ECE520 – VLSI Design

Lecture 4: Device Scaling Issues & Advanced MOS

Payman Zarkesh-Ha

Office: ECE Bldg. 230B
Office hours: Wednesday 2:00-3:00PM or by appointment
E-mail: pzarkesh@unm.edu

Review of Last Lecture

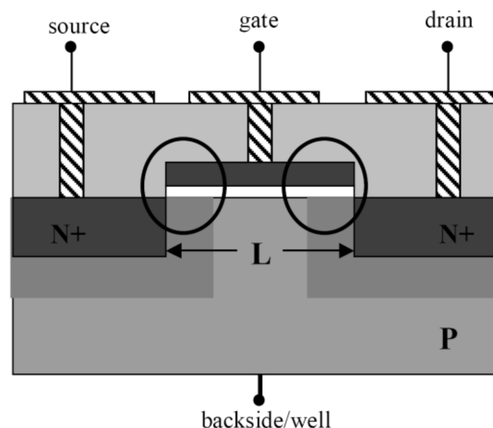
- Overview of “Static Parameters of Long Channel MOSFET”
- “Dynamic Parameters of Long Channel MOSFET”
- MOSFET Parasitic Capacitances
  - Overlap capacitances
  - Channel capacitances
  - Junction capacitances

Today’s Lecture

- ❑ **Device Modeling Issues for “Short Channel” MOS**
- ❑ **Short Channel Effects**
  - **Source-Drain resistance**
  - **Subthreshold conduction**
  - **Velocity saturation**
  - **Mobility degradation**
  - **Threshold voltage rolloff**
  - **DIBL effect**
  - **Punch through**
  - **Hot electron**
  - **Narrow channel effect**
- ❑ **Device Scaling Issues**

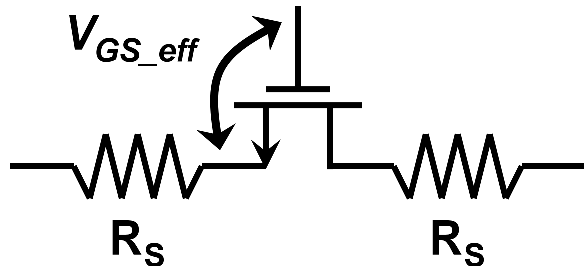
Device Scaling Issues

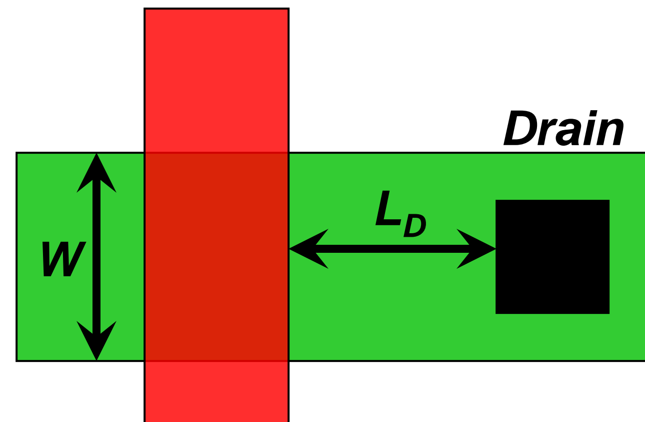
- When the channel length becomes comparable to other dimensions like depth of the source/drain junctions and the width of the depletion regions the "long-channel" approximations made previously break down
  - Assumptions such as, current flows only on surface, electric field is only in the direction of current flow, etc., are no longer true
  - Such a short channel device can not be adequately described by simple one dimensional model. Hence, a two dimensional model is widely used.
  - With transistor scaling, junctions are made shallower & contacts windows are made smaller while their depth is increased



Source-Drain Resistance

- With transistor scaling, junctions are made shallower & contacts windows are made smaller while their depth is increased

