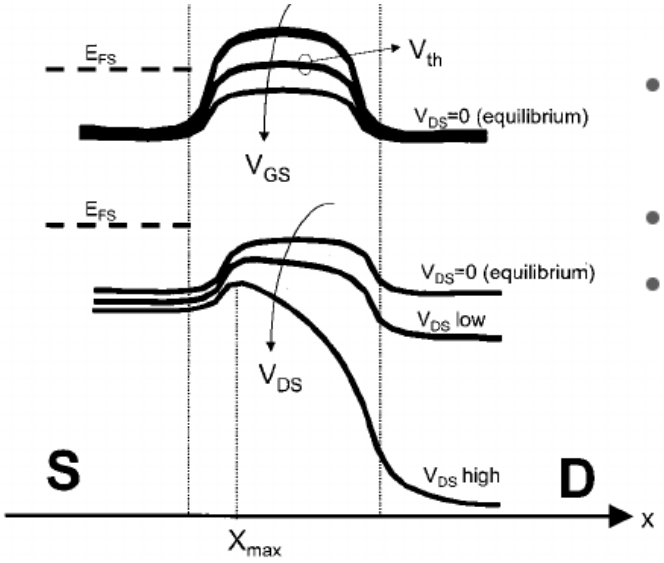
Long Channel



Short Channel

CMOS scaling:

- Gate length scaling



Short Channel

Drain Induced Barrier Lowering (DIBL)

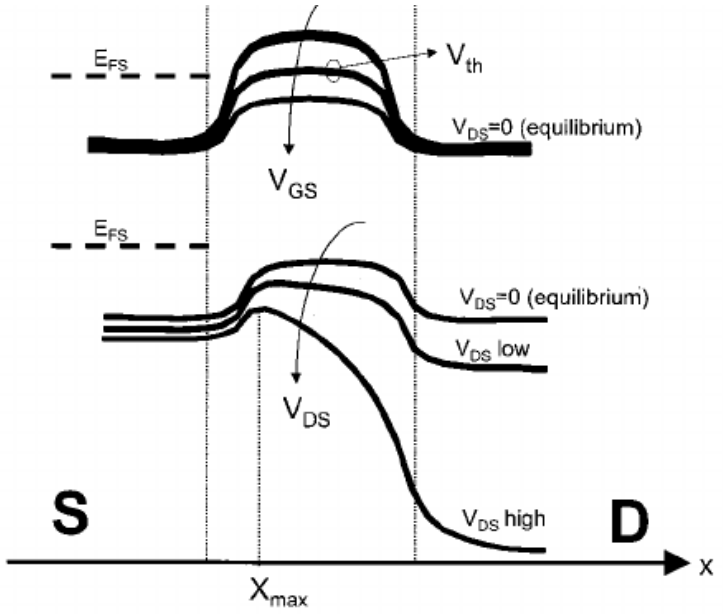
- At **high** V_{DS}, the electric field from the drain reduces the source-channel potential barrier.
- Effectively, the **threshold voltage** (V_T) **decreases** when V_{DS} increases.
- So the MOSFET turns on **at a lower** V_{GS} than expected.

Consequences

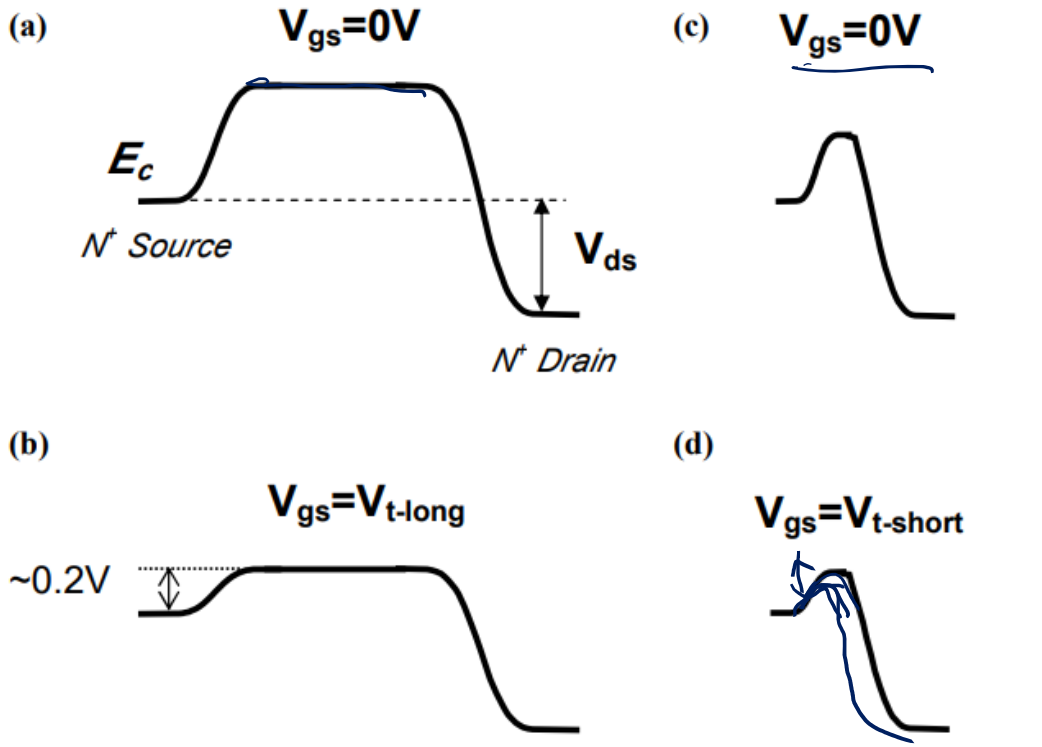
- **Threshold roll-off:** V_T reduces with shorter channel lengths and higher V_{DS}.
- **Increased off-state current:** Device leaks even when gate voltage is below nominal threshold → subthreshold leakage rises.
- **Reduced control:** Gate loses dominance over channel, hurting device scaling.

CMOS scaling:

- Gate length scaling



Drain Induced Barrier Lowering (DIBL)

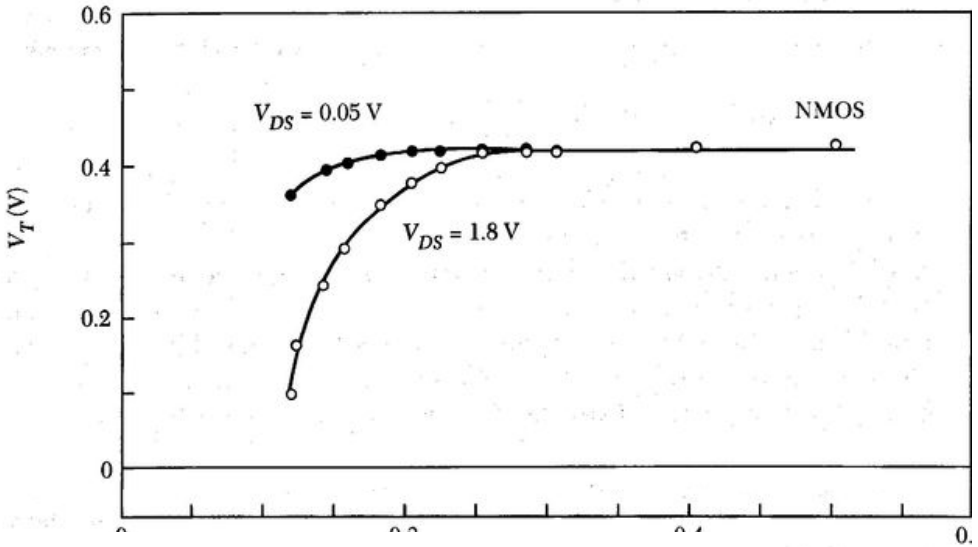


(a)-(d): Energy-band diagram from source to drain when V_{gs}=0V and V_{gs}=V_t

CMOS scaling:

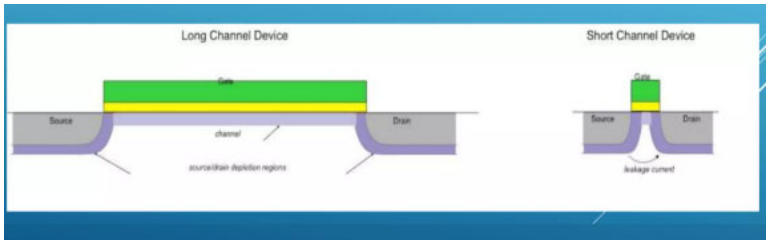
- Gate length scaling

Threshold Voltage: (V_{th}) Variation



Drain-Induced Barrier Lowering (DIBL)

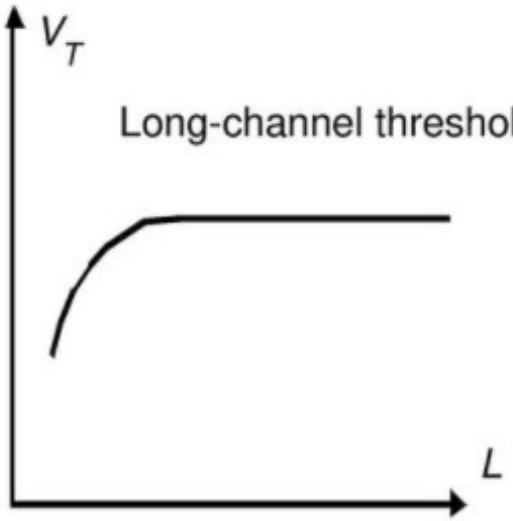
- At high drain bias, the drain electric field reduces the source-channel potential barrier.
- Effectively reduces V_t further (especially at short L).
- Leads to higher subthreshold leakage.



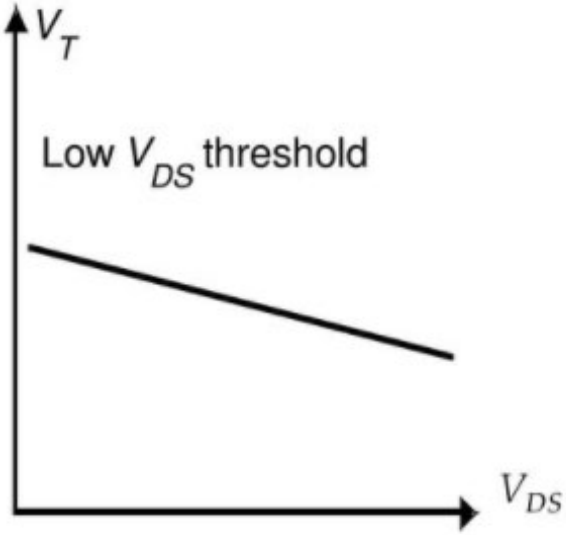
https://www.slideshare.net/slideshow/short-channel-effects/50101267

Gate length scaling

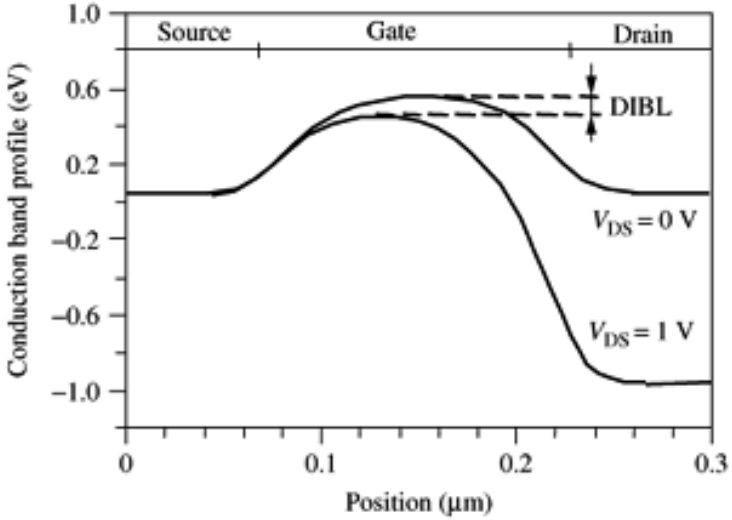
Threshold Voltage: (V_{th}) and DIBL



Threshold as a function of the length (for low V_{DS})



Drain-induced barrier lowering (for low L) …(DIBL)