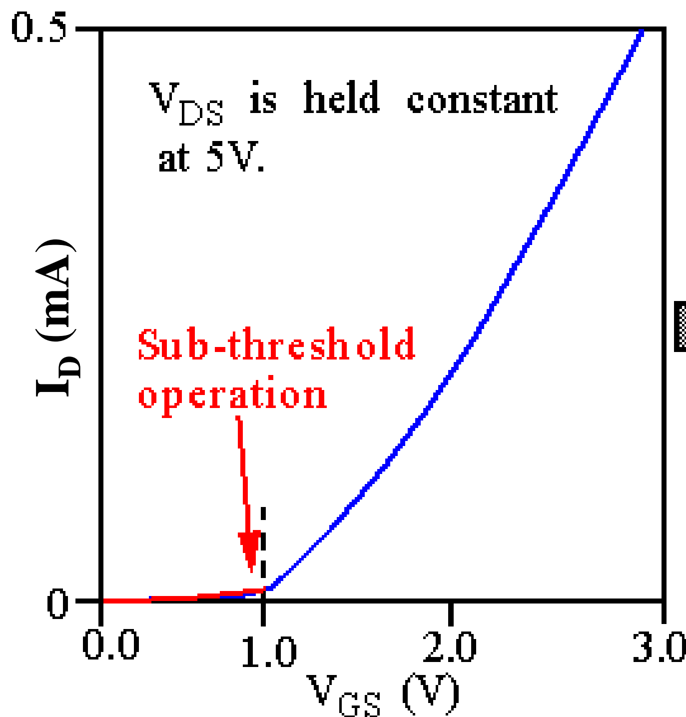


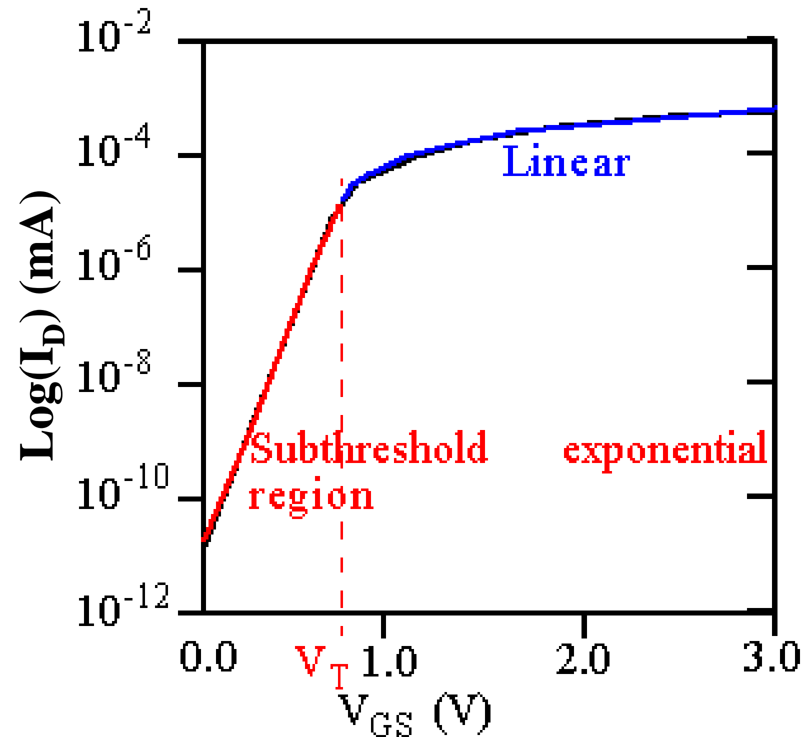


- Technology and design objective is to reduce source-drain resistance
- Often source drain regions are covered by titanium or tungsten (silicidation) to reduce the resistance

Subthreshold Conduction

- I_{DS} does not equal to zero even with $V_{GS} = 0$
  - To get I_{DS}=0 need V_{DS}=0
- This is known as subthreshold conduction





Subthreshold Conduction

- Subthreshold conduction happens because of parasitic bipolar transistor [i.e. n+(source),p(bulk),n+(drain)]
  - The current in subthreshold region can be approximated by:

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

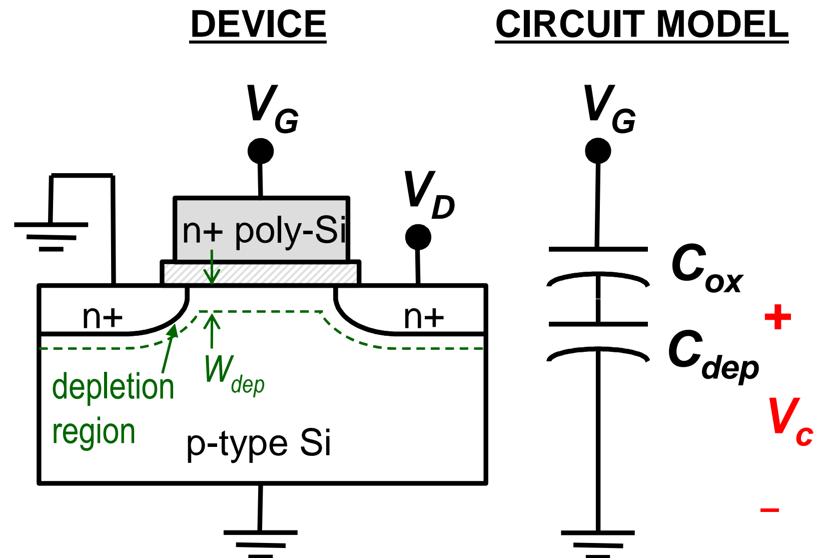
- I_{off} is defined as the I_{DS} when V_{GS}=0 and high V_{DS}
  - This is the transistor source to drain leakage (dominant leakage)
  - Sum of I_{off} across the die is the static power component
  - Static power is a limiting component of the total power (particularly for mobile markets)
- Key points:
  - I_{DS} depends exponentially on V_{GS} below threshold
  - I_{off} increases 3-10x per process generation (high performance)
  - I_{off} increases 8-12x per 100ºC

Qualitative Explanation

- The channel *Vc* (at the Si surface) is capacitively coupled to the gate voltage V_{GS}

  - Using the capacitive voltage divider formula:

$$\Delta V_c = \frac{C_{ox}}{C_{ox} + C_{dep}} \Delta V_G$$



  - The forward bias on the channel-source pn junction increases with V_{GS} scaled by the factor Cox / (Cox+Cdep)

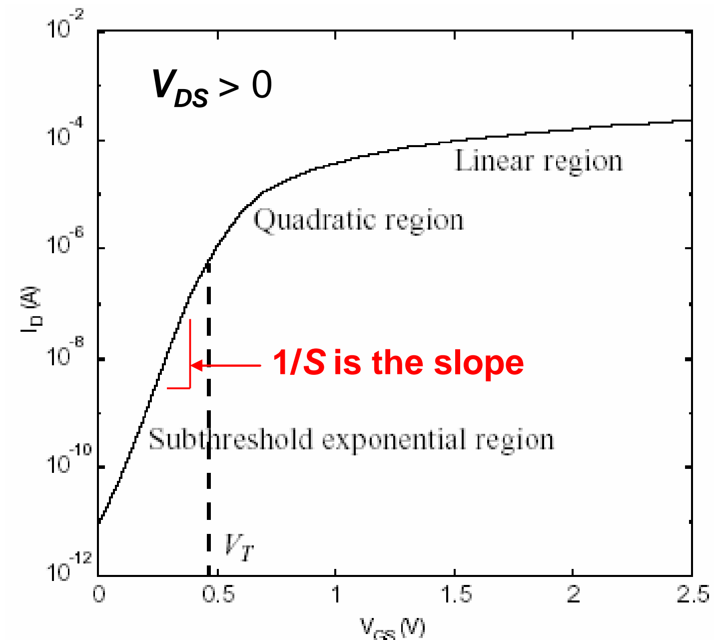
$$\Rightarrow \quad n = \frac{C_{ox} + C_{dep}}{C_{ox}} = 1 + \frac{C_{dep}}{C_{ox}}$$