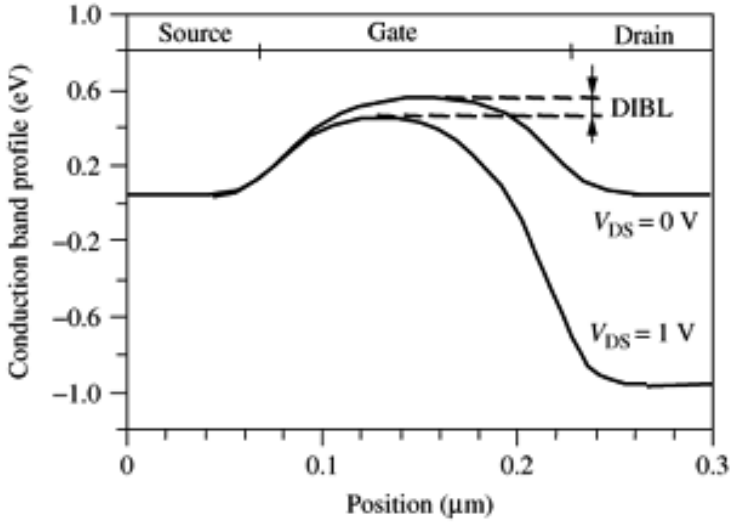
J. Ho, "Introduction and Short Channel Effects", 2014, Semiconductors

Gate length scaling

DIBL

Mitigation Techniques

- Use **lightly doped drain (LDD) structures**.
- **Halo/pocket implants** to control channel doping near source/drain.
- Move to **FinFETs / GAA FETs** for better electrostatic gate control.
- Reduce supply voltage to minimize drain fields.

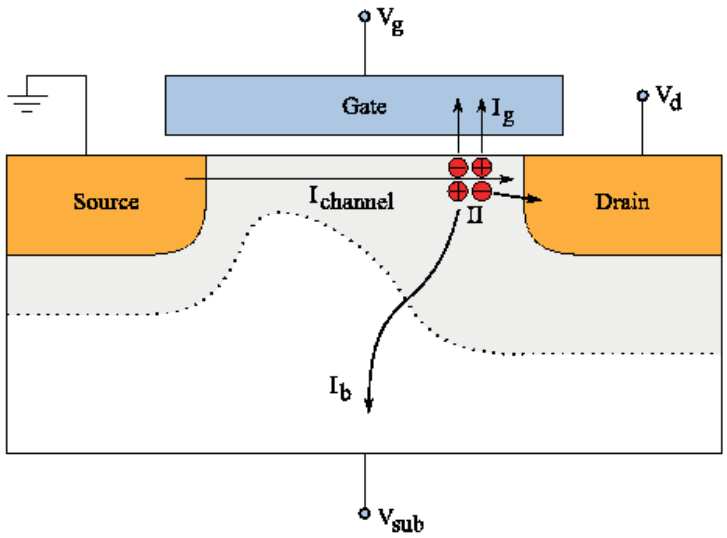


J. Ho, "Introduction and Short Channel Effects", 2014, Semiconductors

Hot carrier injection

High electric field near the substrate-oxide interface energizes the electrons or holes and they cross the substrate-oxide interface to enter the oxide layer. This phenomenon is known as hot carrier injection.

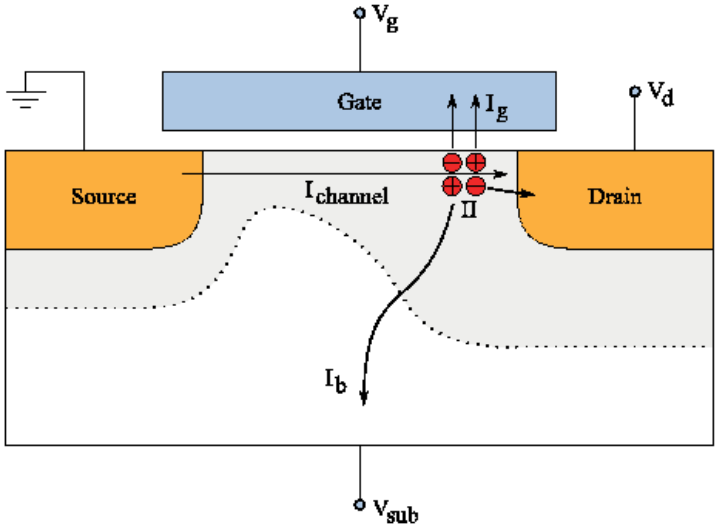
- When MOSFET operates at **high** V_{DS} (especially near saturation), the electric field near the **drain end of the channel** becomes very large.
- Carriers gain **very high kinetic energy** → become "hot carriers."



https://www.iue.tuwien.ac.at/phd/entner/node21.html

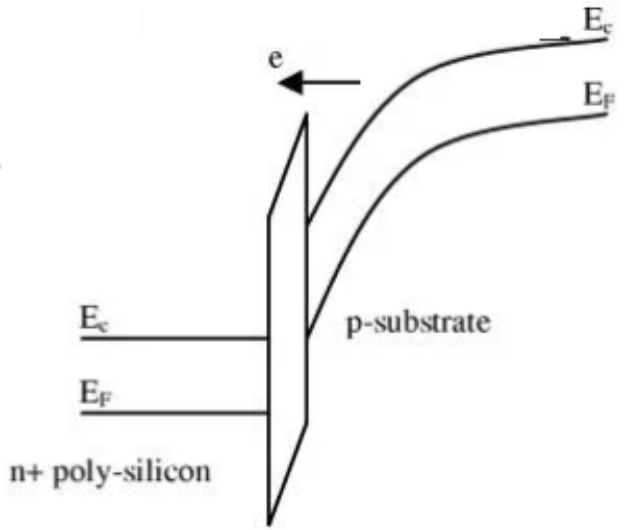
Hot carrier injection

- Cause **impact ionization** → generate electron–hole pairs.
- Get trapped in the oxide or at the **Si–SiO_2 interface**.



Consequences of HCI

- **Threshold voltage shift** (ΔV_T): Due to trapped charges.
- **Mobility degradation**: Carriers scatter more from interface states.
- **Drive current degradation**: MOSFET becomes weaker over time.
- **Long-term reliability issue**: Device lifetime is reduced.