Subthreshold Swing

- Subthreshold swing, *S*, is the inverse rate of decrease in current (in decade) with respect to V_{GS}

$$S = \frac{dV_G}{d(Log_{10} I_D)} = Ln(10)\frac{dV_G}{d(Ln I_D)} \quad \Rightarrow \quad S = \frac{KT}{q} Ln(10)\left(1 + \frac{C_{dep}}{C_{ox}}\right)$$

- Subthreshold swing is an indicator of the transistor quality
- Ideally subthreshold swing is 60mV/decade, but in real transistors it is always more (due to non-vanishing C_{dep})

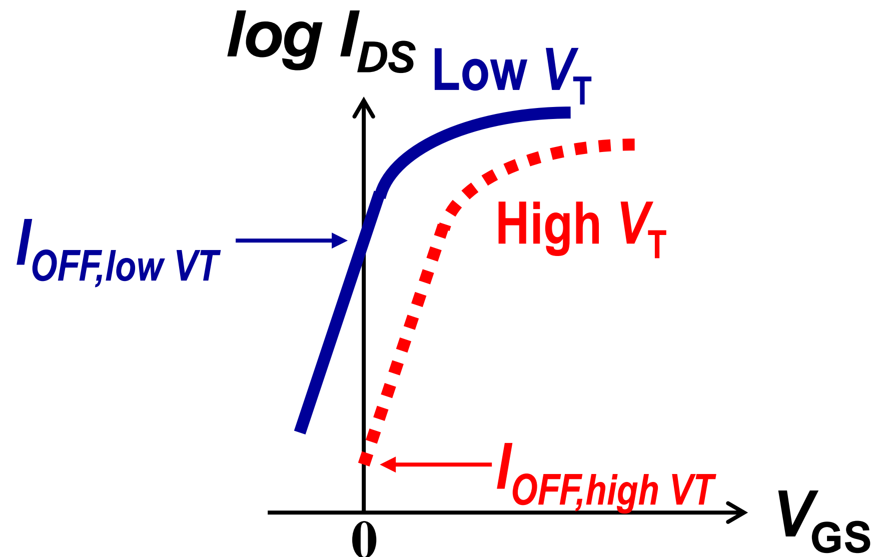


V_T Design Trade-off

- Low V_T is desirable for **high ON current**

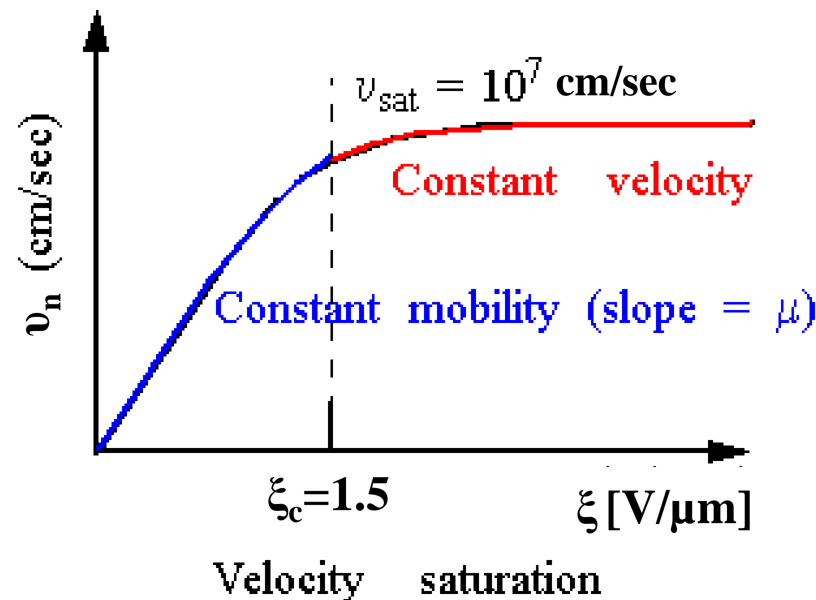
$$I_{Dsat} \propto (V_{DD} - V_T)^2$$

- But high V_T is needed for **low OFF current**



Velocity Saturation

- Velocity of carriers (*v*) proportional to the electric field (ξ) is true only for values of ξ less than 1.5V/μm.
- Consider a 0.6μ device velocity saturation will start to occur for a V_{DS} of 1V.



Velocity saturation

Velocity Saturation Modeling

- Assume that the velocity saturates abruptly at ξ_c:

$$\upsilon = \mu_n \xi \qquad \textit{for} \quad \xi \le \xi_c$$

$$\upsilon = \upsilon_{sat} = \mu_n \xi_c \qquad \textit{for} \quad \xi \ge \xi_c$$

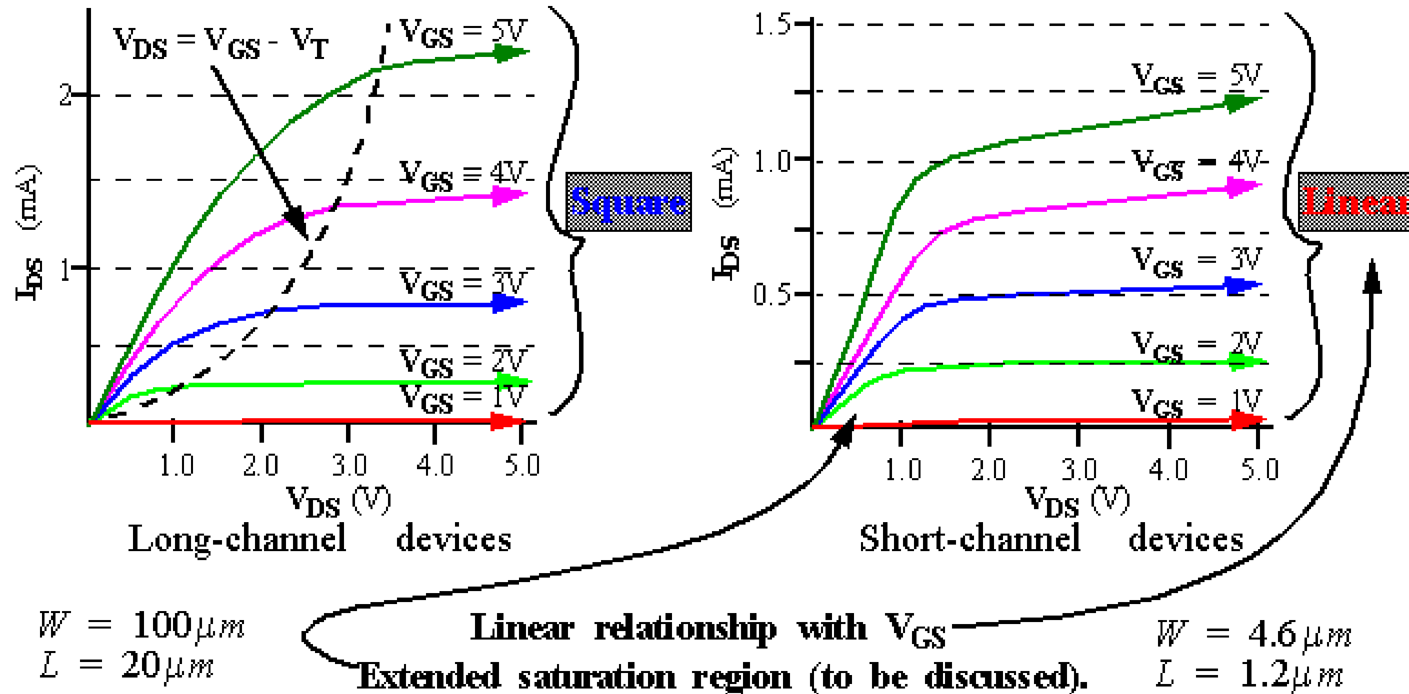
- Then, the drain-source voltage at saturation becomes:

$$V_{Dsat} \approx L\xi_c = \frac{L\nu_{sat}}{\mu_n}$$

- This results in the saturation region drain current of:

$$I_{Dsat} \approx \nu_{sat} C_{ox} W \left(V_{GS} - V_T - \frac{V_{Dsat}}{2} \right)$$

Velocity Saturation Modeling



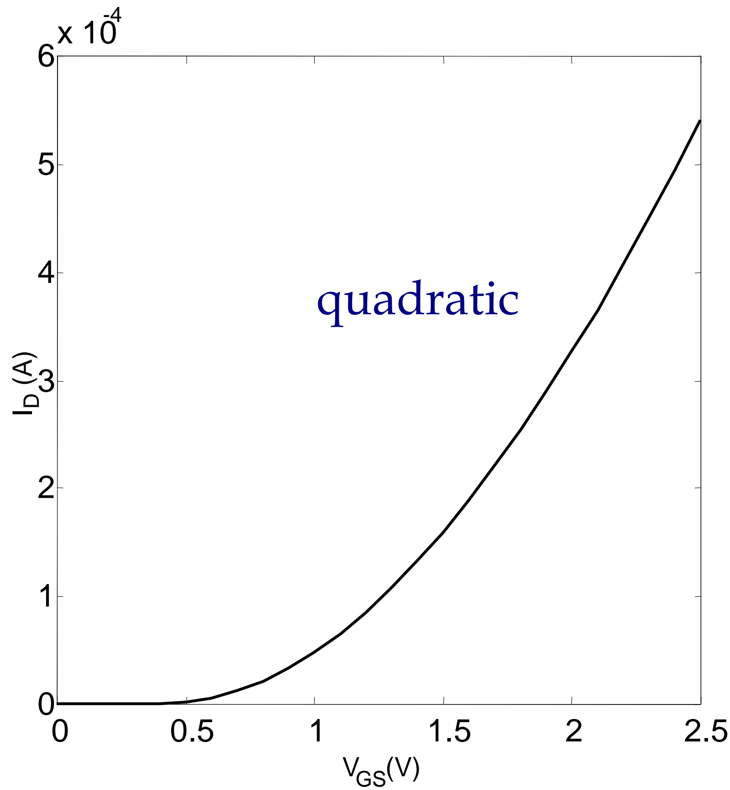
- These I-V curves were derived from devices in the same technology (1.2um) but different sizes.
- Long-channel device on the left has W=100um and L=20um while short-channel device on the right has W=4.6um and L=1.2um.
- Velocity saturation reduces the drain current by 53% for V_{GS} = 5.0V and V_{DS} = 5V, (1.2mA versus 2.3mA).