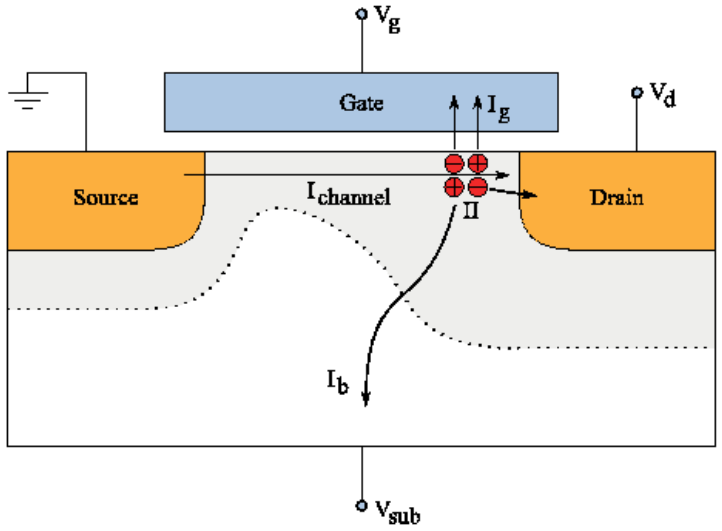


https://www.iue.tuwien.ac.at/phd/entner/node21.html

Hot carrier injection

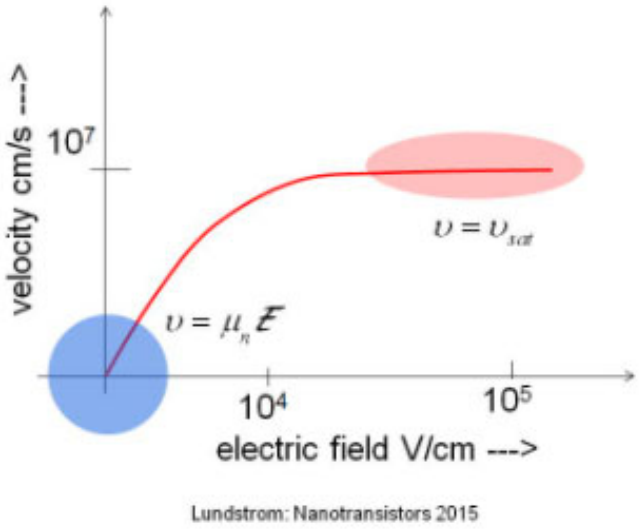
Mitigation Techniques

- **Lightly Doped Drain (LDD) structure**: Reduces electric field near drain.
- **Graded junctions** to spread out the high field region.
- **Lower VDD**: Reduces energy gained by carriers.
- **High-k dielectrics** with larger barrier heights.
- **FinFET/GAA structures**: Better electrostatic control, lower peak fields.



https://www.iue.tuwien.ac.at/phd/entner/node21.html

Velocity saturation



Lundstrom: Nanotransistors 2015

Carrier Transport in Long-Channel MOSFETs

- In long-channel devices, drift velocity of carriers follows:

$$v = \mu E$$

where μ = mobility, E = electric field.

- This linear relationship holds as long as the electric field is **moderate**.

Saturation Velocity (v_{sat})

- For **electrons in Si**: ~ 10^7 cm/s.
- For **holes in Si**: ~ 0.7×10^7 cm/s.

- Velocity of charge carriers is linearly proportional to the electric field and the proportionality constant is called as mobility of carrier.

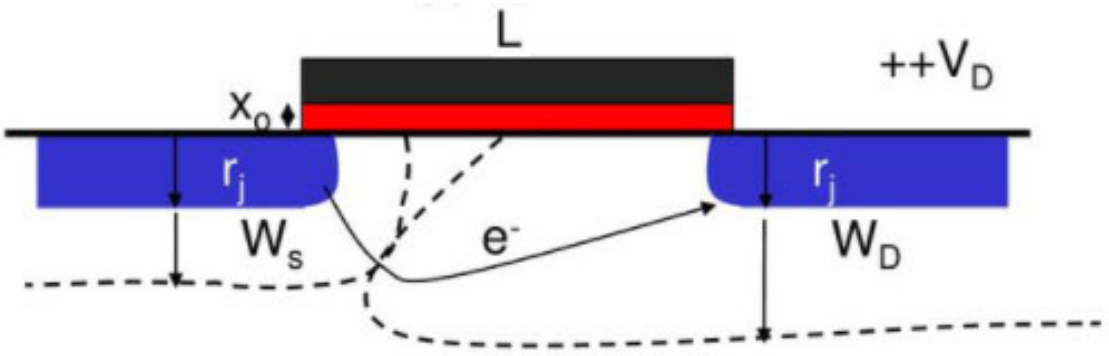
- But when we increase the electric field beyond certain velocity called as the thermal velocity or saturated velocity the velocity of the charge carrier does not change with electric field. The loss of energy is because of the collisions of carriers called as scattering effect.

Punch through effect

In a long-channel MOSFET:

- The **source and drain depletion regions** extend slightly into the channel.
- They remain separated by a central portion of neutral channel controlled by the **gate voltage**.
- Current flow only occurs when the gate creates an inversion layer.

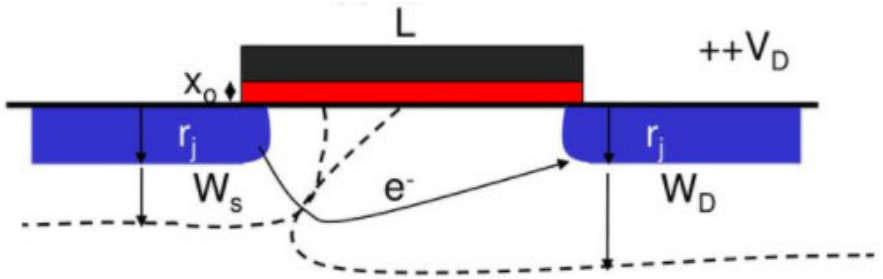
- In short channel devices, due to the proximity of drain and source terminals, the depletion region of both the terminals come together and eventually merge. In such a condition, "punch-through" is said to have taken place.