Velocity Saturation Modeling

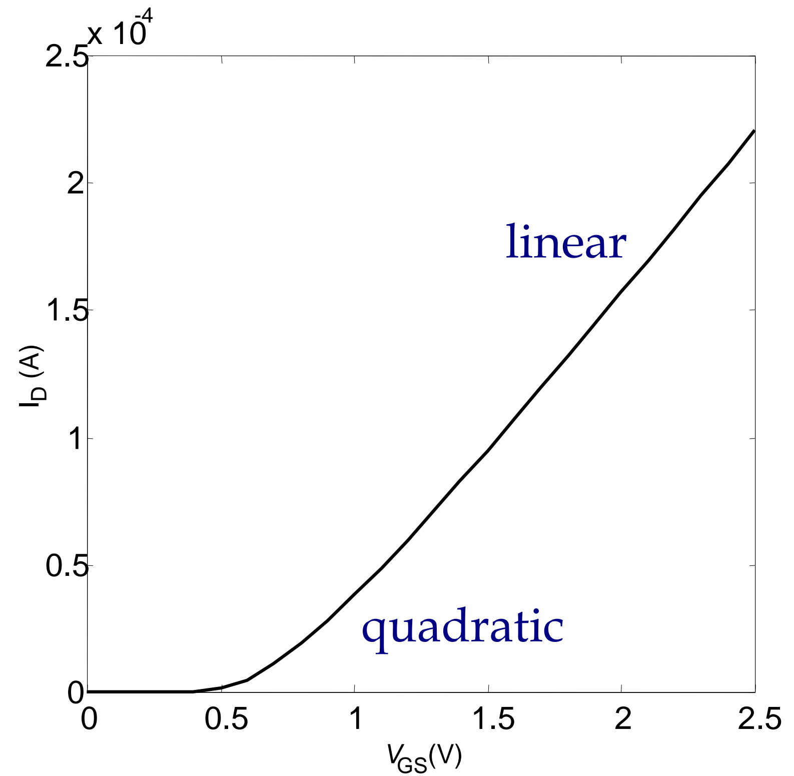
- **Key points:**
  - I_{Dsat} is a linear function of (V_{GS}-V_T) - reducing the operating voltage *does not* have such a significant effect in submicron devices as it would for long-channel devices (which is good)
  - I_{Dsat} is independent of L - current drive *cannot* be further improved by decreasing the channel length (which is bad)

$$I_{Dsat} \approx \nu_{sat} C_{ox} W \left(V_{GS} - V_T - \frac{V_{Dsat}}{2} \right)$$

I_D Versus V_{GS}

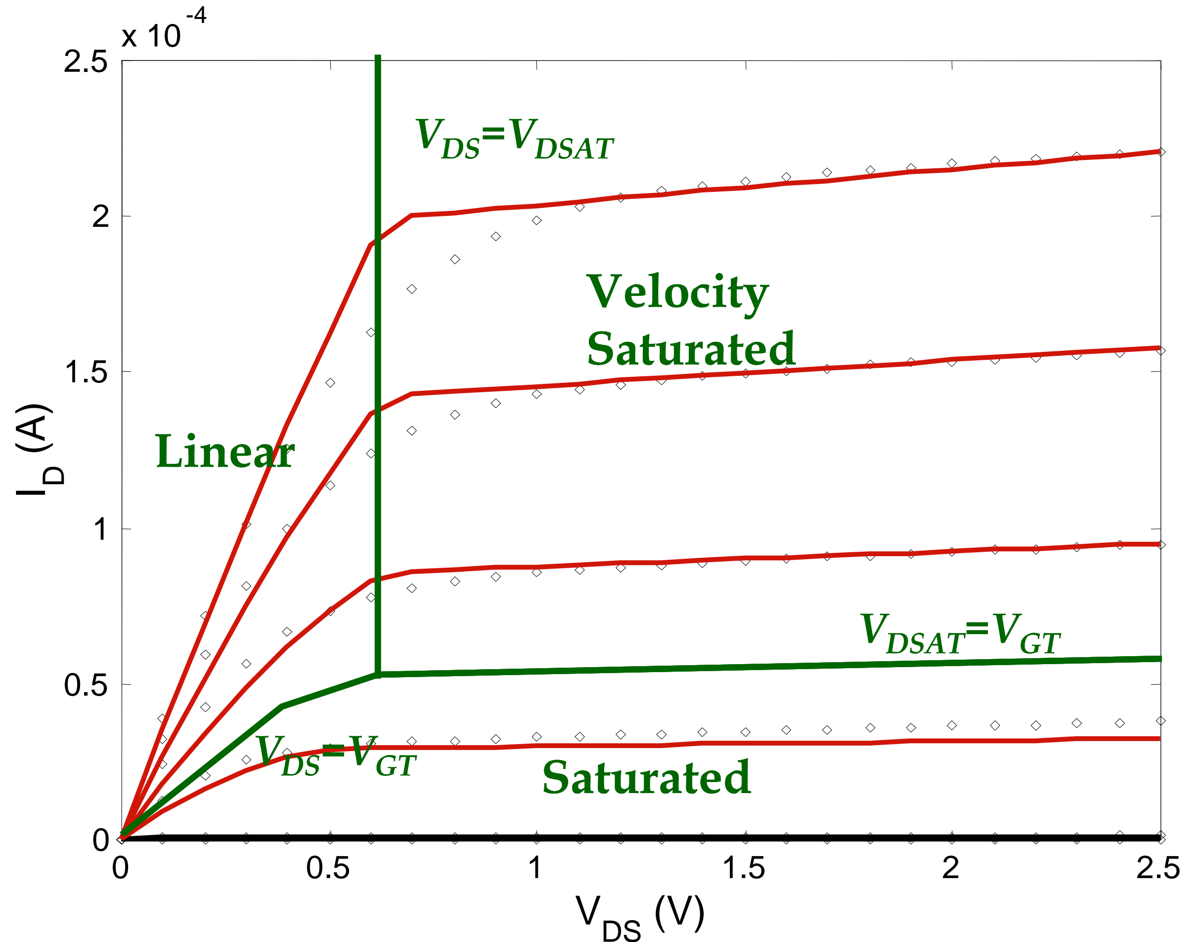


Long Channel

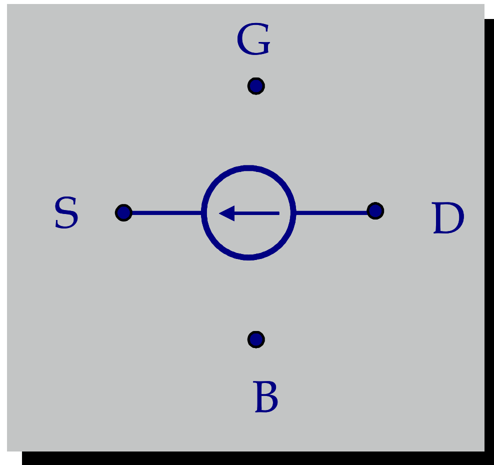


Short Channel

Revised MOS Region of Operation



Unified Model for Manual Analysis



$$I_D = 0 \text{ for } V_{GT} \leq 0$$

$$I_D = k'\frac{W}{L}\left(V_{GT}V_{min} - \frac{V_{min}^2}{2}\right)(1 + \lambda V_{DS}) \text{ for } V_{GT} \geq 0$$

with $V_{min} = \min(V_{GT}, V_{DS}, V_{DSAT})$,

$V_{GT} = V_{GS} - V_T$,

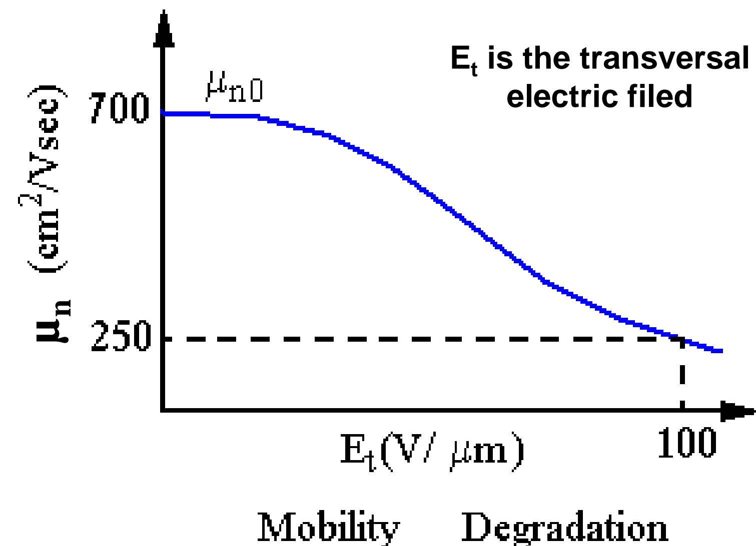
and $V_T = V_{T0} + \gamma(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|-2\phi_F|})$

Mobility Degradation

- Smaller geometries of devices result in greater electric fields, e.g. E = V/distance
- The *vertical* components of the field through the channel (due to both the drain and the gate) are strong and can affect the mobility of the carriers in the channel (surface mobility)
- Mobility degradation can be approximated by:

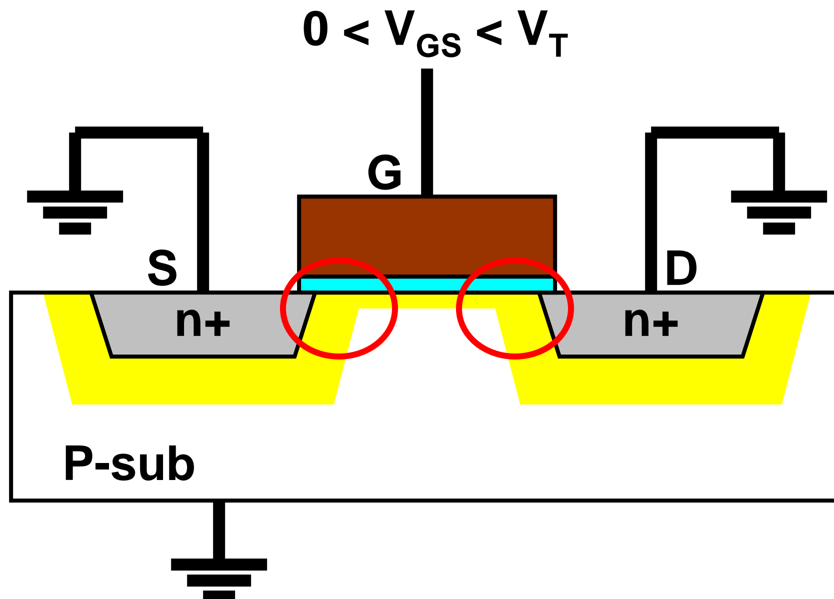
$$\mu_{eff} = \frac{\mu_0}{1 + \eta(V_{GS} - V_T)}$$

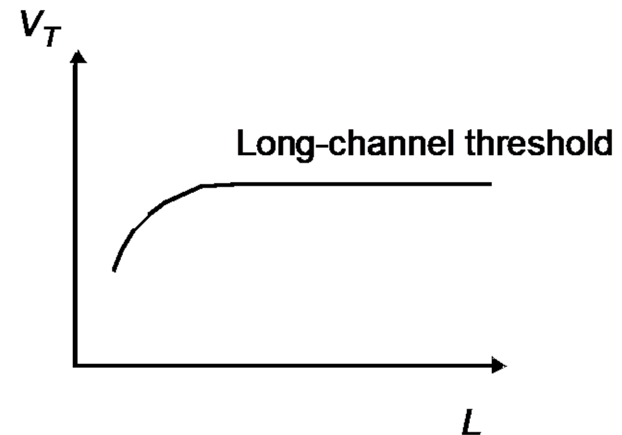
where μo is the low field mobility and η is an empirical factor



Threshold Voltage Rolloff

- Because of the partial channel depletion caused by Source and Drain, a smaller threshold is enough to create strong inversion.
- Therefore, the threshold voltage is a function of channel length, L.