S&D depletions touch
– punch through

J. Ho, "Introduction and Short Channel Effects", 2014, Semiconductors

Punch through effect

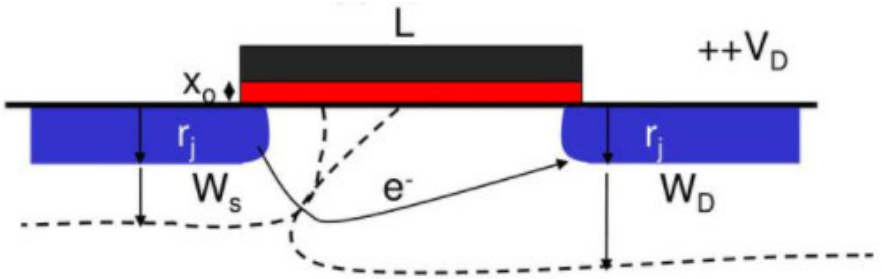


S&D depletions touch
– punch through

- The **drain and source depletion regions expand** and may overlap inside the channel.
- This creates a **continuous depletion path** from source to drain, **bypassing gate control**.
- Drain current no longer controlled by gate voltage
- Drain current does not saturate
- High subthreshold current

J. Ho, “Introduction and Short Channel Effects”, 2014, Semiconductors

Punch through effect



S&D depletions touch
– punch through

Mitigation Techniques

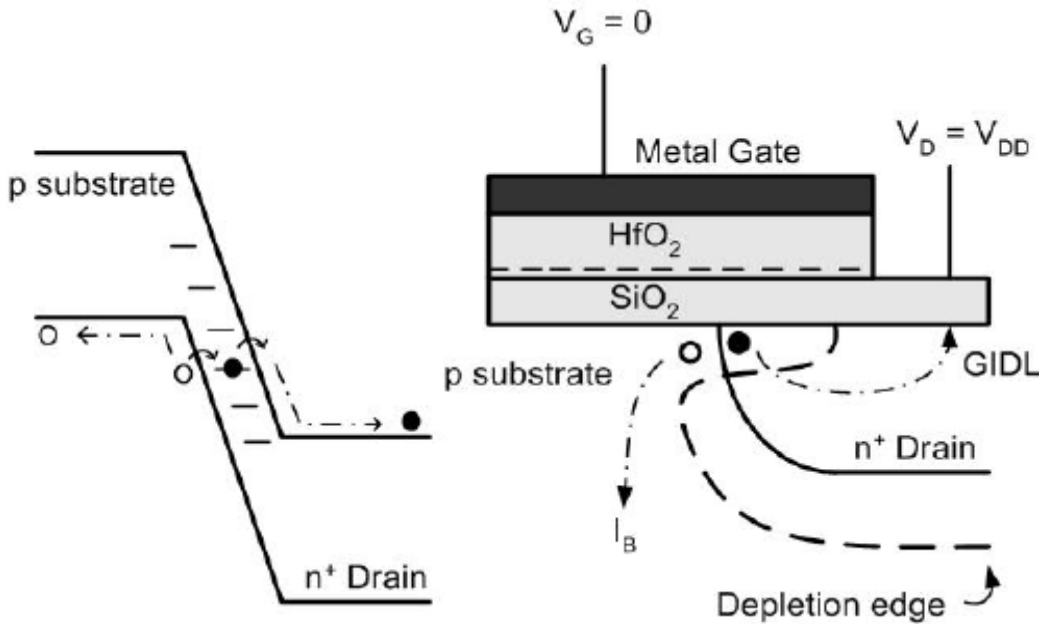
- **Increase channel doping** → narrows depletion regions.
- Use **retrograde wells** or **halo implants** near source/drain to suppress depletion overlap.
- Move to **multi-gate devices (FinFET, GAA)** for better gate electrostatic control.
- **Reduce drain bias** to minimize field penetration.

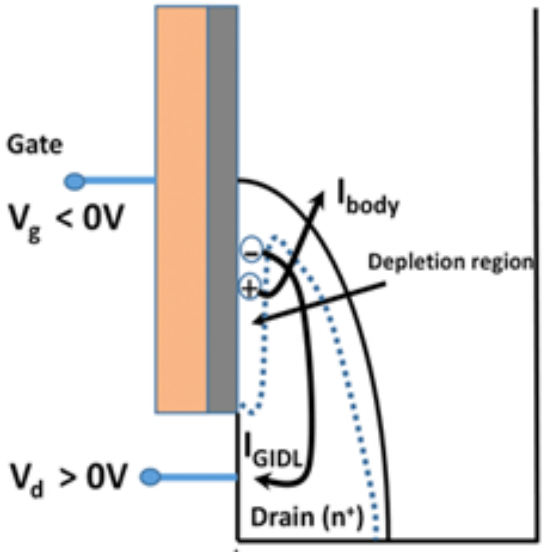
J. Ho, “Introduction and Short Channel Effects”, 2014, Semiconductors

Gate-Induced Drain Leakage (GIDL)

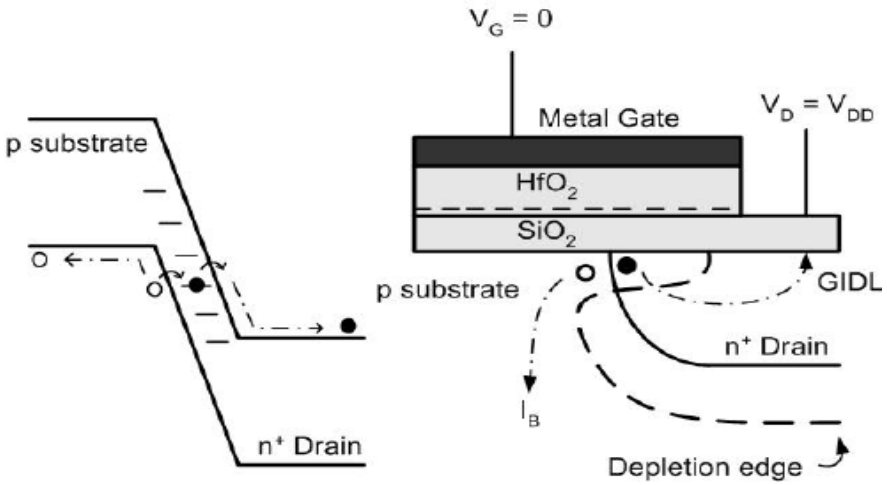
Where It Occurs

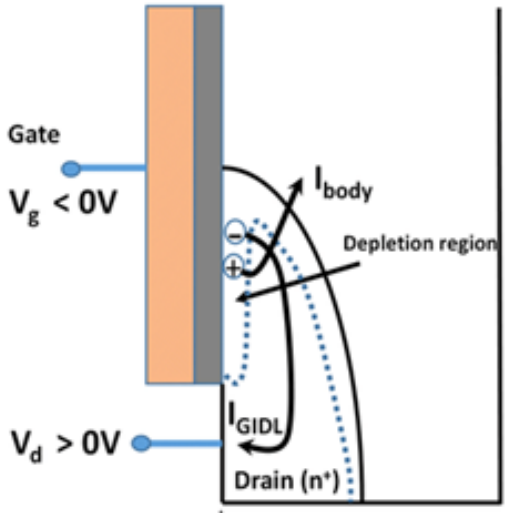
- Happens in the **drain–gate overlap region** of a MOSFET.
- Most significant when the **gate is held at a low (negative or 0 V) voltage** while the **drain is at a high voltage**.
- Common in **OFF state leakage** in short-channel devices.