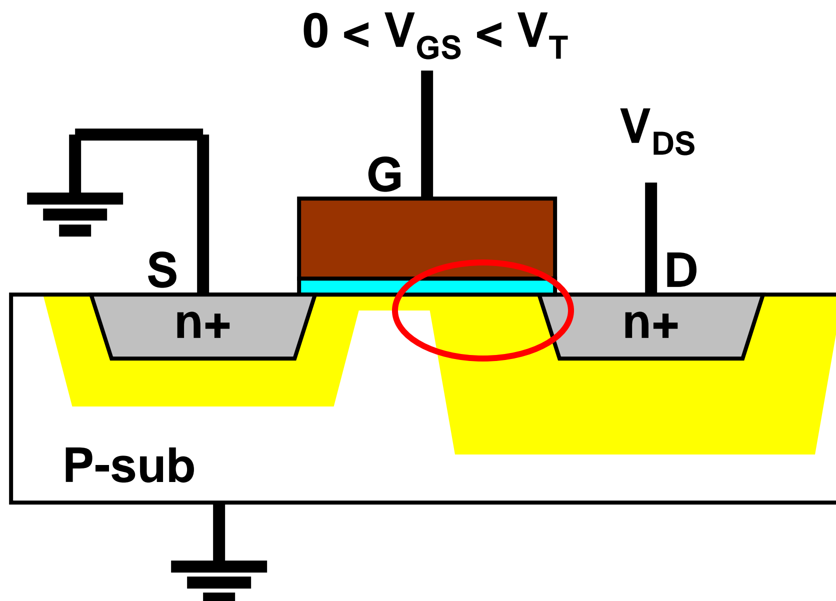


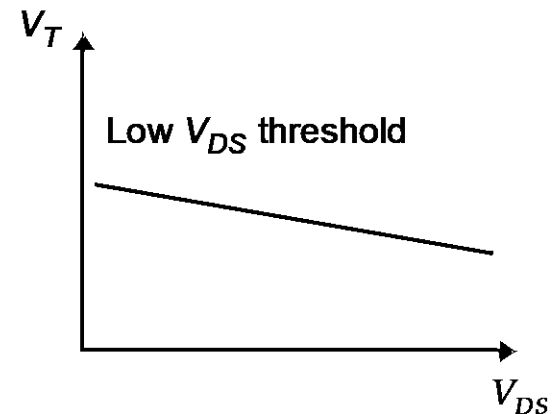


Threshold as a function of the length (for low V_{DS})

DIBL Effect

- Raising the drain potential increases the drain junction depletion region, reducing threshold voltage furthermore.
- This is called Drain Induced Barrier Lowering (DIBL).
- Therefore, threshold voltage in short channel device becomes a function of operating voltages.

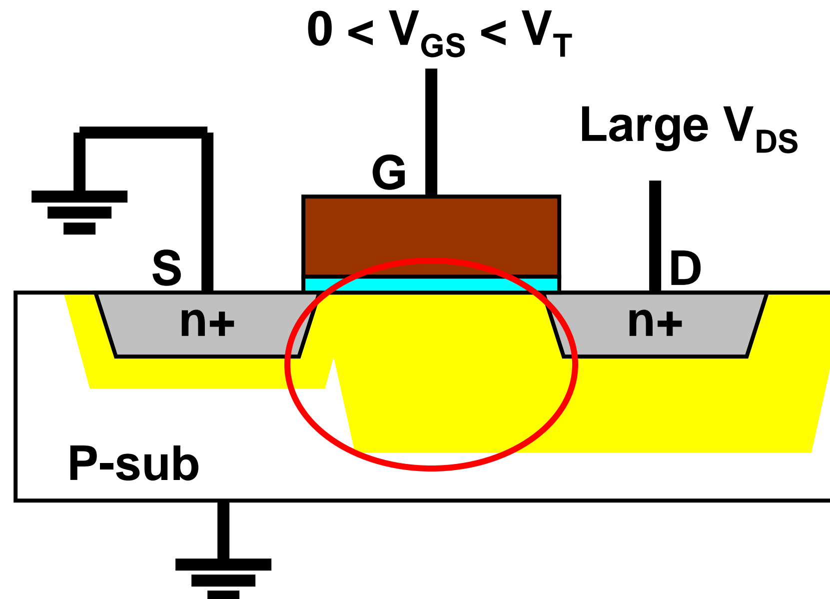




Drain-induced barrier lowering (for low L)

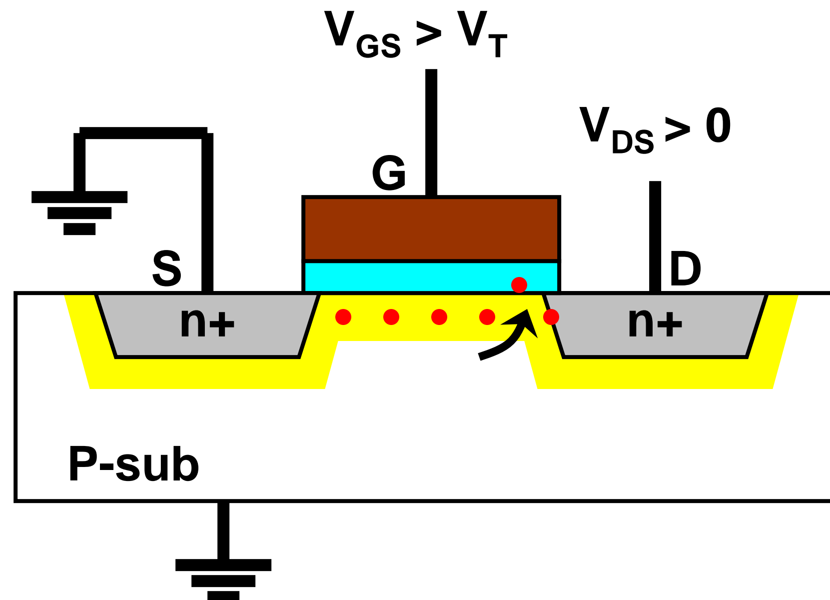
Punch Through

- For high enough Drain voltage, the depletion region of Drain reaches the Source
- When this happens carriers from the source will get injected into this depletion region and swept to the drain, resulting in a high current.
- This is called Punch Through.



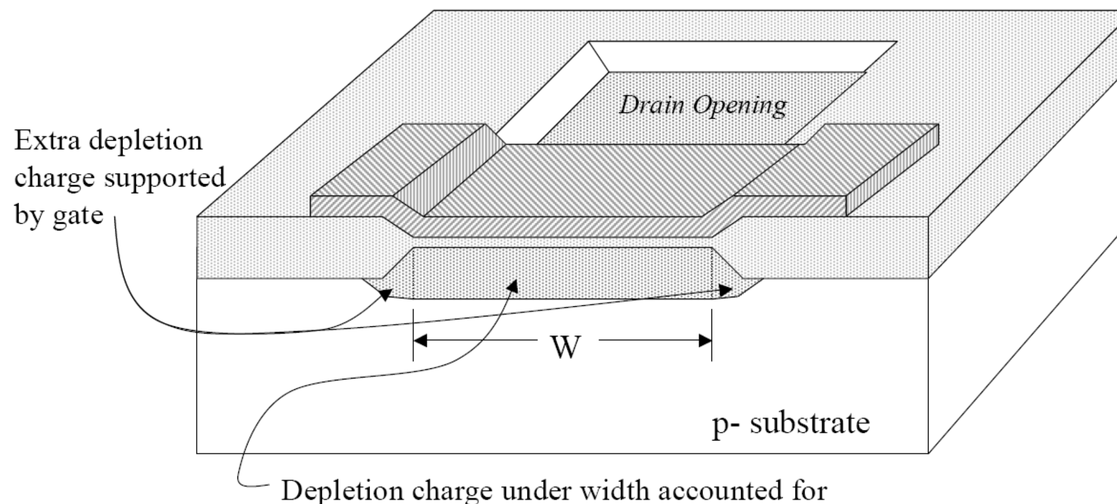
Hot Electrons

- Increase in lateral electric field causes an increasing velocity of electron
- The lateral field is strongest in the pinch off region where V_{DS}-V_{DSAT} is dropped over the distance of the channel that is pinched off.
- This high energy electrons (hot electrons) can leave channel and tunnel into the gate oxide
- Electrons trapped in the gate oxide creates a fixed charge that increases the threshold voltage of NMOS, which result is degradation of device transcoductance



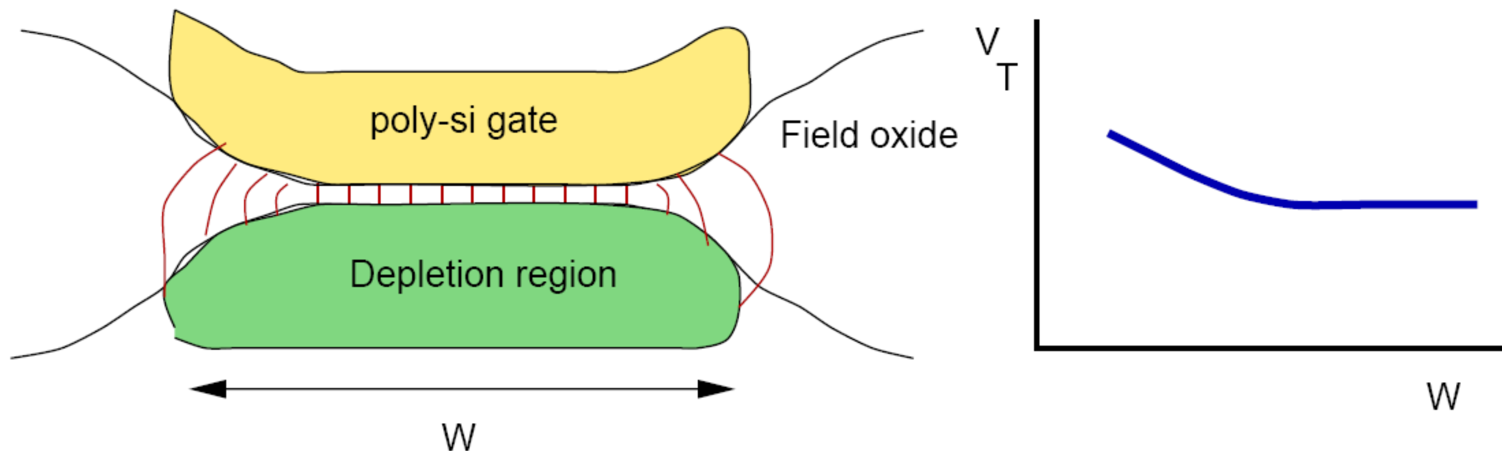
Narrow Channel Effect

- Depletion charge supported by gate is not just directly under the gate
- The field from the gate fringes and supports some depletion charge on each side of the channel
- This will impact the threshold voltage



Narrow Channel Effect

- Depletion region is not limited to the area just under the thin oxide
- The gate having to support this extra charge at the side of the channel effectively increases the VT
  - If W is large: part of the depletion region on the sides is small fraction and may be neglected
  - If W is small: gate also depletes the sides, hence significantly larger V_T



Summary

- **Static Parameters for Long Channel MOS**
  - Threshold voltage (impact of body bias)
  - Drain current (cutoff, linear, saturation)
- **Dynamic Parameters for MOS**
  - Overlap capacitances
  - Channel capacitances
  - Junction capacitances
- **Short Channel Effects**
  - Velocity saturation
  - Threshold voltage rolloff
  - etc.
- **Scaling Issues**