Gate-Induced Drain Leakage (GIDL)





Physical Mechanism

A **strong electric field** develops between the drain and gate in the overlap region when:

- Gate is at low potential, and Drain is at high potential.

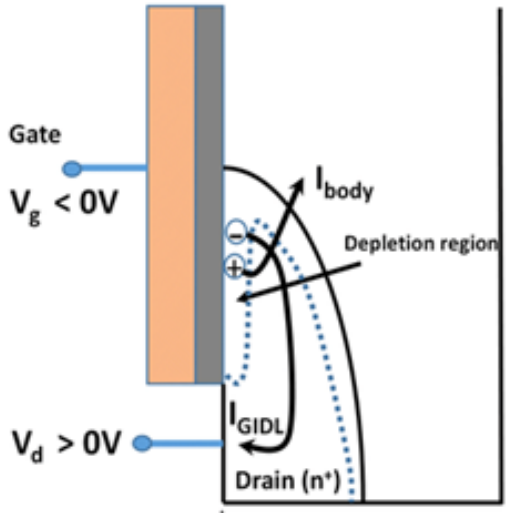
This high vertical electric field causes **band-to-band tunneling (BTBT)**:

- This generates electron–hole pairs.
- Electrons go into the drain → leakage current appears even though MOSFET is “off.”

Gate-Induced Drain Leakage (GIDL)

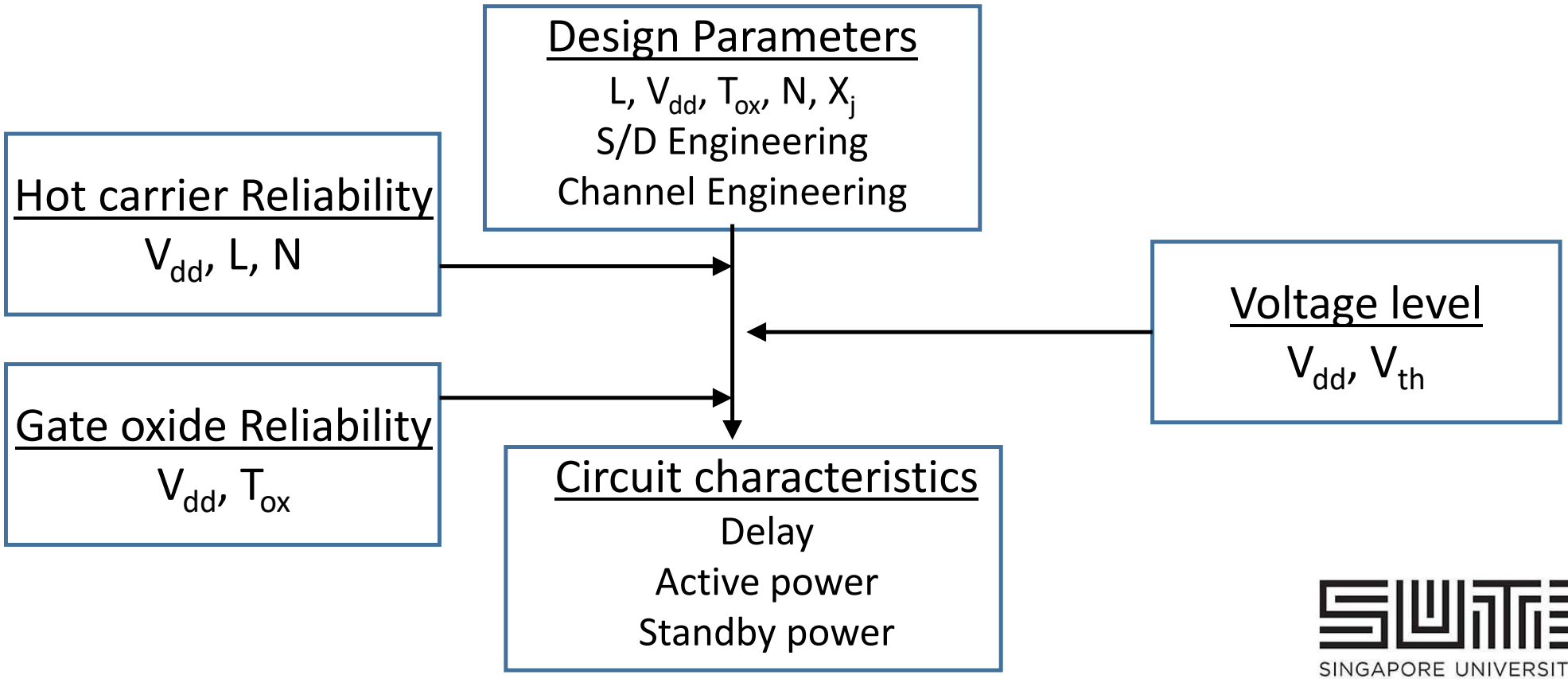
Mitigation Techniques

- Lightly Doped Drain (LDD) structures → reduce electric field.
- High-k dielectrics / thicker oxides at drain-edge → reduce tunneling.
- Device engineering (FinFET, GAA) improves electrostatics → suppresses GIDL.



Nano-MOSFET:

MOSFET design methodology for Logic (Digital)Technology



Ref: N. Bhat, CeNSE, IISc Bangalore- Nanoelectronic Device Technology

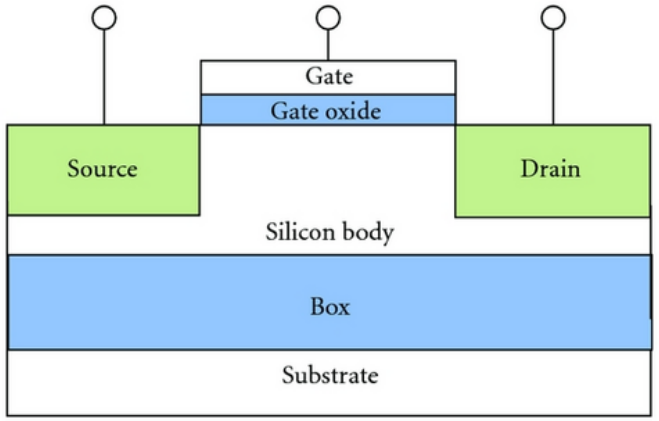
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Why SOI

- With Silicon-On-Insulator (SOI) wafers, transistors are formed in thin layers of silicon that are isolated from the main body of the wafer by a layer of electrical insulator, usually silicon dioxide.

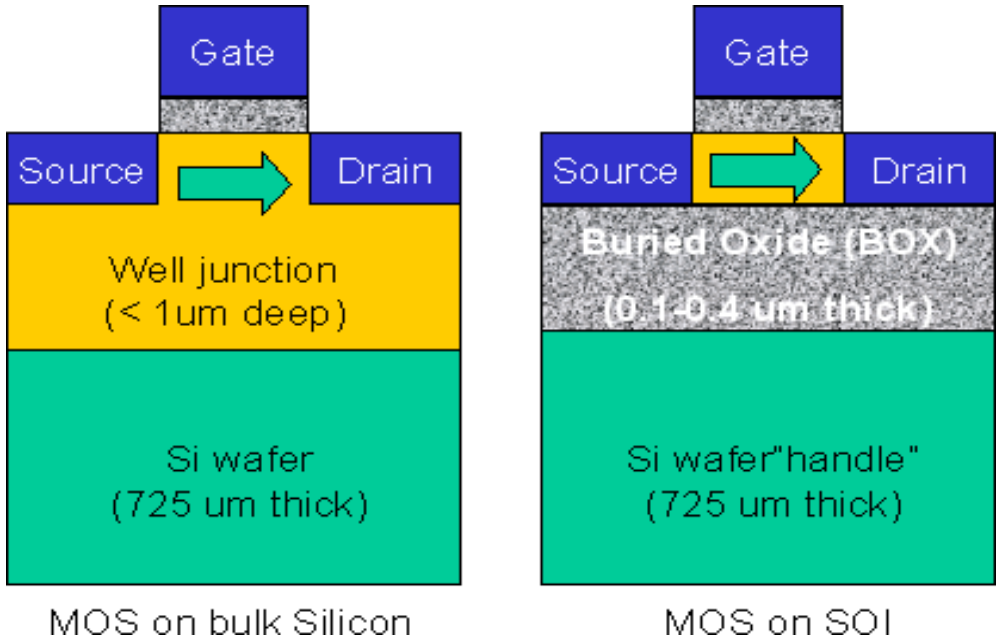
- The silicon layer thickness ranges from several microns for electrical power switching devices to less than 50 nm for high-performance microprocessors.

- Isolating the active transistor from the rest of the silicon substrate reduces the electrical current leakage that would otherwise degrade the performance of the transistor.



https://doi.org/10.1155/2012/565827

Silicon-On-Insulator (SOI) MOSFET

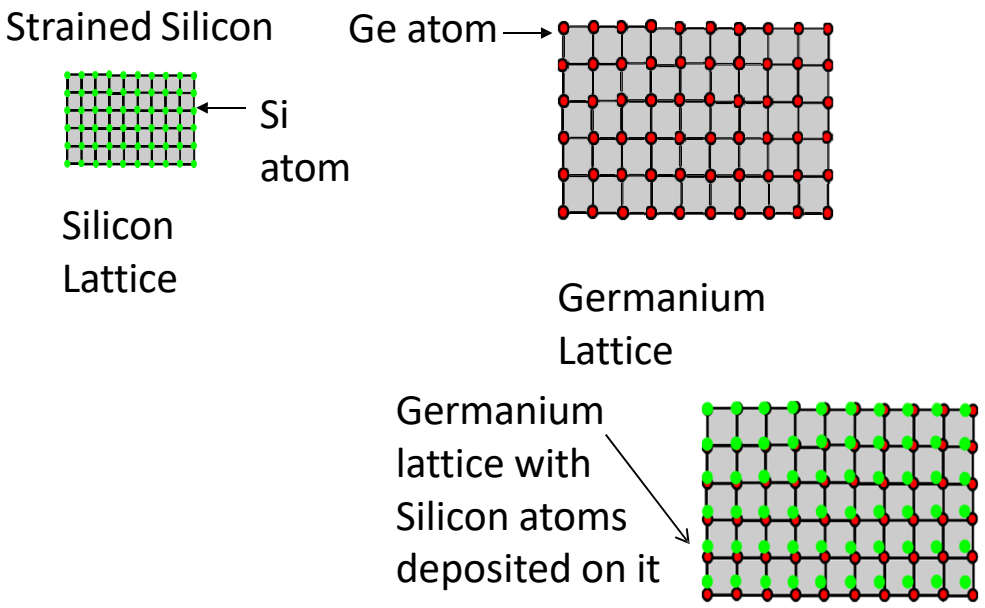


Silicon-on-Insulator (SOI) wafers consist of three layers: a thin (20 nm to several microns, depending on the application) layer of single-crystal silicon on a thick (100 to 400 nm) silicon dioxide layer that is bonded to a conventional "handle" wafer. The entire transistor is located in the thin top layer of silicon and electrically isolated from the bulk wafer by the buried oxide (BOX) layer.

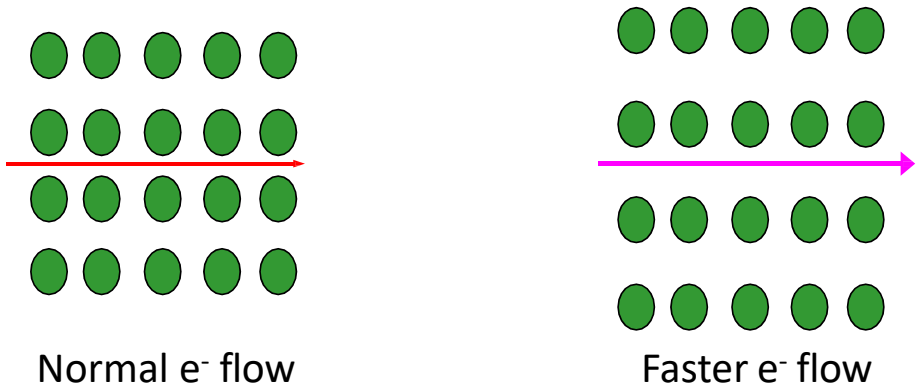
Strained Silicon Wafers

Strained-Si on SiGe virtual substrate:

- Grow a relaxed SiGe layer → lattice mismatch stretches thin Si grown on top.
- Provides **tensile strain** → boosts **NMOS electron mobility**.
- Used in early 90 nm–65 nm processes.



Higher Mobility in Strained Si

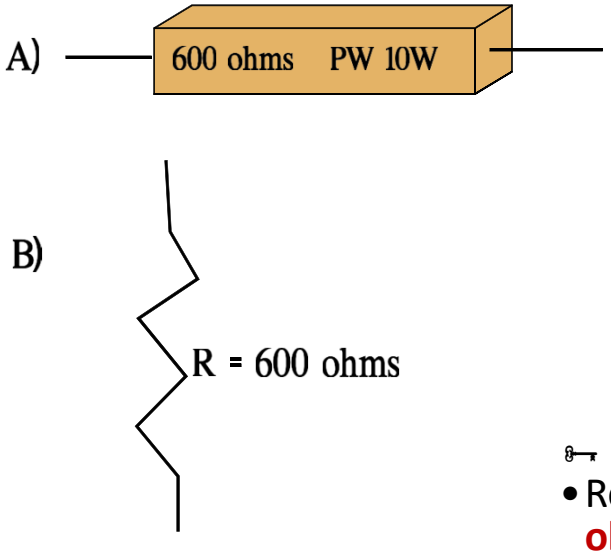


Stressing, stretching, or straining, the Si lattice allows electrons to flow with less resistance. Lower resistance means that transistors switch faster.

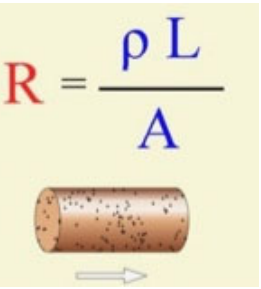
Resistors on ICs

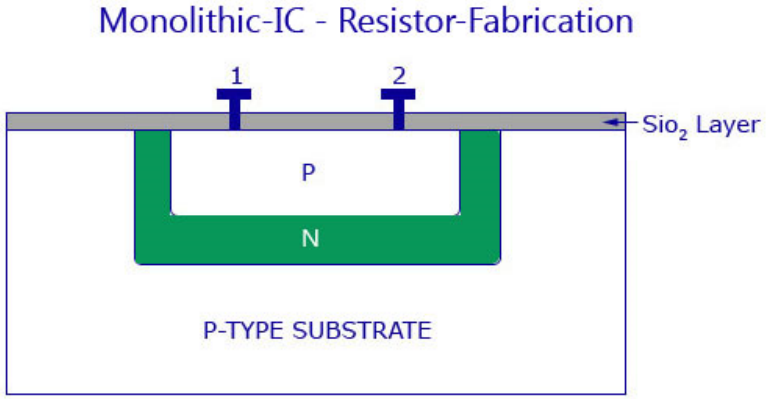
Resistors are fabricated on ICs using

- either the diffused resistor method- doping specific regions of the semiconductor material,
- or the thin-film resistor method, where a resistive thin film is deposited and patterned on the substrate



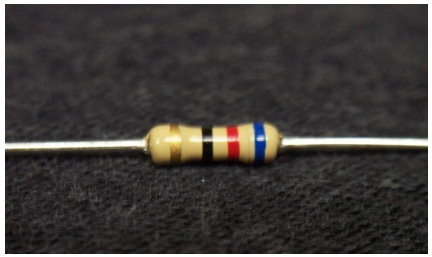
$$R = \frac{\rho L}{A}$$





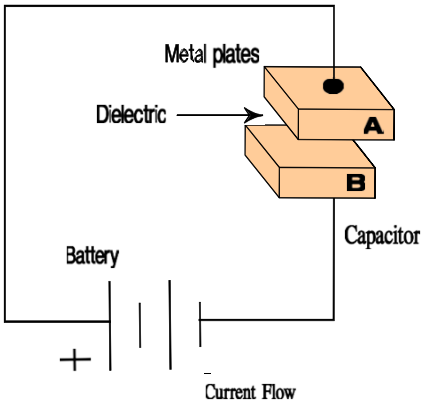
🗝 Key Theme:

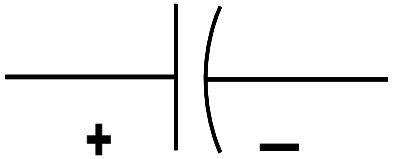
- Resistance is measured in **ohms**



Capacitors on ICs

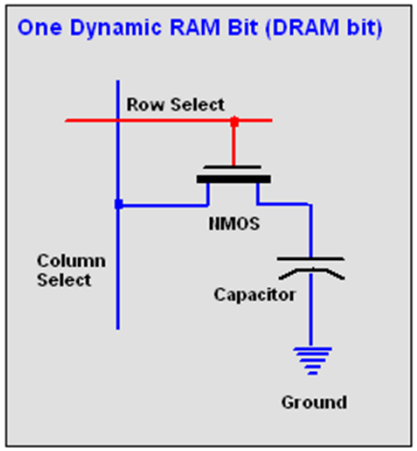
- Capacitance is measured in **farads** and contains two conductive plates and a dielectric between the two plates

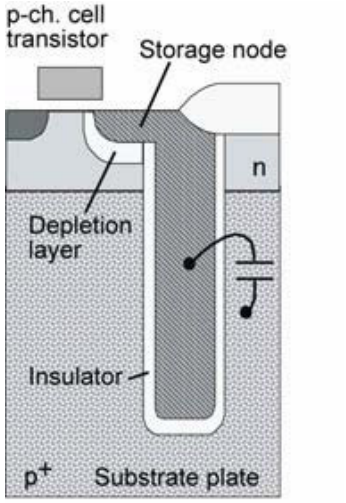




Symbol for a capacitor

Capacitor in DRAM





https://www.intechopen.com/chapters/9855